



Los capítulos del presente libro son co-creados por los autores al poner atención e intención en su participación en el Primer Congreso Internacional de Matemáticas y sus Aplicaciones (1CIMA). El 1CIMA emana de la fortuna de contar con el mejor comité organizador que ha designado la Academia de Matemáticas de la Facultad de Ciencias Físico Matemáticas, he aquí los resultados que promueve la riqueza matemática, trabajos tenaces que lograron sobreponerse a los inexorables jueces y fueron autorizados después de un arbitraje riguroso.

BUAP **FCFM** **CSMA**



textos
Científicos



Matemáticas y sus aplicaciones 6



Matemáticas y sus aplicaciones 6

Fernando Macías Romero
Editor

Matemáticas y sus aplicaciones 6

*Dedicado al Profesor Alejandro Illanes Mejía
por su 60 aniversario*



Benemérita Universidad Autónoma de Puebla
Facultad de Ciencias Físico Matemáticas

Fernando Macías Romero
Editor

BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

José Alfonso Esparza Ortiz

Rector

René Valdiviezo Sandoval

Secretario General

Ygnacio Martínez Laguna

Vicerrector de Investigación y Estudios de Posgrado

Flavio Marcelino Guzmán Sánchez

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

Ana María Dolores Huerta Jaramillo

Directora de Fomento Editorial

José Ramón Enrique Arrazola Ramírez

Director de la Facultad de Ciencias Físico Matemáticas

Primera edición, 2015

ISBN: 978-607-487-975-9

© Benemérita Universidad Autónoma de Puebla

Dirección de Fomento Editorial

2 Norte 1404, C.P. 72000

Puebla, Pue.

Teléfono y fax: 01 222 246 8559

Impreso y hecho en México

Printed and made in Mexico

Matemáticas y sus aplicaciones 6

Selección bajo arbitraje riguroso de algunos trabajos presentados en el Primer Congreso Internacional de Matemáticas y sus Aplicaciones 2014, FCFM, BUAP.

Editor

Fernando Macías Romero

Comité científico internacional

José Gerardo Ahuatzi Reyes (BUAP), María de la Paz Álvarez Scherzer (UNAM), Fernando Baltazar Larios (UNAM), Carolina Barriga Montoya (UNAM), Hugo Adán Cruz Suárez (BUAP), Gudelia Figueroa Preciado (UNISON), Judy Kennedy (LU, USA), Alejandro Illanes (UNAM), Fernando Macías Romero (BUAP), Javier Páez Cárdenas (UNAM), Ana Irene Ramírez Galarza (UNAM), José Gregorio Rodríguez Nieto (UNC, CO), Roberto Torres Hernández (UAQ).

Contenido

Presentación	1
Profesor Alejandro Illanes Mejía	
<i>Breve semblanza. El editor</i>	5
Ecuaciones diferenciales, sistemas dinámicos y modelación matemática	
Capítulo 1. Bifurcación de Andronov Hopf en la dinámica neuronal	13
<i>Lucía Cervantes Gómez, Ana Luisa González Pérez, Julio Erasto Poisot Macías, Leonardo Remedios Santiago</i>	
Capítulo 2. Estabilidad y bifurcación para algunos modelos Depredador-Presa	33
<i>Evodio Muñoz Aguirre, Francisco Sergio Salem Silva</i>	
Filosofía, Divulgación e Historia de las Matemáticas	
Capítulo 3. Una aproximación geométrica al cálculo diferencial	57
<i>José Juan Angoa Amador, Fernando Cocolletzi Adame</i>	
Capítulo 4. Los números perfectos y los números primos de Mersenne, a través del tiempo	107
<i>Aarón Aparicio Hernández</i>	
Geometría	
Capítulo 5. La magia de los cuadriláteros cíclicos	131
<i>Aarón Aparicio Hernández</i>	
Probabilidad y Estadística	

Capítulo 6. Análisis del Modelo Constante por tramos con Puntos de Cambio en la Función de Riesgo	153
<i>Yoanna Arenas Martínez, Hugo Adán Cruz Suárez, Hortensia Reyes Cervantes, Francisco Tajonar Sanabria</i>	
Capítulo 7. El modelo de Bandidos Armados, Programación Dinámica y la incorporación de nuevo personal	179
<i>Hugo Adán Cruz Suárez, Víctor Hugo Vázquez Guevara</i>	
Capítulo 8. Diseño y análisis de un experimento de uso eficiente de nitrógeno en maíz	195
<i>Bulmaro Juárez Hernández, Jenaro Reyes Matamoros, Thomas Lübberstedt</i>	
Topología	
Capítulo 9. Continuos localmente conexos sin hiperespacio único $C_n(X)$	215
<i>José Gerardo Ahuatzi Reyes, David Herrera Carrasco, Fernando Macías Romero</i>	
Capítulo 10. Acerca de continuos con n-ésimo producto simétrico único	241
<i>Mauricio Esteban Chacón Tirado, Luis Alberto Guerrero Méndez, David Herrera Carrasco, Fernando Macías Romero</i>	
Índice de autores	263

Presentación

Tenemos la fortuna de hacer un recorrido hacia un proceso creativo sin precedentes. Ha llegado el momento de compartir esta sabiduría, de invitar a todo el mundo a embarcarse en el navío que nos conduce hacia la Fuente de todo lo creado. Esta es la razón por la cual editamos el libro que tienen en sus manos. La felicidad que propone este libro por su divulgación, investigación e intercambio de ideas se debe a la generosidad de muchísimos matemáticos que participaron en el denominado Primer Congreso Internacional de Matemáticas y sus Aplicaciones (1CIMA), un esfuerzo profesional consolidado que ha permitido la participación de grandes personajes de diversas universidades, nacionales y extranjeras, tanto en el desarrollo del 1CIMA como en su memoria escrita, que es el presente libro. La base ha sido un comité organizador especializado, entusiasta y vigoroso emanado de la Academia de Matemáticas de la Facultad de Ciencias Físico Matemáticas de la BUAP. Este producto no es ni siquiera siete veces normal, es normal, de por lo menos nueve meses de trabajo constante. Por el amor a la matemática es que ha nacido este ejemplar que nos brinda la sabiduría necesaria para mostrarles parte de nuestros quehaceres cotidianos.

Los capítulos de este libro están agrupados por secciones de acuerdo al área temática en el 1CIMA. Dichos capítulos fueron sometidos a arbitraje riguroso.

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

Fernando Macías Romero
Editor

Profesor Alejandro Illanes Mejía

Breve semblanza

Alejandro Illanes Mejía nació en la ciudad de México el 13 de noviembre de 1955. Realizó sus estudios de Licenciatura en Matemáticas de 1975 a 1979 y estudió la Maestría en Ciencias de 1979 a 1981, todo en la Facultad de Ciencias de la UNAM. Su doctorado lo realizó en el Instituto de Matemáticas de la UNAM, obteniendo el grado el 30 de marzo de 1984, bajo la dirección del Dr. Adalberto García Máynez quien es considerado como el «Caballero de la Topología de Conjuntos en México». Alejandro se incorporó, como investigador, al Instituto de Matemáticas de la UNAM en abril de 1984.

Alejandro ha obtenido varios premios, distinciones y reconocimientos a lo largo de su impresionante trayectoria: la medalla «Gabino Barreda», en 1984; el Reconocimiento Distinción Universidad Nacional para jóvenes académicos en Docencia, en 1994; el premio UNAM correspondiente al área de Ciencias Exactas en Docencia, en 2005; entre otros. Actualmente es Investigador Titular «C» en el Instituto de Matemáticas de la UNAM y es «Nivel III» del Sistema Nacional de Investigadores.

Las Olimpiadas de Matemáticas, organizadas por la Sociedad Matemática Mexicana, son parte fundamental de sus esfuerzos de divulgación. Ha sido un gran impulsor de estos eventos desde todos los ámbitos; sea como organizador, entrenador, representante de delegaciones mexicanas o diseñador de problemas para los concursos. El objetivo —ha dicho Alejandro— es fomentar o afianzar vocaciones científicas, así como detectar jóvenes estudiosos, para desarrollar su talento.

Alejandro es evaluador de un sin número de proyectos de investigación, así como de artículos de investigación en muchísimas revistas de circulación internacional. Asimismo, ha sido director de 28 tesis de licenciatura, 5 tesis de maestría y 12 tesis de doctorado; hasta el momento tiene publicados 121 artículos de investigación.

A fines de 1984, Alejandro fue invitado al seminario en el que se reunían Isabel Puga, Luis Montejano, Adalberto García Máynez y Sergio Macías a leer el libro *Hyperspaces of Sets* de Sam. B. Nadler Jr. Desde el primer momento los hiperespacios lo conquistaron, sintió que había encontrado un tema que realmente lo satisfacía. Alejandro dice:

... hurgando en muchos temas de topología, ninguno me había llenado completamente; algunos era demasiado abstractos y en otros, se usaba demasiada estructura. En general, me gusta mucho la geometría pero también me gusta la flexibilidad que ofrece la topología y en los hiperespacios primero y la teoría de los continuos después encontré la combinación precisa de geometría y flexibilidad que me llena y se acomoda

mejor a mi manera de hacer matemáticas.

*Los hiperespacios no solo me gustaron sino que se portaron generosos conmigo, después de unas semanas de estudiarlos pude contestar uno de los problemas abiertos que contenía el libro de Nadler y desde entonces, lo que había sido amor a primera vista se convirtió en una relación llena de satisfacciones que, como se puede ver en la lista de publicaciones que tengo, ha perdurado hasta ahora.*¹

El mencionado seminario surgió por el interés que en ese entonces tenían Luis Montejano y Sergio Macías. De una manera u otra estos seminarios continúan hasta la fecha en el Instituto de Matemáticas de la UNAM. Por estos seminarios han desfilado personalidades, ahora de mucho peso.

A partir de 1985 hubo un crecimiento espectacular en el grupo de teoría de los continuos de la UNAM debido, entre otros, a los siguientes eventos:

1. La llegada a México de los profesores polacos Janusz J. Charatonik y su hijo Włodzimierz J. Charatonik.
2. El generoso apoyo del programa de PAPIIT de la DGAPA de la UNAM.
3. La invitación que le hizo Sam B. Nadler Jr. para escribir un libro acerca de hiperespacios (cuyo propósito fue presentar una monografía actualizada de los conocimientos sobre el tema de 1978 a 1998 que además sirviera como un texto para estudiar). Este libro lo publicó Marcel Dekker en 1999.
4. La incorporación a los seminarios de los profesores poblanos encabezados por Raúl Escobedo, al mismo tiempo se conjuntó en la Facultad de Ciencias y de otras universidades un excelente grupo de alumnos deseosos de estudiar la teoría de continuos.

Actualmente en la BUAP, UAEMéx, UNACH, UAQ, UMSNH y la UNISON se desarrollan con profundidad estos temas.

Alejandro encabeza anualmente un taller de investigación (entre junio y julio) en diferentes universidades de México. Para este taller, se conforma una lista de problemas abiertos previamente solicitados por él, después los organiza y comenta y establece una lista de problemas posibles por atacar. Ya en el Taller, se conforman grupos para atacar dichos problemas; en todos estos equipos, Alejandro tiene muchísimo que decir en cada problema abierto; de hecho, es la única persona que participa de esta manera. Esta creación de Alejandro es la escuela más grande del mundo en esta área de la topología.

¹Introducción, Hiperespacios de Conjuntos, Aportaciones Matemáticas de la Sociedad Matemática Mexicana, Serie Textos, vol. 28, 2004, pág. 1.

El apoyo que Alejandro ha brindado particularmente a los «poblanos» ha sido enorme pues tiene varios artículos publicados en coautoría con ellos. Alejandro ha apoyado a egresados y profesores de la BUAP como Dr. Raúl Escobedo Conde, Dra. María de J. López Toriz, Dr. Fernando Macías Romero, Dr. David Herrera Carrasco, Dr. Juan Carlos Macías Romero, Dr. Jesús Fernando Tenorio Arvide, Dr. Florencio Corona Vázquez, Dr. Hugo Villanueva Méndez, Dra. Alicia Santiago Santos, Dr. Mauricio Esteban Chacón Tirado y la Dra. Claudia Guadalupe Domínguez López. Ha participado en la BUAP con conferencias y como jurado de tesis de maestría y doctorado. En los primeros tres congresos internacionales en teoría de los continuos, los únicos que se han realizado hasta ahora y de los cuales la BUAP fue sede, participó como conferenciante. Alejandro ha afirmado también que un universitario completo es aquel que hace investigación y docencia. «Le doy casi el mismo tiempo a una cosa que a la otra», él ha dicho.

David Herrera Carrasco
Fernando Macías Romero



De pie, de izquierda a derecha: Luis Alberto Guerrero Méndez, Eduardo Jacobo Villegas, José Gerardo Ahuatzí Reyes, Vianey Córdova Salazar, Agustín Contreras Carreto, Javier Páez Cárdenas, María Elena, David Herrera Carrasco, Leonardo Espinoza Pérez, Mauricio Esteban Chacón Tirado, Javier Sánchez Martínez, Hugo Villanueva Méndez, Fernando Macías Romero, Lázaro Flores de Jesús, Manuel Ibarra Contreras, Armando Martínez García.

Sentados, de frente, de izquierda a derecha: Martha Leticia Ruiz Zavala, Raúl Escobedo Conde, Claudia Guadalupe Domínguez López, Jorge Marcos Martínez Montejano, Alejandro Illanes Mejía, Verónica Martínez de la Vega y Mansilla, María de Jesús López Toriz, Włodzimierz J. Charatonik, Ana María Reyes Crispín, Juana Onofre Cortés, Karen Clemente Robles.

Sentados, de espaldas, de izquierda a derecha: Patricia Pellicer Covarrubias, María Cristina, Ángel Tamariz Mascarúa, José Juan Angoa Amador, Fernando Velázquez Castillo.

Ecuaciones diferenciales, sistemas dinámicos y modelación matemática

Capítulo 1

Bifurcación de Andronov Hopf en la dinámica neuronal

Lucía Cervantes Gómez, Ana Luisa González Pérez, Julio Erasto Poiset Macías, Leonardo Remedios Santiago
Facultad de Ciencias Físico Matemáticas BUAP

Resumen

Se estudia la bifurcación de Andronov Hopf mediante la forma normal topológica. La bifurcación de Andronov Hopf es una de las bifurcaciones presentes en la respuesta electrofisiológica neuronal, el objetivo del trabajo es mostrar como se aplica esta teoría en la clasificación de la dinámica neuronal.

1 Introducción

El buen funcionamiento de las neuronas es fundamental para percibir el medio que nos rodea y para reaccionar a los estímulos del exterior. Este funcionamiento adecuado está relacionado con la capacidad de las células para transmitir los impulsos eléctricos con velocidad constante y sin distorsión.

A mediados de siglo, el trabajo de Hodgkin y Huxley describió el proceso de transmisión eléctrica a lo largo del axón gigante de calamar, el cual tiene un diámetro de .5 a 1 mm. y por estas dimensiones fue posible realizar sus experimentos. Uno de los resultados importantes de estos estudios es que el proceso anteriormente explicado se puede ver como un sistema dinámico. Ellos fueron los pioneros en realizar un estudio de las bifurcaciones en la dinámica neuronal antes de que la teoría de bifurcaciones se hubiera desarrollado. Sus estudios, publicados en 1952 se basaron en la estimulación del axón con impulsos de varias amplitudes e identificaron tres clases de respuestas de la excitabilidad neuronal.

Durante la segunda mitad del siglo *XX*, después de los trabajos de Hodgkin y Huxley empezó a aumentar la cantidad de experimentos y resultados sobre la respuesta eléctrica, teniendo un crecimiento exponencial en las últimas décadas, surgiendo una gran variedad de patrones distintos; entre otros fenómenos se observó que células con corrientes similares exhibían dinámicas completamente diferentes, con lo cual la clasificación de Hodgkin ya no era suficiente. El artículo de Rinzel y Ermentrout publicado en 1989 resumió el enfoque de varios trabajos de matemáticos

y planteó las bases de la nueva visión de la utilización de conceptos de la teoría cualitativa y de la teoría de bifurcaciones para entender, predecir e interpretar estos fenómenos de la dinámica neuronal.

En general las neuronas son excitables porque están cerca de una bifurcación del reposo al estado de disparo, de esta forma, la bifurcación determina las propiedades excitables de la neurona, el tipo depende de la electrofisiología de la misma. Una observación sorprendente es que puede haber millones de mecanismos electrofisiológicos de excitabilidad y disparo, pero solo han encontrado cuatro tipos diferentes de bifurcación del equilibrio en sistemas de dimensión 2, a las cuales se puede someter un sistema y una vez que hayamos entendido estas bifurcaciones se tienen las bases para comprender las bifurcaciones de dimensiones mayores.

Las cuatro bifurcaciones son:

I. *Bifurcación Silla-Nodo.*

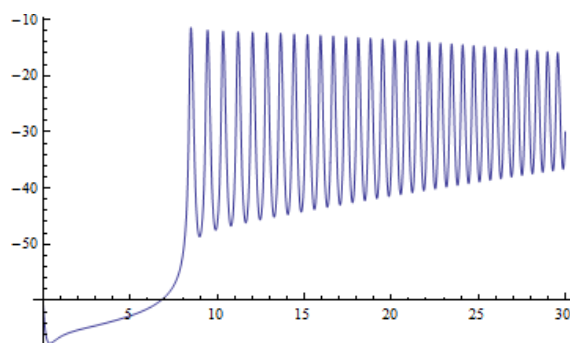


Figura 1: Comportamiento del voltaje en la bifurcación silla nodo

II. *Bifurcación silla nodo en un círculo invariante.*

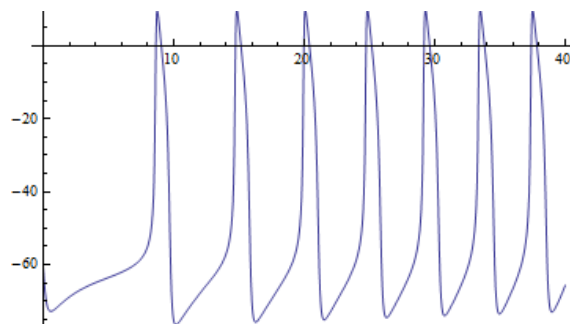


Figura 2: Comportamiento del voltaje en la bifurcación silla nodo en un círculo invariante

III. *Bifurcación Subcrítica de Andronov Hopf.*

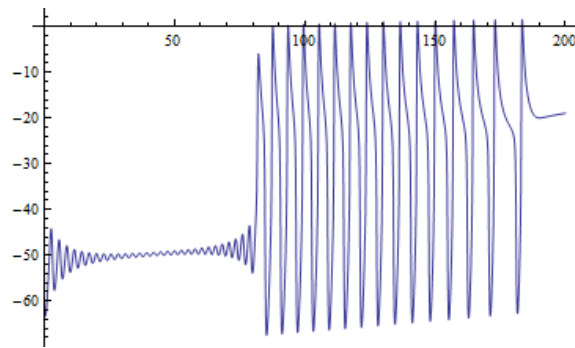


Figura 3: Comportamiento del voltaje en la bifurcación subcrítica de Andronov Hopf

IV. Bifurcación de Andronov Hopf Hipercrítica.

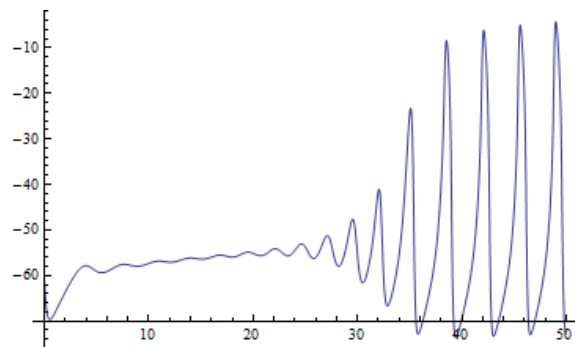


Figura 4: Comportamiento del voltaje en la bifurcación hipercrítica de Andronov Hopf

Los sistemas que se someten a las bifurcaciones de Andronov Hopf (hipercrítica o subcrítica) presentan pequeñas oscilaciones del potencial de membrana mientras que los sistemas que presentan las bifurcaciones de silla-nodo (ya sea en un círculo invariante o no) no muestran estas oscilaciones. La existencia de oscilaciones de pequeña amplitud crea la posibilidad de resonancia a la frecuencia de los impulsos entrantes y otras características interesantes. Vamos a referirnos a las neuronas con pequeñas oscilaciones subumbrales como *resonadores* y a las que no tienen esta propiedad como *integradores*. De esta misma forma vamos a referirnos a las neuronas que exhiben la coexistencia de los estados de reposo y de disparo al menos cerca de la transición del reposo al estado de disparo como *biestables* y a las que no presentan esta característica como *monoestables*.

En este trabajo nos centraremos en el estudio de la bifurcación de Andronov Hopf.

2 Equivalencias y bifurcaciones en sistemas dinámicos

Empezaremos recordando la definición de sistema dinámico.

Definición 2.1. Un *sistema dinámico* es una tripleta $\{T, X, \varphi^t\}$ donde T es un conjunto que expresa el tiempo (discreto o continuo) X es el espacio de estados y $\varphi^t : X \rightarrow X$ es una familia de operadores de evolución parametrizados por $t \in T$ que satisface las siguientes propiedades:

- I. $\varphi^0 = id$
- II. $\varphi^{t+s} = \varphi^t \circ \varphi^s$, donde $t, s \in T$

Debido a que podría ser que $\varphi^t x$ no estuviera definido para toda pareja $(x, t) \in X \times T$, se exige que se cumpla II cuando ambos miembros de la ecuación estén definidos.

Para comparar los diferentes estados del sistema dinámico, es necesario que el conjunto X satisfaga algunas propiedades; una caracterización general con muchas aplicaciones es que X sea un espacio métrico; por otra parte, el conjunto que expresa el tiempo es algún subconjunto $T \subset \mathbb{R}$, por ejemplo puede ser \mathbb{Z} en el caso de los sistemas dinámicos de tiempo discreto o \mathbb{R} o algún subintervalo de \mathbb{R} en el caso de sistemas dinámicos de tiempo continuo. Para conocer el estado x_t del sistema al tiempo t , es necesario conocer el estado inicial o condición inicial x_0 . Los sistemas dinámicos con un operador de evolución definido para $t \geq 0$ y para $t < 0$ se llaman *invertibles*, y en tales sistemas la condición inicial x_0 no solo define los estados futuros del sistema sino también su comportamiento pasado. También es posible que $\varphi^t x_0$ esté definido localmente en el tiempo, es decir, para $0 \leq t < t_0$ donde t_0 depende de $x_0 \in X$. En este trabajo vamos a tratar solo con los sistemas de tiempo continuo.

Una forma muy común mediante la cual podemos generar un sistema dinámico de tiempo continuo es a partir de ecuaciones diferenciales. Supongamos que el espacio de estados de un sistema es $X = \mathbb{R}^n$ con coordenadas (x_1, x_2, \dots, x_n) . Muy a menudo la ley de evolución del sistema está dada implícitamente, en términos de las velocidades \dot{x}_i , como funciones de las coordenadas (x_1, x_2, \dots, x_n) :

$$\dot{x}_i = f_i(x_1, x_2, \dots, x_n), i = 1, 2, \dots, n,$$

o en forma vectorial:

$$\dot{x} = f(x), \tag{1}$$

donde la función vectorial $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es diferenciable (suave). La función en el lado derecho de (1) se llama *campo vectorial*, ya que asigna un vector $f(x)$ a cada punto x .

Una de las tareas más importantes es el estudio de las características cualitativas de los sistemas dinámicos y su comportamiento, para poder realizar una clasificación y comparar los sistemas. La comparación de objetos se basa en una relación de equivalencia, lo cual permitirá definir clases de equivalencia entre los objetos.

Definición 2.2. Un sistema dinámico $\{T_1, \mathbb{R}^n, \varphi^t\}$ es *topológicamente equivalente* a un sistema dinámico $\{T_2, \mathbb{R}^n, \psi^t\}$ si existe un homeomorfismo $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$ que envía órbitas del primer sistema a órbitas del segundo sistema, preservando la dirección de las órbitas determinada por el avance del tiempo en el sentido que si una órbita Γ del primer sistema se dirige de x_1 a x_2 , entonces sus imágenes se dirigen a una órbita Υ del segundo sistema de $h(x_1)$ a $h(x_2)$.

A veces el estudio de los sistemas dinámicos se realiza de manera local, es decir, no en todo el espacio \mathbb{R}^n , pero sí en una región $U \subset \mathbb{R}^n$; tal región puede ser una vecindad del punto de equilibrio o de un ciclo. De esta manera, las equivalencias anteriormente definidas se pueden modificar para hacer un estudio local, por ejemplo la siguiente modificación de la definición (2.2) es muy útil para un estudio local.

Definición 2.3. Un sistema dinámico $\{T, \mathbb{R}^n, \varphi^t\}$ es *localmente equivalente topológicamente* cerca de un punto de equilibrio x_0 a un sistema dinámico $\{T, \mathbb{R}^n, \psi^t\}$ cerca de un punto de equilibrio y_0 si existe un homeomorfismo $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$ que:

- I. Está definido en una pequeña vecindad $U \subset \mathbb{R}^n$ de x_0
- II. Satisface que $y_0 = h(x_0)$
- III. Envía órbitas del primer sistema en U a órbitas del segundo sistema en $V = f(U) \subset \mathbb{R}^n$ preservando su dirección.

Ahora consideremos un sistema dinámico que depende de parámetros, que escribiremos de la siguiente manera:

$$\dot{x} = f(x, \alpha),$$

donde $x \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}^m$ representan las variables fase y los parámetros respectivamente; consideremos el retrato fase del sistema cuando el parámetro varía. Hay dos posibilidades: que el sistema permanezca topológicamente equivalente al sistema original o no.

Definición 2.4. La aparición de retratos fase no equivalentes topológicamente bajo la variación de parámetros se llama *bifurcación*.

Definición 2.5. Un *diagrama de bifurcación* de un sistema dinámico es una estratificación de su espacio de parámetros inducida por la equivalencia topológica junto con los retratos fase representativos de cada estrato.

Definición 2.6. La *codimensión* de una bifurcación en un sistema $\dot{x} = f(x)$ es la diferencia entre la dimensión del espacio de parámetros y la dimensión de la correspondiente frontera de bifurcación.

Equivalentemente, la codimensión es el número de condiciones independientes que determinan la bifurcación.

La siguiente definición es equivalente a la definición 2.2 con las modificaciones necesarias debido a la dependencia del parámetro.

Definición 2.7. Un sistema dinámico

$$\dot{x} = f(x, \alpha), x \in \mathbb{R}^n, \alpha \in \mathbb{R}^m \quad (2)$$

es *equivalente topológicamente* a un sistema dinámico

$$\dot{y} = g(y, \beta), y \in \mathbb{R}^n, \beta \in \mathbb{R}^m \quad (3)$$

si:

- I. Existe un homeomorfismo del espacio de parámetros $p : \mathbb{R}^m \rightarrow \mathbb{R}^m$ tal que $\beta = p(\alpha)$.
- II. Hay un homeomorfismo dependiente del parámetro del espacio fase $h_\alpha : \mathbb{R}^n \rightarrow \mathbb{R}^n, y = h_\alpha(x)$ que envía órbitas del sistema (2) a valores del parámetro α en órbitas del sistema (3) a valores del parámetro $\beta = p(\alpha)$ preservando la dirección.

Al igual que la definición 2.2, la definición 2.7 puede modificarse para realizar un estudio local.

Definición 2.8. Dos sistemas (2) y (3) son *localmente equivalentes topológicamente* cerca del origen si existe una función $(x, \alpha) \rightarrow (h_\alpha(x), p(\alpha))$ definida en una pequeña vecindad de $(x, \alpha) = (0, 0)$ en el producto directo $\mathbb{R}^n \times \mathbb{R}^m$ y tal que:

- I. $p : \mathbb{R}^m \rightarrow \mathbb{R}^m$ es un homeomorfismo definido en una pequeña vecindad de $\alpha = 0, p(0) = 0$.
- II. $h_\alpha : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es un homeomorfismo dependiente del parámetro definido en una pequeña vecindad U_α de $x_0, h_0(0) = 0$ y envía órbitas de (2) en U_α a órbitas de (3) en $h_\alpha(U_\alpha)$ preservando la dirección.

Sea

$$\dot{\xi} = g(\xi, \beta; \sigma), \quad \xi \in \mathbb{R}^n, \quad \beta \in \mathbb{R}^k, \quad \sigma \in \mathbb{R}^l \quad (4)$$

un sistema simple (polinomial en ξ_i) que tiene en $\beta = 0$ un punto de equilibrio $\xi = 0$ que satisface k condiciones de bifurcación que determinan una bifurcación de codimensión k para este equilibrio. Aquí σ es un vector de los coeficientes $\sigma_i, i = 1, 2, \dots, l$, de los polinomios involucrados en (4). Junto con el sistema (4) vamos a considerar un sistema

$$\dot{x} = f(x, \alpha), \quad x \in \mathbb{R}^n, \quad \alpha \in \mathbb{R}^k, \quad (5)$$

el cual tiene en $\alpha = 0$ un equilibrio $x = 0$.

Definición 2.9 (forma normal topológica). El sistema (4) es una *forma normal topológica* para la bifurcación si cualquier sistema genérico (5) con equilibrio $x = 0$ que satisface las mismas condiciones de bifurcación en $\alpha = 0$ es localmente equivalente topológicamente cerca del origen a (4) para algunos valores de los coeficientes σ_i .

3 Bifurcación de Andronov Hopf

Consideremos el siguiente sistema de ecuaciones diferenciales dependientes de un parámetro

$$\begin{cases} \dot{x}_1 = \alpha x_1 - x_2 - x_1(x_1^2 + x_2^2), \\ \dot{x}_2 = x_1 + \alpha x_2 - x_2(x_1^2 + x_2^2) \end{cases} \quad (6)$$

el sistema tiene un punto de equilibrio $x_1 = x_2 = 0$ para todo α con matriz jacobiana,

$$A = \begin{pmatrix} \alpha & -1 \\ 1 & \alpha \end{pmatrix},$$

la cual tiene valores propios $\lambda_{1,2} = \alpha \pm i$. La representación polar de este sistema es:

$$\begin{cases} \dot{p} = p(\alpha - p^2), \\ \dot{\varphi} = 1 \end{cases} \quad (7)$$

Las bifurcaciones del sistema se pueden estudiar fácilmente en la forma polar, ya que las ecuaciones para p y φ están desacopladas. La primera ecuación (definida solo para $p \geq 0$) solo tiene un equilibrio $p = 0$ para todo α . El equilibrio es linealmente estable para $\alpha < 0$; se mantiene estable pero no de manera lineal en $\alpha = 0$ y para $\alpha > 0$ se vuelve inestable. Además hay un punto de equilibrio $p_0(\alpha) = \sqrt{\alpha}$ que es estable para $\alpha > 0$. La segunda ecuación describe una rotación con velocidad

constante. El equilibrio $p = 0$ es un foco estable para $\alpha < 0$ y un foco inestable para $\alpha > 0$; en el valor crítico del parámetro $\alpha = 0$ es estable no lineal y es topológicamente equivalente al foco. A veces se denomina foco débilmente atractor. Este equilibrio está rodeado por una órbita periódica única y estable; este ciclo es un círculo de radio $\sqrt{\alpha}$. Todas las órbitas con condición inicial fuera o dentro del ciclo excepto en el origen tienden a este cuando $t \rightarrow \infty$.

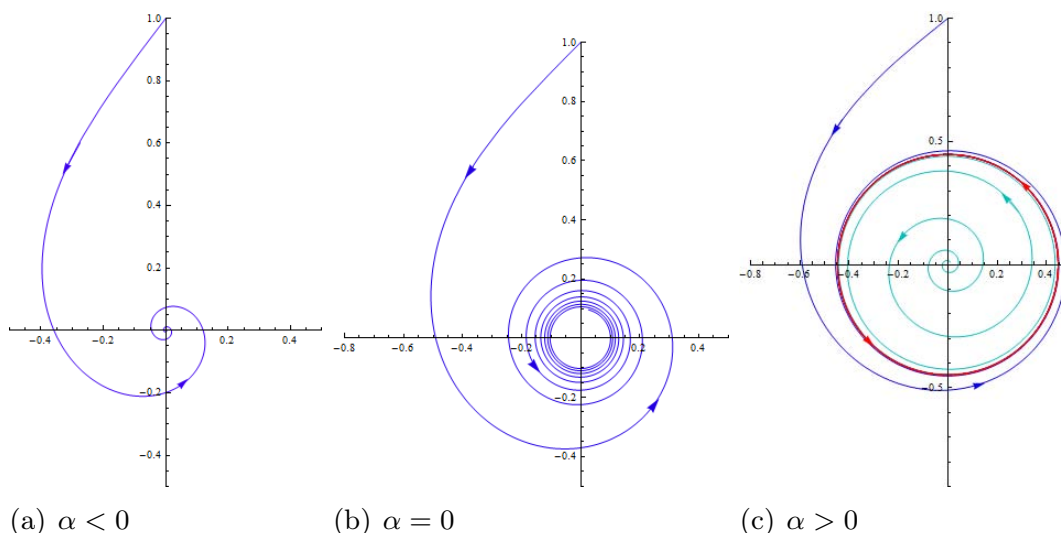


Figura 5: Bifurcación hipercrítica de Andronov Hopf.

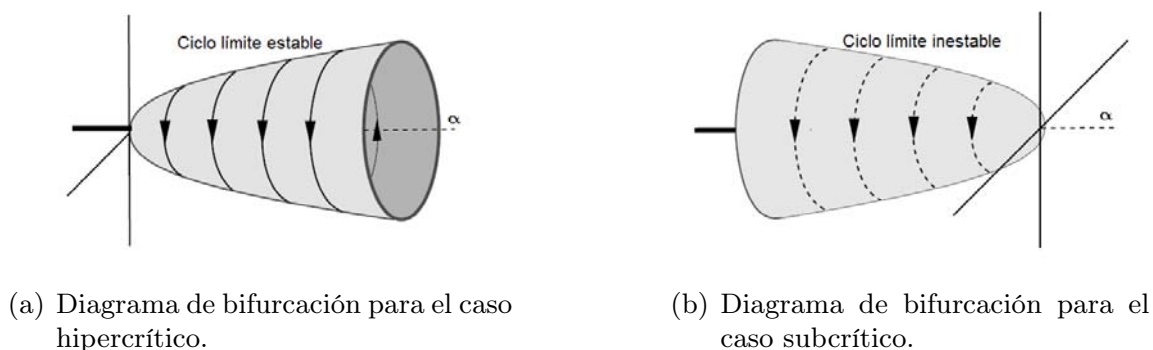


Figura 6: Diagramas de bifurcación [4].

Ahora consideremos el sistema con signo contrario en los términos no lineales

$$\begin{cases} \dot{x}_1 = \alpha x_1 - x_2 + x_1 (x_1^2 + x_2^2), \\ \dot{x}_2 = x_1 + \alpha x_2 + x_2 (x_1^2 + x_2^2) \end{cases} \quad (8)$$

el cual, bajo un cambio de variable similar al del caso anterior tiene la siguiente forma compleja:

$$\dot{z} = (\alpha + i)z + z|z|^2.$$

El análisis de este sistema es similar al del caso anterior; este sistema tiene un ciclo límite inestable para $\alpha < 0$ que desaparece cuando α pasa de negativo a positivo. El origen es estable para $\alpha < 0$ y es inestable para $\alpha > 0$; a diferencia del caso anterior, el equilibrio es inestable de forma no lineal en $\alpha = 0$. Esta es la bifurcación de Andronov Hopf. Esta bifurcación se llama hipercrítica cuando el ciclo existe para valores positivos de α , es decir, después de la bifurcación (*i.e.*, después de $\alpha = 0$) y es subcrítica cuando el ciclo existe antes de la bifurcación.

Vamos a agregar algunos términos de orden superior y a escribir el sistema (6) en forma de vectores

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} \alpha & -1 \\ 1 & \alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - (x_1^2 + x_2^2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + O(\|x^4\|), \quad (9)$$

donde $O(\|x^4\|)$ son términos que pueden depender suavemente de α .

Lema 3.1. *El sistema (9) es localmente equivalente topológicamente cerca del origen al sistema (6).*

Consideremos un sistema

$$\dot{x} = f(x, \alpha), \quad x = (x_1, x_2)^T \in \mathbb{R}^2, \quad \alpha \in \mathbb{R}$$

con f una función suave, y que tiene en $\alpha = 0$ el equilibrio $x = 0$ con valores propios $\lambda_{1,2} = \pm iw_0$, $w_0 > 0$. Por el teorema de la función implícita, el sistema tiene un único equilibrio $x_0(\alpha)$ en alguna vecindad del origen para todo $|\alpha|$ suficientemente pequeño, ya que $\lambda = 0$ no es un valor propio de la matriz jacobiana. Podemos realizar un cambio de coordenadas para trasladar este equilibrio al origen, por lo tanto podemos asumir que $x = 0$ es el punto de equilibrio para $|\alpha|$ suficientemente pequeño. Así, el sistema puede escribirse como:

$$\dot{x} = A(\alpha)x + F(x, \alpha), \quad (10)$$

donde F es una función vectorial cuyas componentes $F_{1,2}$ tienen expansiones de Taylor en x empezando con al menos los términos cuadráticos, $F = O(\|x^2\|)$. La matriz jacobiana $A(\alpha)$ puede escribirse como:

$$A(\alpha) = \begin{pmatrix} a(\alpha) & b(\alpha) \\ c(\alpha) & d(\alpha) \end{pmatrix}$$

con funciones suaves de α como sus elementos. Sus valores propios tienen la siguiente representación:

$$\lambda_1(\alpha) = \lambda(\alpha), \quad \lambda_2(\alpha) = \overline{\lambda(\alpha)},$$

donde $\lambda(\alpha) = \mu(\alpha) + iw(\alpha)$, $\mu(0) = 0$ y $w(0) = w_0 > 0$.

Lema 3.2. *Con la introducción de una variable compleja z , el sistema (10) puede ser escrito para $|\alpha|$ suficientemente pequeño como una sola ecuación*

$$\dot{z} = \lambda(\alpha)z + g(z, \bar{z}, \alpha), \quad (11)$$

donde $g = O(|z|^2)$ es una función suave de (z, \bar{z}, α)

Demostración. Sea $q(\alpha) \in \mathbb{C}^2$ correspondiente al valor propio $\lambda(\alpha)$, i.e., $A(\alpha)q(\alpha) = \lambda(\alpha)q(\alpha)$, y consideremos $p(\alpha) \in \mathbb{C}^2$ un vector propio de la matriz transpuesta, i.e., $A^T(\alpha)p(\alpha) = \overline{\lambda(\alpha)}p(\alpha)$. Supongamos sin pérdida de generalidad que $\langle q(\alpha), p(\alpha) \rangle = 1$, donde $\langle \cdot, \cdot \rangle$ es el producto interno de \mathbb{C}^2 . Cualquier vector $x \in \mathbb{R}^2$ puede ser representado de manera única para cualquier α pequeño como

$$x = zq(\alpha) + \bar{z}\bar{q}(\alpha) \quad (12)$$

para algún complejo z , siempre que se especifiquen los vectores propios. Podemos determinar z de la siguiente manera:

$$z = \langle x, p(\alpha) \rangle$$

en efecto, tomando producto escalar en ambos lados de (12) tenemos

$$\begin{aligned} \langle x, p(\alpha) \rangle &= \langle zq(\alpha) + \bar{z}\bar{q}(\alpha), p(\alpha) \rangle \\ &= z\langle q(\alpha), p(\alpha) \rangle + \bar{z}\langle \bar{q}(\alpha), p(\alpha) \rangle \\ &= z + \bar{z}\langle \bar{q}(\alpha), p(\alpha) \rangle \end{aligned}$$

Solo resta ver que $\langle \bar{q}(\alpha), p(\alpha) \rangle = 0$, nótese que:

$$\begin{aligned} \langle \bar{q}(\alpha), p(\alpha) \rangle &= \left\langle \frac{1}{\lambda} A\bar{q}(\alpha), p(\alpha) \right\rangle \\ &= \frac{1}{\lambda} \langle A\bar{q}(\alpha), p(\alpha) \rangle \\ &= \frac{1}{\lambda} \langle \bar{q}(\alpha), A^T p(\alpha) \rangle \\ &= \frac{1}{\lambda} \langle \bar{q}(\alpha), \bar{\lambda}p(\alpha) \rangle \\ &= \frac{\lambda}{\lambda} \langle \bar{q}(\alpha), p(\alpha) \rangle \end{aligned}$$

y por lo tanto tenemos que

$$\left(1 - \frac{\lambda}{\bar{\lambda}}\right) \langle \bar{q}(\alpha), p(\alpha) \rangle = 0$$

dado que $\bar{\lambda} \neq \lambda$ para $\|\alpha\|$ suficientemente pequeño, entonces $\langle \bar{q}(\alpha), p(\alpha) \rangle = 0$. Así, tomando el producto interno en ambos lados de (10) tenemos que

$$\begin{aligned} \dot{z} &= \langle \dot{x}, p(\alpha) \rangle = \langle A(\alpha)x + F(x, \alpha), p(\alpha) \rangle \\ &= \langle A(\alpha)x, p(\alpha) \rangle + \langle F(x, \alpha), p(\alpha) \rangle \end{aligned}$$

haciendo el cambio (12) en $\langle A(\alpha)x, p(\alpha) \rangle$ tenemos que

$$\dot{z} = \lambda(\alpha)z + \langle F(x, \alpha), p(\alpha) \rangle$$

y finalmente, si hacemos $g(z, \bar{z}, \alpha) = \langle F(x, \alpha), p(\alpha) \rangle$, entonces

$$\dot{z} = \lambda(\alpha)z + g(z, \bar{z}, \alpha). \quad \square$$

No es necesario que g sea una función analítica de z , (*i.e.*, independiente de \bar{z}), así que vamos a escribir g en serie de Taylor en 2 variables complejas

$$g(z, \bar{z}, \alpha) = \sum_{k+l \geq 2} \frac{1}{k!l!} g_{kl}(\alpha) z^k \bar{z}^l,$$

donde $g_{kl} = \frac{\partial^{k+l}}{\partial z^k \partial \bar{z}^l} \langle F(zq(\alpha) + \bar{z}\bar{q}(\alpha), \alpha), p(\alpha) \rangle |_{z=0}$

Supongamos que en $\alpha = 0$ la función $F(x, \alpha)$ se puede representar como

$$F(x, 0) = \frac{1}{2}B(x, x) + \frac{1}{6}C(x, x, x) + O(\|x\|^4),$$

donde $B(x, y)$ y $C(x, y, u)$ son funciones vectoriales simétricas y multilineales de $x, y, u \in \mathbb{R}^2$, desarrollando tenemos

$$B_i(x, y) = \sum_{j,k=1}^2 \frac{\partial^2 F_i(\xi, 0)}{\partial \xi_j \partial \xi_k} |_{\xi=0} x_j y_k \quad i = 1, 2$$

y

$$C_i(x, y, u) = \sum_{j,k,l=1}^2 \frac{\partial^3 F_i(\xi, 0)}{\partial \xi_j \partial \xi_k \partial \xi_l} |_{\xi=0} x_j y_k u_l \quad i = 1, 2$$

Así, realizando el cambio (12) en $B(x, y)$ tenemos que

$$B(zq(\alpha) + \bar{z}\bar{q}(\alpha), zq(\alpha) + \bar{z}\bar{q}(\alpha)) = z^2 B(q, q) + 2z\bar{z}B(q, \bar{q}) + \bar{z}^2 B(\bar{q}, \bar{q})$$

donde $q = q(0)$, por lo tanto, los coeficientes de Taylor $g_{k,l}$, $k+l = 2$ de los términos cuadráticos de $g(z, \bar{z}, \alpha)$ podemos expresarlos de la siguiente manera:

$$g_{2,0} = \langle B(q, q), p \rangle, \quad g_{1,1} = \langle B(q, \bar{q}), p \rangle, \quad g_{0,2} = \langle B(\bar{q}, \bar{q}), p \rangle$$

y similarmente,

$$g_{2,1} = \langle C(q, q, \bar{q}), p \rangle.$$

Ahora estudiaremos los términos cuadráticos y cúbicos de $g(z, \bar{z}, \alpha)$ mediante el siguiente lema:

Lema 3.3 (Forma normal de Poincaré para la bifurcación de Hopf). *La ecuación*

$$\dot{z} = \lambda z + \sum_{2 \leq k+l \leq 3} \frac{1}{k!l!} g_{k,l} z^k \bar{z}^l + O(|z|^4),$$

donde $\lambda = \lambda(\alpha) = \mu(\alpha) + i\omega(\alpha)$, $\mu(0) = 0$, $\omega(0) = \omega_0 > 0$ y $g_{i,j} = g_{i,j}(\alpha)$ se puede transformar por un cambio de coordenadas invertible y dependiente del parámetro

$$z = w + \frac{h_{2,0}}{2} w^2 + h_{1,1} w \bar{w} + \frac{h_{0,2}}{2} \bar{w}^2 + \frac{h_{3,0}}{6} w^3 + \frac{h_{1,2}}{2} w \bar{w}^2 + \frac{h_{0,3}}{6} \bar{w}^3 \quad (13)$$

para todo $|\alpha|$ suficientemente pequeño en una ecuación con solo el término cúbico resonante

$$\dot{w} = \lambda w + c_1 w^2 \bar{w} + O(|w|^4), \quad (14)$$

donde $c_1 = c_1(\alpha)$.

Lema 3.4. *Consideremos la ecuación*

$$\frac{dw}{dt} = (\mu(\alpha) + i\omega(\alpha)) w + c_1(\alpha) w |w|^2 + O(|w|^4),$$

donde $\mu(0) = 0$ y $\omega(0) = \omega_0 > 0$. Supongamos que $\mu'(0) \neq 0$ y $\operatorname{Re} c_1(0) \neq 0$ mediante una transformación lineal de coordenadas dependiente del parámetro, un rescalamiento del tiempo y una reparametrización del tiempo no lineal se puede transformar en una ecuación de la forma

$$\frac{du}{d\theta} = (\beta + i) u + s u |u|^2 + O(|u|^4),$$

donde u es una nueva coordenada compleja y θ, β son el nuevo tiempo y parámetro respectivamente y $s = \operatorname{signo} \operatorname{Re} c_1(0) = \pm 1$

Demostración.

Paso 1 (Escalamiento lineal del tiempo). Introduciendo la nueva variable del tiempo $\tau = \omega(\alpha)t$ la dirección del tiempo se preserva, ya que $\omega(\alpha) > 0$ para todo $|\alpha|$ suficientemente pequeño, así

$$\frac{dw}{d\tau} = (\beta + i)w + d_1(\beta)w|w|^2 + O(|w|^4),$$

donde $\beta = \beta(\alpha) = \frac{\mu(\alpha)}{\omega(\alpha)}$, $d_1(\beta) = \frac{c_1(\beta(\alpha))}{\omega(\beta(\alpha))}$. Podemos considerar a β como un nuevo parámetro porque:

$$\beta(0) = 0 \text{ y } \beta'(0) = \frac{\mu'(0)}{\omega(0)} \neq 0$$

y por lo tanto el teorema de la función inversa garantiza la existencia local y suavidad de α como función de β . Nótese que $d_1(\beta)$ es complejo.

Paso 2 (Reparametrización del tiempo no lineal). Cambiando la parametrización del tiempo a lo largo de las órbitas mediante la introducción del nuevo tiempo $\theta = \theta(\tau, \beta)$ donde

$$d\theta = (1 + e_1(\beta)|w|^2) d\tau$$

con $e_1 = \text{Im } d_1(\beta)$. El cambio del tiempo es una transformación parecida a la identidad en una pequeña vecindad del origen. Usando esta nueva definición del tiempo obtenemos

$$\frac{dw}{d\theta} = (\beta + i)w + l_1(\beta)w|w|^2 + O(|w|^4),$$

donde $l_1(\beta) = \text{Re } d_1(\beta) - \beta e_1(\beta)$ es real y $l_1(0) = \frac{\text{Re } c_1(0)}{\omega(0)}$.

Paso 3 (Escalamiento lineal de coordenadas). Finalmente vamos a introducir una nueva variable compleja u :

$$w = \frac{u}{\sqrt{|l_1(\beta)|}}$$

lo cual es posible debido a que $\text{Re } c_1(0) \neq 0$ y así $l_1(0) \neq 0$, de este modo la ecuación toma la forma requerida

$$\begin{aligned} \frac{du}{d\theta} &= (\beta + i)u + \frac{l_1(\beta)}{|l_1(\beta)|}u|u|^2 + O(|u|^4) \\ &= (\beta + i)u + su|u|^2 + O(|u|^4) \end{aligned}$$

con $s = \text{signo } l_1(0) = \text{signo } \text{Re } c_1(0)$. □

Definición 3.5. La función real $l_1(\beta)$ es el *primer coeficiente de Liapunov*.

Se sigue que el primer coeficiente de Liapunov en $\beta = 0$ se puede calcular mediante la fórmula

$$l_1(0) = \frac{1}{2w_0} \operatorname{Re}(ig_{2,0}g_{1,1} + w_0g_{2,1}). \quad (15)$$

Teorema 3.6. *Supongamos que un sistema bidimensional*

$$\frac{dx}{dt} = f(x, \alpha), \quad x \in \mathbb{R}^2, \quad \alpha \in \mathbb{R}$$

con f suave tiene $\forall |\alpha|$ suficientemente pequeño el equilibrio $x = 0$ con valores propios

$$\lambda_{1,2}(\alpha) = \mu(\alpha) \pm iw(\alpha)$$

donde $w(0) = w_0 > 0$, $\mu(0) = 0$ si satisface las siguientes condiciones:

(B.1) $l_1(0) \neq 0$, donde l_1 es el primer coeficiente de Liapunov.

(B.2) $\mu'(0) \neq 0$.

Mediante un cambio de coordenadas, de parámetros y una reparametrización del tiempo que transforma el sistema en

$$\frac{d}{d\tau} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \alpha & -1 \\ 1 & \alpha \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \pm (y_1^2 + y_2^2) \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + O(\|y\|^4).$$

Usando el lema (3.1) podemos eliminar los términos de $O(\|y\|^4)$ y finalmente llegar al siguiente resultado general.

Teorema 3.7 (Forma normal topológica para la bifurcación de Hopf). *Cualquier sistema genérico bidimensional que depende de un parámetro*

$$\dot{x} = f(x, \alpha)$$

que tiene en $\alpha = 0$ el equilibrio $x = 0$ con valores propios

$$\lambda_{1,2}(0) = \pm iw_0, \quad w_0 > 0$$

es localmente equivalente topológicamente cerca del origen a una de las siguientes formas normales:

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = \begin{pmatrix} \alpha & -1 \\ 1 & \alpha \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \pm (y_1^2 + y_2^2) \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Las condiciones de genericidad asumidas en el teorema (3.7) son la condición de no degeneración **(B.1)** y la condición de transversalidad **(B.2)** del teorema (3.6).

4 Modelo Neuronal

Uno de los modelos fundamentales en la neurociencia computacional es el que incluye una corriente persistente de sodio y una corriente de potasio $I_{Na,p} + I_K$, donde la corriente de Na^+ es rápida y la de K^+ es relativamente más lenta:

$$\begin{aligned} C\dot{V} &= I - \overbrace{g_L(V - E_L)}^{\text{débil } I_L} - \overbrace{g_{Na}m(V - E_{Na})}^{I_{Na,p}} - \overbrace{g_K n(V - E_K)}^{I_K} \\ \dot{m} &= (m_\infty(V) - m) / \tau_m(V) \\ \dot{n} &= (n_\infty(V) - n) / \tau_n(V) \end{aligned}$$

Este modelo equivale en muchos aspectos al modelo $I_{Ca} + I_K$ propuesto por Morris y Lecar para describir las oscilaciones del voltaje en la fibra muscular del percebe gigante. Una suposición razonable basada en las observaciones experimentales es que la variable $m(t)$ de la puerta de Na^+ es mucha más rápida que la variable de voltaje $V(t)$ así que m aproxima prácticamente de manera instantánea el valor asintótico $m_\infty(V)$, en este caso podemos sustituir $m = m_\infty(V)$ en la ecuación del voltaje y reducir el sistema tridimensional anterior a un sistema plano

$$\begin{aligned} C\dot{V} &= I - \overbrace{g_L(V - E_L)}^{\text{débil } I_L} - \overbrace{g_{Na}m_\infty(V - E_{Na})}^{\text{instantánea } I_{Na,p}} - \overbrace{g_K n(V - E_K)}^{I_K} \\ \dot{n} &= (n_\infty(V) - n) / \tau_n(V) \end{aligned} \quad (16)$$

Muchas características interesantes del comportamiento de la variación del voltaje local de una sola neurona se pueden explicar o ilustrar mediante estos sistemas bidimensionales, aún los disparos en ráfagas, los cuales se presentan en neuronas con mayor variedad de canales y que se modelan mediante sistemas con mayor cantidad de ecuaciones, se pueden entender a través del análisis de las bifurcaciones en estos sistemas. En este caso trabajaremos un modelo $I_{Na,p} + I_K$ con bajo umbral que presenta la aparición de un ciclo límite vía la bifurcación de Andronov Hopf.

Empezaremos el análisis de este sistema. En este modelo las funciones m_∞ y n_∞ toman la siguiente forma:

$$\begin{aligned} m_\infty(V) &= \frac{1}{1 + \exp\left\{\left(V_{\frac{1}{2}} - V\right)/k\right\}}, \\ n_\infty(V) &= \frac{1}{1 + \exp\left\{\left(V_{\frac{1}{2}} - V\right)/k\right\}}. \end{aligned}$$

Los parámetros en el modelo son: $C = 1$, $E_L = -78mV$, $g_L = 8$, $g_{Na} = 20$, $g_K = 10$, $m_\infty(V)$ tiene $V_{\frac{1}{2}} = -20$ y $k = 15$, $n_\infty(V)$ tiene $V_{\frac{1}{2}} = -45$, $k = 5$ y $\tau(V) = 1$,

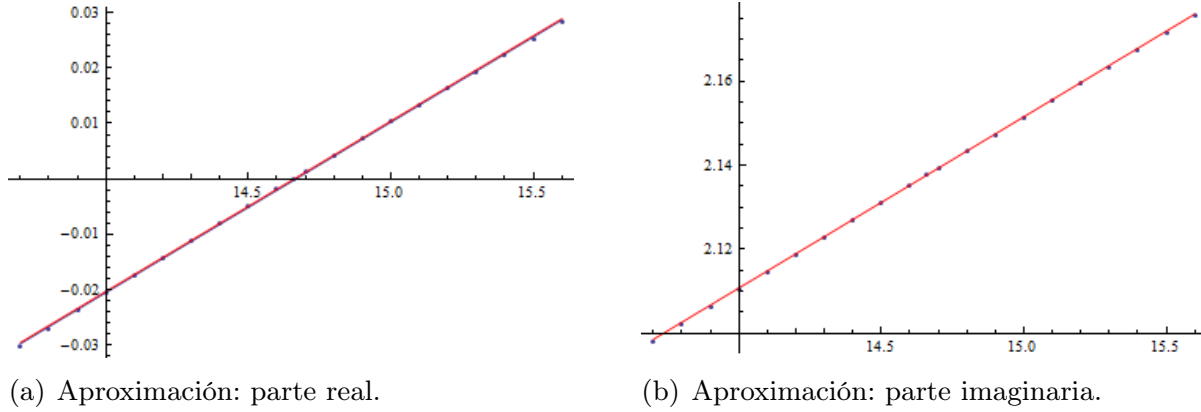


Figura 7: Aproximaciones numéricas

$E_{Na} = 60mV$ y $E_K = -90mV$; así pues tenemos el siguiente sistema:

$$\begin{aligned}\dot{V} &= I - 8(V + 78) - 20m_\infty(V - 60) - 10n(V + 90) \\ \dot{n} &= (n_\infty(V) - n) / \tau_n(V).\end{aligned}\tag{17}$$

Debido a la complejidad de las funciones, un análisis idéntico al del modelo depredador presa es un poco complicado, por lo cual, mediante simulaciones, se ha encontrado que el valor crítico para esta bifurcación es $I_0 = 14.659$. Con este parámetro, resolviendo numéricamente las ecuaciones para obtener el punto de equilibrio, tenemos que $E_0 = (-56.4815, .0914299)$ y cuya matriz jacobiana evaluada en el punto de equilibrio tiene la forma

$$A = \begin{pmatrix} .999998 & -335.185 \\ .0166141 & -1 \end{pmatrix}$$

con valores propios $\lambda_{1,2} = \pm 2.13748i$. Por la complejidad de las funciones involucradas en este modelo, se obtendrán numéricamente algunos valores para los valores propios de la matriz al cambiar el parámetro que en este caso es la corriente aplicada en una pequeña vecindad de I_0 y después procederemos a utilizar el método de mínimos cuadrados para obtener así una aproximación para las funciones de la parte real e imaginaria de los valores propios de la matriz jacobiana (v. Figura 7). De esta forma tenemos que $\mu(I) = .0307066(I - 14.659)$ y $w(I) = 2.13748(I - 14.569)$ y así tenemos que $\mu'(I_0) = .0307066 \neq 0$, por lo tanto, se cumple la condición de transversalidad **(B.2)** del teorema (3.6).

Para calcular el primer coeficiente de Liapunov realizaremos primero un cambio de coordenadas para trasladar el punto de equilibrio al origen y fijaremos el valor

del parámetro en I_0 , con lo cual, haciendo $V = -56.4815 + \gamma_1$ y $n = .0914299 + \gamma_2$ tenemos el nuevo sistema:

$$\begin{aligned}\dot{\gamma}_1 &= -185.135 + \frac{2329.63 - 20\gamma_1}{1 + 11.3828e^{\frac{-\gamma_1}{15}}} - 8.9143\gamma_1 - 335.185\gamma_2 - 10\gamma_1\gamma_2, \\ \dot{\gamma}_2 &= -.0914299 + \frac{1}{1 + e^{\frac{1}{5}(11.4815 - \gamma_1)}} - \gamma_2.\end{aligned}\quad (18)$$

Este nuevo sistema se puede escribir de la forma

$$\dot{\gamma} = A\gamma + \frac{1}{2}B(\gamma, \gamma) + \frac{1}{6}C(\gamma, \gamma, \gamma) + O(\|\gamma\|^4).$$

De esta forma la matriz A y las funciones multilineales B y C tienen la forma:

$$\begin{aligned}A &= \begin{pmatrix} .999998 & -335.185 \\ .0166141 & -1 \end{pmatrix}, \\ B(\gamma, \eta) &= \begin{pmatrix} -10(\gamma_1\eta_2 + \eta_1\gamma_2) \\ 0 \end{pmatrix}, \\ C(\gamma, \eta, \xi) &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}.\end{aligned}$$

Sea

$$q = \begin{pmatrix} .999975 \\ .00300479 + .00636687i \end{pmatrix}$$

un vector propio asociado a $\lambda = 2.13748i$ y consideremos p un vector asociado a A^T y a $\bar{\lambda}$,

$$p = \begin{pmatrix} -.002983 + .00637684i \\ .999975 \end{pmatrix}$$

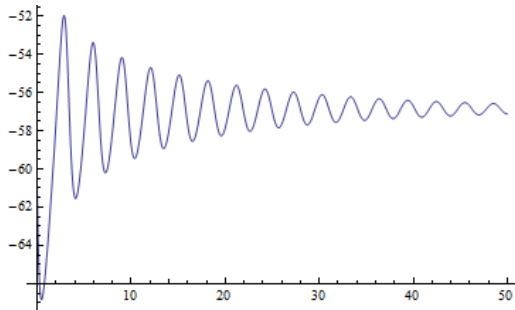
para la normalización $\langle q, p \rangle = 1$ tomaremos a

$$p = \begin{pmatrix} .500013 + .233927i \\ -78.4087i \end{pmatrix}$$

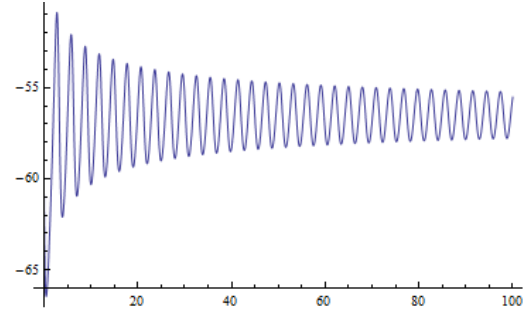
ahora podemos hacer los siguientes cálculos:

$$g_{2,0} = \langle B(q, q), p \rangle = .0777259i$$

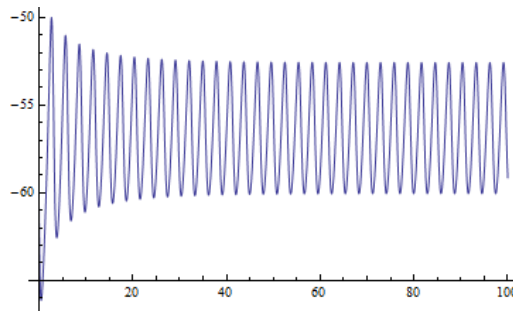
$$g_{1,1} = \langle B(q, \bar{q}), p \rangle = .0298335 + .0139573i$$



(a) Antes del valor crítico.



(b) En el valor crítico.



(c) Después del valor crítico.

Figura 8: Comportamiento del voltaje durante la bifurcación.

nótese que en este modelo, por la forma de la función multilineal C , tenemos que $g_{2,1} = 0$; utilizando la fórmula (15) tenemos que

$$l_1(I_0) = \frac{1}{2w^2} \operatorname{Re}(ig_{2,0}g_{1,1} + wg_{2,1}) = .000253767$$

así la condición **(B.1)** del teorema (3.6) se cumple y por lo tanto en este modelo se presenta una bifurcación de Andronov Hopf hipercrítica en el valor del parámetro I_0 .

Veamos el comportamiento del voltaje durante la bifurcación, así como también el comportamiento en el plano fase.

La figura 1.8(a) muestra el comportamiento del voltaje en el sistema cuando el parámetro (en este caso es la corriente aplicada) está antes del valor de bifurcación I_0 ; se nota que el voltaje presenta oscilaciones que van disminuyendo su amplitud hasta converger al punto de equilibrio, por lo tanto este es un foco estable. En 1.8(b) se observa el comportamiento del voltaje cuando el valor del parámetro se encuentra en el valor crítico I_0 ; notamos un comportamiento parecido al del caso anterior pero con la diferencia que la velocidad de convergencia es mucho más lenta lo que implica que el punto de equilibrio es un foco débilmente atractor. Finalmente

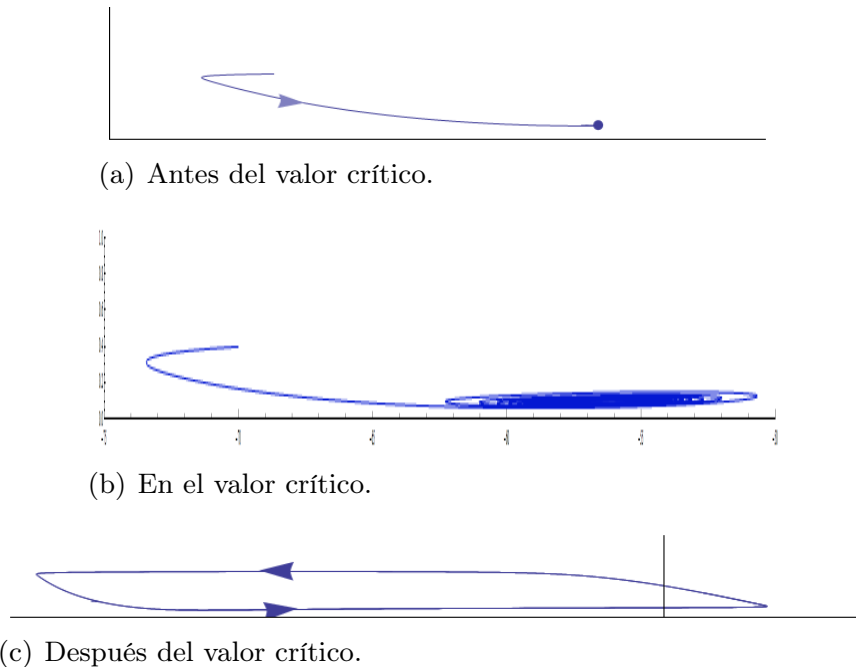


Figura 9: Comportamiento del voltaje durante la bifurcación.

observamos en 1.8(c) que la amplitud del voltaje converge a un valor constante, lo cual nos dice que existe un ciclo límite al cual van a converger las órbitas cuando el parámetro se encuentra después del valor I_0 .

En la figura 9 observamos el comportamiento de las trayectorias en el plano fase; en 1.9(a) antes del valor crítico I_0 vemos cómo la trayectoria converge al punto de equilibrio; en 1.9(b) vemos que la trayectoria converge al punto de equilibrio igual que en el caso anterior, pero con una velocidad de convergencia mucho menor y finalmente en 1.9(c) vemos la existencia del ciclo límite al cual converge la trayectoria cuanto $t \rightarrow \infty$.

Bibliografía

- [1] Fernández Pérez, C., Vázquez Hernández, F. J. y Vegas Montaner, J. M. (2003) *Ecuaciones diferenciales y en diferencias. Sistemas dinámicos*. Editorial Thomson. Págs. 124-130.
- [2] Friedberg, S. H., Insel, A. J. y Spence, L. E. (1982) *Álgebra Lineal*. Editorial Publicaciones Cultural, S.A. Pág. 235
- [3] Izhikevich, E. M. (2001) “Resonate and fire neurons”. *Neural Networks*, 14.

- [4] Izhikevich, E. M. (2007) *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. Editorial The MIT Press. Págs. 8-16, 89, 172-174.
- [5] Kuznetsov, Y. A. (1998) *Elements of Applied Bifurcation Theory*. Editorial Springer. Caps. 1-3
- [6] Munkres, J. R. (1974) *Topology A first course*. Editorial Prentice-Hall. Págs. 22, 104
- [7] Perko L. (2001) *Differential Equations and Dynamical Systems*. Editorial Springer. Págs 211-216.
- [8] Rinzel, J. y Ermentrout, G. B. (1989) “Analysis of neural excitability and oscillations”, en Christof Koch e Idan Segev *Methods in Neural Modeling*, Cap. 7.

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

Nota: Los autores están en orden alfabético de acuerdo al formato requerido, no indica nivel de participación. Autor principal: Leonardo Remedios Santiago, hostem.darkleo@gmail.com

lcervant@fcfm.buap.mx
anilu_g_65@hotmail.com
jpoisot@fcfm.buap.mx
214470282@fcfm.buap.mx

Capítulo 2

Estabilidad y bifurcación para algunos modelos Depredador-Presa

Evodio Muñoz Aguirre, Francisco Sergio Salem Silva
Facultad de Matemáticas, Universidad Veracruzana

Resumen

En este capítulo se describen los modelos de crecimiento de poblaciones unidimensionales de Volterra y Verhulst para con ello dar paso a la explicación de las ecuaciones bidimensionales que componen el modelo de Lotka-Volterra o Depredador-Presa, para cuando éstas se presenten. Posteriormente se muestra una clasificación de los modelos depredador-Presa de acuerdo al tipo de respuesta funcional que contienen. Más adelante se realiza un análisis de la estabilidad de los puntos de equilibrio de dos modelos conocidos en la literatura. Por último se analiza la bifurcación de los mismos sistemas considerando como parámetro de bifurcación un solo parámetro. Aquí se presenta un modelo de tres ecuaciones diferenciales ordinarias que involucra estructura sexual.

1 Introducción

Existe una gran variedad de estudios sobre el crecimiento de una población. De entre los más conocidos y utilizados en muchos campos, principalmente en Biología, se tienen algunos modelos clásicos, como el de Volterra y Verhulst. Estos modelos aparecen muy frecuentemente en la literatura, sobre todo la correspondiente al área de Biología Matemática, área en la cual se explican con más detalles estos sistemas y se hace un análisis a profundidad del comportamiento de éstos en términos de su estabilidad y su bifurcación, ver por ejemplo [8]. La estabilidad de estos sistemas se ha estudiado de diferentes puntos de vista: por aproximación, mediante construcción de funciones de Lyapunov, etc. Además, se ha visto que la existencia de un punto de equilibrio único es equivalente a estudiar la estabilidad global del sistema, en algunos casos incluso se han construido funciones de Lyapunov para probar esta aseveración, tal como se menciona en [9].

Para el caso de la interacción de dos poblaciones es muy útil utilizar el modelo particular Depredador-Presa de Lotka Volterra, con las hipótesis que se requieran en cada caso. Para realizar el análisis, sistemas de esta clase han sido objeto de

estudio por varios autores, entre ellos Sze-Bi Hsu y Tsy-Wei Huang [10] hicieron un análisis de la estabilidad usando el criterio de Dulac y algunas funciones de Lyapunov; Shunyi Li y Zuoliang Xion [11] hacen un análisis de bifurcación a un sistema aumentándole una estructura de sexo y favoritismo sexual, además de considerar ecuaciones diferenciales con retardo.

Particularmente, en este capítulo se hace una breve descripción de los modelos de crecimiento de poblaciones clásicos de Malthus-Verhulst. Si bien el objetivo es realizar un estudio de algunos sistemas de Lotka-Volterra que describen la evolución de dos poblaciones, es imprescindible mencionar que cada ecuación del mismo tienen una analogía con ecuaciones de una sola población y su formulación se hace de manera análoga. Una vez detallado el sistema depredador-presa y después de dar la clasificación de éstos según Leslie, se realiza un breve análisis de la estabilidad y bifurcación de dos modelos tipo Leslie y un modelo con estructura sexual. Por último se menciona la aparición de retardo en los sistemas con estructura de sexo, lo que da origen al estudio de la bifurcación de los mismos, considerando el retardo como un parámetro. El tratamiento desde este punto de vista no se detalla puesto que aparecen ecuaciones diferenciales con retardo en el modelo y su tratamiento está fuera del alcance de este material.

Cabe aclarar que la clasificación de los modelos presentados en este trabajo no es la única, algunos autores, como [5] y [6] indicados en las referencias, hacen otra clasificación y análisis utilizando funciones que las llaman inclusive con el mismo nombre, pero las proponen un tanto más generales.

2 El Modelo de Malthus-Verhulst para una especie

Existen leyes de conservación en algunas áreas de la ciencia, por lo regular aparecen más frecuentemente en la Química y en la Física, por ejemplo leyes sobre conservación de los momentos y la masa o incluso conservación de la energía. En la Biología también existe una ecuación o ley de conservación para la razón de cambio de una población $x(t)$ respecto al tiempo t , ésta dice: *el cambio que sufre una población respecto del tiempo en un medio ambiente se da considerando el número de nacimientos, menos el número de muertes, más la población que llega al medio o sale de él*. Explícitamente puede formularse matemáticamente mediante la fórmula:

$$\frac{dx}{dt} = \text{nacimientos} - \text{muertes} + \text{migración}.$$

Ésta tiene varias formas, la más simple es:

$$\frac{dx}{dt} = bx - dx.$$

Implicando que, cuando se tiene una condición inicial $x(t_0) = x_0$, la solución se describe por la función,

$$x(t) = x_0 e^{(b-d)t},$$

donde b, d son constantes positivas.

Lo que se explica en los pasados renglones, no es más que la conocida ley de crecimiento de poblaciones: *El cambio que sufre una población respecto del tiempo, es proporcional a la población misma en ese instante de tiempo.* Para el caso explicado, la constante de proporcionalidad es $b - d$.

Un modelo más realista es el modelo conocido como restrictivo (de Verhulst) o Modelo Logístico. Este modelo se diferencia del anterior, en el sentido de que su solución no tiene un comportamiento exponencial o que no tienda a infinito cuando el tiempo tiende a infinito, justo como sucede con la solución de la ecuación del modelo anterior. Es razonable pensar que el comportamiento de las poblaciones en la realidad, no debe ser de esta manera. El modelo Logístico se describe como:

$$\frac{dx}{dt} = rx \left(1 - \frac{x}{K}\right),$$

donde r y K son constantes positivas.

Características del Modelo Logístico:

- En este modelo, el índice de nacimiento per capita es $r(1 - x/K)$.
- K es la capacidad de carga del medio, ésta es determinada por la cantidad de recursos disponibles que suele consumir la población para sobrevivir.
- Se puede constatar inmediatamente que el sistema cuenta con dos puntos de equilibrio, a saber: $x = 0$ que es un punto de equilibrio estable y $x = K$ que es un punto de equilibrio inestable.
- Si se supone una población inicial en el tiempo $t = 0$ como $x(0) = x_0$, la solución es descrita por la función:

$$x(t) = \frac{x_0 K e^{rt}}{[K + x_0(e^{rt} - 1)]}.$$

- Una característica muy peculiar para esta solución es que la población tiende a K cuando t tiende a infinito.
- En general el modelo puede tener en su segundo miembro a una función de cualquier tipo que dependa de x , es decir, se puede escribir: $\frac{dx}{dt} = f(x)$.

Los modelos anteriores nos ayudarán a comprender el modelo Depredador-presa que se presenta en la siguiente sección. Aunque este modelo contiene dos ecuaciones, cada una de ellas representa la evolución de una sola de las poblaciones considerando las interacciones con la otra.

3 Modelo Lotka-Volterra

En esta sección se presenta el modelo clásico de Lotka-Volterra que describe la evolución de dos poblaciones compuestas por los depredadores y sus presas. Según [8], el modelo tiene la forma:

$$\begin{aligned}\frac{dx}{dt} &= x(a - by) \\ \frac{dy}{dt} &= -y(d - cx),\end{aligned}\tag{1}$$

donde x es el número de presas (por ejemplo una población de conejos), y es el número de algún depredador (por ejemplo una población de lobos); y cada una de las ecuaciones representan la razón de crecimiento de cada una de las dos poblaciones en el tiempo t . Los parámetros a, b, c y d que aparecen en las ecuaciones describen la interacción entre las especies.

El modelo hace un número de suposiciones acerca del desarrollo y evolución de las poblaciones del depredador y de la presa:

1. La población de la presa encuentra comida todo el tiempo.
2. El suministro de comida para la población depredador depende completamente de la población presa.
3. La razón de cambio de la población es proporcional a su tamaño.
4. Durante el proceso, el medio ambiente no cambia en favor de una especie y la adaptación genética es suficientemente suave.
5. Como se utilizan ecuaciones diferenciales la solución es determinista y continua. Esto implica que las generaciones de ambas se desarrollan continuamente.

3.1 Explicación de las ecuaciones

A continuación se explica cada uno de los términos que aparecen en las ecuaciones del sistema (1).

- a) Se supone que la presa tiene un suministro de alimentos ilimitada y se reproduce exponencialmente a menos que esté sujeta a depredación; este crecimiento exponencial está representado en la primera de las ecuaciones diferenciales del sistema con el término ax , aquí a es la constante de proporcionalidad requerida si el crecimiento de la población es de este tipo. La razón de depredación sobre la presa se supone proporcional a la razón a la cual los depredadores y las presas se encuentran, lo que se representa en la ecuación (1) mediante el término bxy , donde b es la constante de proporcionalidad. Así, la primera ecuación del sistema se interpreta como que, el cambio en el número de presas es determinado por su crecimiento propio menos la tasa a la cual es depredada.
- b) En la segunda ecuación diferencial del mismo sistema y correspondiente al crecimiento del depredador, cxy representa el crecimiento de la población depredador, c es la constante de proporcionalidad; notemos la similitud de la razón de depredación, sin embargo se usa una constante diferente ya que la razón a la cual la población del depredador crece, no necesariamente es igual a la razón a la cual éste consume la presa. El parámetro dy representa la tasa de pérdida de los depredadores debido a muerte natural o emigración, en este caso d es una constante que indica el índice de muerte o emigración. Por lo tanto, la ecuación indicada expresa el cambio en la población depredador como crecimiento debido al suministro de comida, menos muerte natural o emigración.

3.2 Sistema adimensionalizado

Haciendo un cambio de variable utilizando las variables $u(\tau) = \frac{cx(t)}{d}$, $v(\tau) = \frac{by(t)}{a}$, $\tau = at$ y $\alpha = \frac{d}{a}$. Después de sustituir estas variables haciendo algunas operaciones básicas y regresando a la notación respecto a las variables más conocidas x, y, t , se llega al siguiente sistema:

$$\begin{aligned}\frac{dx}{dt} &= x(1 - y) \\ \frac{dy}{dt} &= \alpha y(x - 1).\end{aligned}\tag{2}$$

El sistema (2) es más simple que el sistema (1) puesto que solo contiene un parámetro y es más fácil de analizar en el plano fase x, y . Mejor aún, es posible convertir este sistema a una ecuación diferencial ordinaria en la variable dependiente

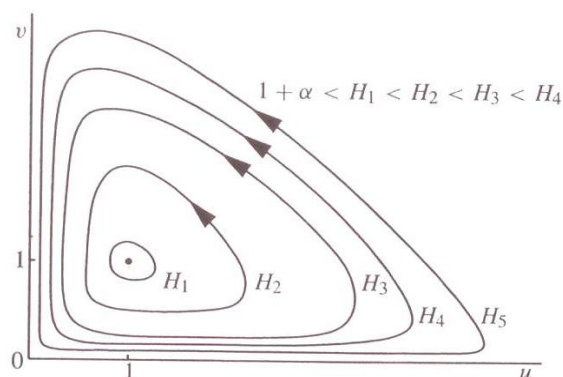


Figura 1

y respecto a la variable independiente x , de la forma:

$$\frac{dy}{dx} = \alpha \frac{y(x-1)}{x(1-y)},$$

con puntos singulares en $x = y = 0$, $x = y = 1$.

Es verificable inmediatamente que esta ecuación diferencial ordinaria se resuelve fácilmente por el método de separación de variables, dando inmediatamente de forma explícita las trayectorias:

$$\alpha x + y - \ln x^\alpha y = H,$$

donde $H > H_{min}$ es una constante; $H_{min} = 1 + \alpha$ es el mínimo de H sobre todo (x, y) que se encuentre en el primer cuadrante en donde tiene sentido biológico el problema, y ocurre cuando $x = y = 1$. Para $H > 1 + \alpha$ dado, las trayectorias son como se muestran en la Figura 1. Para hacer la simulación de las trayectorias se consideraron los valores: $H_1 = 2.1$, $H_2 = 2.4$, $H_3 = 3.0$, $H_4 = 4$.

Se observa del comportamiento de las trayectorias en la Figura 1 que son curvas cerradas y periódicas; aún más, se observa que el punto de equilibrio $(0, 0)$ es un punto inestable, mientras el punto de equilibrio $(1, 1)$ es equilibrio estable, resultado coincidente con el que se obtiene cuando se hace un análisis del sistema lineal, como se puede verificar en [7].

4 Modelo depredador-presa realista

El modelo de Lotka-Volterra aunque es un poco realista, sugiere que las interacciones depredador-presa pueden presentar comportamientos periódicos de las poblaciones, de acuerdo al razonamiento que detenidamente se desprende del modelo y que se describe en las siguientes observaciones:

- Si la población de presas se incrementa y como es el único alimento de la población depredador, entonces se favorece al crecimiento de este último.
- Es lógico razonar que si hay más depredadores, éstos consumirán más presas, y la población de presas comienza a disminuir.
- En cambio, la pérdida de comida para la población depredador, hace que el número de depredadores disminuya.
- Cuando la población depredador está en su punto más bajo, permite que se incremente la población de presas, y así sucesivamente.

Dependiendo de los detalles que presente el sistema de acuerdo a los parámetros que aparecen en él, se concluye que éste evoluciona de acuerdo a oscilaciones y que éstas suelen crecer o decaer hacia algún ciclo límite estable o a algún ciclo límite inestable, o exhibir comportamiento caótico (sobre todo cuando el sistema se describe, por ejemplo, con un sistema de tres ecuaciones diferenciales).

Una de las suposiciones poco realistas del modelo de Lotka-Volterra es que el crecimiento de la presa es acotado en ausencia de la depredación. Para el modelo indicado por el sistema (2) de la sección anterior, los términos entre paréntesis describen la tasa de crecimiento per capita y dependen de la densidad de la otra población. Para ser más realistas, esta tasa de crecimiento puede depender de las densidades de la presa y del depredador, así el modelo puede tener la forma:

$$\begin{aligned}\frac{dx}{dt} &= xF(x, y) \\ \frac{dy}{dt} &= -yG(x, y),\end{aligned}\tag{3}$$

donde las funciones de dos variables F y G pueden depender de la interacción entre las especies, de las mismas especies particulares, y en algunos casos de otras constantes.

Como un primer paso razonable para encontrar un modelo más realista, se puede esperar que la presa satisfaga un crecimiento logístico, digamos que en la ausencia de depredadores, ésta tiene un crecimiento dinámico el cual tiende a una capacidad de carga máxima, otra analogía con el modelo de Verhulst. Luego, una ecuación más realista para la población presa puede tener la forma:

$$\frac{dx}{dt} = xF(x, y), \quad F(x, y) = r\left(1 - \frac{x}{K}\right) - yR(x)$$

donde la respuesta funcional $R(x)$ contiene uno de los términos de depredación ilustrados en la Figura 2, mientras que K es la constante de capacidad de carga para la población presa cuando la población depredador satisface $y = 0$.

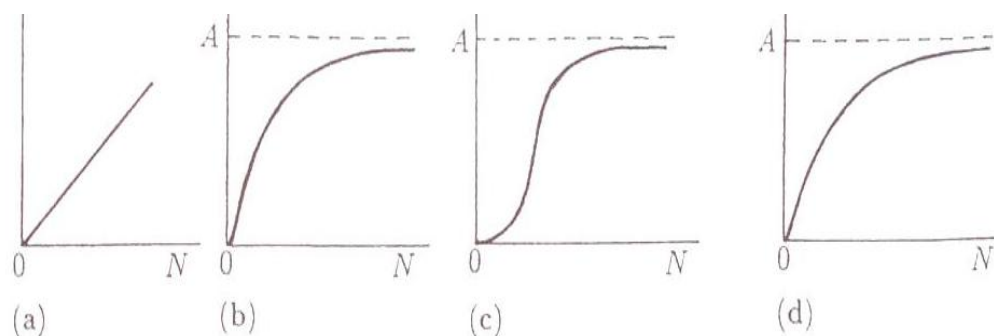


Figura 2

El término de depredación (en este caso $R(x)$), se conoce en la literatura con el nombre de **respuesta funcional** del depredador para cambiar a la densidad de la presa y generalmente presenta algún efecto de saturación, tal como las gráficas de la Figura 2. La cuestión es que en vez de tomar una respuesta del depredador bxy como en el modelo de Lotka Volterra, se puede tomar $yxR(x)$ donde $xR(x)$ se satura para x grande. Algunos ejemplos de respuesta funcional tiene las siguientes formas especiales:

$$R(x) = \frac{A}{x+B}, \quad R(x) = \frac{Ax}{x^2+B^2}, \quad R(x) = \frac{A[1-e^{-ax}]}{x}.$$

donde A, B y a son constantes positivas, y las saturaciones para cada caso son ilustradas en la Figura 2 en los incisos (b) a (d). Los ejemplos mostrados en la figura correspondientes a (b) y (c) son aproximadamente lineales en x para densidades pequeñas. la saturación para x grande es un reflejo de la capacidad limitada del depredador, o perserverancia, cuando la presa es abundante.

Igualmente, existen otras posibles formas para la función G presente en la segunda ecuación del sistema (3), de entre ellas se tiene las siguientes:

$$G(x, y) = k \left(1 - \frac{hy}{x}\right), \quad G(x, y) = -d + eR(x)$$

donde k, h, d y e son constantes positivas, y $R(x)$ como se ha descrito anteriormente. La primera de las dos ecuaciones indicadas anteriormente, nos dice que la capacidad de carga para el depredador es directamente proporcional a la densidad de la presa.

Algo análogo sucede con la ecuación de la población depredador, por ejemplo, la función G puede tener una de las siguientes formas:

$$G(x, y) = k \left(1 - \frac{hy}{x}\right), \quad G(x, y) = -d + eR(x)$$

donde k , h , d y e son constantes positivas y $R(x)$ es como ya se describió en la ecuación para la presa.

Un ejemplo clásico muy común en la literatura de sistemas tipo Lotka-Volterra se encuentra explícitamente en [9]. Este modelo depredador-presa se describe a continuación.

El modelo se representa por el sistema de ecuaciones:

$$\begin{aligned}\frac{dx}{dt} &= rx\left(1 - \frac{x}{K}\right) - yp(x) \\ \frac{dy}{dt} &= y\left[s\left(1 - \frac{hy}{x}\right)\right],\end{aligned}\tag{4}$$

con $x(0) > 0$, $y(0) > 0$, $r, s, h, K > 0$, donde x es la población de la presa y y la población del depredador. En este sistema se supone que la presa crece logísticamente con capacidad de carga K y razón de crecimiento intrínseco r en ausencia de depredación. El depredador consume la presa de acuerdo a la respuesta funcional $p(x)$ y crece logísticamente con tasa de crecimiento intrínseco s y capacidad de carga proporcional al tamaño de la población presa. El parámetro h es el número de presas requerido para sostener un depredador en equilibrio cuando $y = \frac{x}{h}$.

5 Clasificación según la respuesta funcional

En la correspondiente literatura existe una clasificación en tres tipos de modelos Lotka-Volterra descritos por la ecuación (4), de acuerdo a la forma de la respuesta funcional $p(x)$, tal como se puede constatar en [9]. En esta sección se retoma la misma clasificación realizada en la bibliografía indicada. Todo esto se describe a continuación.

Modelo Tipo I. De Leslie-Gower (LG)

Si la función depredadora tiene la forma $p(x) = mx$, el sistema recibe el nombre de **Leslie Tipo I** o de **Leslie-Gower**, y se escribe como:

$$\begin{aligned}\frac{dx}{dt} &= rx\left(1 - \frac{x}{K}\right) - mxy \\ \frac{dy}{dt} &= y\left[s\left(1 - \frac{hy}{x}\right)\right],\end{aligned}\tag{5}$$

$x(0) > 0$, $y(0) > 0$.

Modelo Tipo II. Holling-Tanner (HT)

En cambio, si la función depredadora se toma como $p(x) = \frac{mx}{A+x}$, el sistema recibe el nombre de **Leslie Tipo II** o de **Holling-Tanner** y sus ecuaciones forman el sistema:

$$\begin{aligned}\frac{dx}{dt} &= rx\left(1 - \frac{x}{K}\right) - \frac{mx}{A+x}y \\ \frac{dy}{dt} &= y\left[s\left(1 - \frac{hy}{x}\right)\right],\end{aligned}\tag{6}$$

$x(0) > 0$, $y(0) > 0$.

Es importante hacer unos comentarios acerca del modelo de Leslie tipo II. La respuesta funcional saturante del depredador $\frac{mx}{A+x}$ es del tipo Michaelis-Menten que se utiliza en la cinética enzima-sustrato en la Físico-química, el parámetro m es la razón máxima específica de la formación del producto. x es la concentración de sustrato y A (la constante de saturación media) es la concentración de sustrato a la cual la tasa de formación del producto es la mitad del máximo. La función fue propuesta por Holling para depredadores sin aprendizaje, el nivel de poco aprendizaje es un poco engañoso porque incluso los depredadores capaces de aprender pueden exhibir este tipo de respuesta cuando se administra un solo tipo de presa que desea buscar o cazar. De acuerdo a la derivación de Holling las constantes son ahora: $m = \frac{1}{t_h}$ y $A = \frac{1}{ct_h}$, donde t_h es el tiempo de manipulación por presa y c es la razón de encuentro por unidad de la densidad de presa.

El modelo de Holling Tipo II proviene de la función depredadora con forma general $p(x) = \frac{x^p}{b+x^p}$ cuando $p = 1$.

Modelo Tipo III

Para esta clase de modelos, la función depredadora se caracteriza por la expresión $p(x) = \frac{mx^2}{(A+x)(B+x)}$, y en este caso, el sistema se escribe de la forma:

$$\begin{aligned}\frac{dx}{dt} &= rx\left(1 - \frac{x}{K}\right) - \frac{mx^2}{(A+x)(B+x)}y \\ \frac{dy}{dt} &= y\left[s\left(1 - \frac{hy}{x}\right)\right],\end{aligned}\tag{7}$$

$x(0) > 0$, $y(0) > 0$.

Aquí, la función depredadora es una curva de forma sigmoidal; es decir, en forma de S, estas funciones son típicas en depredadores que muestran alguna forma de aprendizaje del comportamiento de la presa, en el cual, bajo cierto nivel de umbral de densidad, el depredador no utilizará la presa para comida en una cantidad muy

grande. Sin embargo, arriba de aquel nivel de densidad, el depredador aumenta su razón de alimentación hasta que se alcanza algún nivel de saturación. Holling razonó que estos animales tienden a aprender lentamente y a olvidar el valor de una comida a no ser que la encuentren con bastante frecuencia. Además este científico dió algunas pruebas de campo en donde muestra que esta respuesta funcional en forma de S es típica para depredadores vertebrados con presa alternativa disponible.

6 Otra Clasificación

Existe una clasificación más, ésta se basa en distinguir a los modelos que cuentan con términos de recolección, o a los modelos en donde a las especies se les divide en presas de acuerdo al sexo.

Modelo con recolección

En efecto, algunos autores, como se puede ver en [12], nos muestran otra forma de clasificar a los sistemas Depredador-Presa, y ésta la hacen de acuerdo a que en las poblaciones existe recolección, ya sea de la población presa o de la población depredador. Esto es muy común en especies marítimas, de presas o de ríos, donde se cosechan peces (depredadores) para consumo humano o peces (presa) que sirven de alimento a otros peces; aunque se puede encontrar en algunos otros medios, incluso en áreas donde no son tan comunes, como lo es la Economía. Un modelo con tales características se describe a continuación:

$$\begin{aligned}\frac{dx}{dt} &= x(\alpha - x) - \frac{\beta x^2}{1 + x^2}y - E_1x \\ \frac{dy}{dt} &= \frac{\beta_1 x^2 y}{1 + x^2} - \gamma y - E_2y,\end{aligned}\tag{8}$$

donde $E_1 \geq 0$ y $E_2 \geq 0$ denotan el esfuerzo de recolección, y son los términos que influyen negativamente en el cambio de la población respecto del tiempo.

Modelo con estructura de sexo

La razón de sexo significa la comparación de individuos masculinos e individuos femeninos en una población. Usualmente se supone la razón 1 : 1. Sin embargo, para algunos modos de vida la razón puede cambiar, ésto es debido a varios factores, entre ellos se tienen a las condiciones del ambiente, el comportamiento social, la adaptabilidad a los recursos, la herencia, la estructura de genes, etc. La razón de sexo en los animales puede cambiar muy a menudo durante su evolución. A través

del paso del tiempo, con la edad, la población femenina tiende a decrecer, pero existe una gran variedad de pájaros para los cuales sucede exactamente lo contrario.

En el siguiente modelo, como se puede constatar en [11], los autores hacen una clasificación dividiendo a la población presa (o depredador) en población masculina y femenina, resultando modelos Lotka-Volterra tridimensionales.

$$\begin{aligned}\frac{dm(t)}{dt} &= b_1 f(t) - d_1 m(t) - k(m(t) + f(t))m(t) - c_1 m(t)x(t) \\ \frac{df(t)}{dt} &= f(t)(\beta - k(m(t) + f(t)) - c_1 x(t)) \\ \frac{dx(t)}{dt} &= x(t)(-a - bx(t) + c_2 m(t) + c_2 f(t)),\end{aligned}\tag{9}$$

En el sistema anterior, las variables dependientes del tiempo t , m y f , son las poblaciones individuales masculinas y femeninas de la presa respectivamente, mientras que la variable x representa la población depredador. Los parámetros $a, b, b_1, c_1, c_2, d_1, d_2, k$ son positivos, b_1, d_1 son constantes de proporcionalidad para el crecimiento de la presa femenina y muerte masculina ($b_2 > d_2$), donde $\beta = b_2 - d_2$) a es la constante de proporcionalidad para un depredador, c_1 es el coeficiente de predación para un depredador y $\frac{c_1}{c_2}$ ($0 < \frac{c_1}{c_2} < 1$) es la tasa de encuentros entre una presa y el depredador respectivamente.

7 Estabilidad de algunos Modelos

Los resultados que se presentan a lo largo de esta sección, se escriben sin realizar demostración alguna, sin embargo, el lector interesado en éstas, las puede encontrar detalladamente en [9] y [10].

7.1 Adimensionalización de los modelos Tipo I, II y III

Utilizando el teorema π -Buckingham, como se hace detalladamente en [4], se obtienen las siguientes variables adimensionales:

$$\begin{aligned}\tilde{t} &= rt, & \tilde{x}(\tilde{t}) &= \frac{x(t)}{K}, & \tilde{y}(\tilde{t}) &= \frac{my(t)}{r}, \\ \delta &= \frac{s}{r}, & \beta &= \frac{sh}{mK}.\end{aligned}$$

Al aplicar el cambio de variable usando las variables adimensionales indicadas, el sistema Leslie-Gower mostrado en (5) se reescribe de la siguiente forma adimen-

sional:

$$\begin{aligned}\frac{dx}{dt} &= x(1-x) - xy, \\ \frac{dy}{dt} &= y\left(\delta - \beta\frac{y}{x}\right), \\ x(0) &> 0, \quad y(0) \geq 0.\end{aligned}\tag{10}$$

Ahora se adimensionalizan los sistemas (6) y (7) utilizando el cambio de variable adimensionalizada como se ilustra en [9]:

$$\tilde{t} = rt, \quad \tilde{x}(\tilde{t}) = \frac{x(t)}{K}, \quad \tilde{y}(\tilde{t}) = \frac{my(t)}{rK}, \quad \delta = \frac{s}{r}, \quad \beta = \frac{\rho h}{m}, \quad a = \frac{A}{K}, \quad b = \frac{B}{K},$$

dando como resultado los sistemas:

$$\begin{aligned}\frac{dx}{dt} &= x(1-x) - \frac{x}{a+x}y \\ \frac{dy}{dt} &= y\left(\delta - \beta\frac{y}{x}\right),\end{aligned}\tag{11}$$

$x(0) > 0, y(0) > 0$; y

$$\begin{aligned}\frac{dx}{dt} &= x(1-x) - \frac{x^2}{(a+x)(b+x)}y \\ \frac{dy}{dt} &= y\left(\delta - \beta\frac{y}{x}\right)\end{aligned}\tag{12}$$

$x(0) > 0, y(0) > 0$.

Los sistemas (11) y (12) se pueden expresar en forma general con una función depredadora especial $p(x)$ como:

$$\begin{aligned}\frac{dx}{dt} &= xg(x) - p(x)y \\ \frac{dy}{dt} &= y\left(\delta - \beta\frac{y}{x}\right)\end{aligned}\tag{13}$$

$x(0) > 0, y(0) > 0$.

7.2 Análisis de la estabilidad del sistema general

En esta subsección se van a presentar algunos resultados (lemas y teoremas) sobre la estabilidad del sistema general (13), las demostraciones se pueden encontrar en [9]. Junto a este sistema, se suponen las siguientes dos hipótesis:

$H_1: g(1) = 0$ y $g'(x) < 0$ para $x > 0$.

$H_2: p(0) = 0$ y $p'(x) > 0$ para $x > 0$.

Considerando ambas hipótesis, el siguiente lema nos garantiza que el sistema tiene un muy buen comportamiento.

Lema 7.1. *Supóngase que se cumplen H_1 y H_2 . Entonces las soluciones de (13) son positivas y acotadas, y además, existe $T \geq 0$ tal que $x(t) < 1$, $y(t) < \frac{\delta}{\beta}$ para $t \geq T$.*

La conclusión del lema nos dice que las poblaciones son acotadas, esto concuerda con la realidad del problema biológico, ya que como sabemos, las poblaciones, por muy grandes que sean no tiene crecimiento infinito.

Puntos de equilibrio de (14)

Haciendo un análisis sobre las soluciones de los miembros derechos igualados a cero, se tienen el punto de equilibrio $E_1 = (1, 0)$.

Si se cumplen las hipótesis H_1 y H_2 , entonces las gráficas de $g(x)$ y $p(x)\frac{\delta}{\beta}$ tiene intersección única x^* , satisfaciendo $0 < x^* < 1$. Así el sistema (13) posee un único punto de equilibrio $E^* = (x^*, y^*)$ donde $y^* > 0$ satisface $y^* = \frac{x^*g(x^*)}{p(x^*)} = \frac{\delta}{\beta}x^*$. El punto de equilibrio posee las dos coordenadas positivas ya que también x^* es positivo.

Utilizando el segundo método de Lyapunov con ayuda de una función de Lyapunov apropiada, se puede demostrar el siguiente teorema.

Teorema 7.2. *Supóngase que se cumplen H_1 y H_2 . Si*

$$(x - x^*)\left(\frac{xg(x)}{p(x)} - y^*\right) < 0 \quad \text{para} \quad 0 < x < 1, x \neq x^*$$

entonces las soluciones de (13) satisfacen

$$\lim_{t \rightarrow \infty} x(t) = x^* \quad \text{y} \quad \lim_{t \rightarrow \infty} y(t) = y^*.$$

7.3 Estabilidad del modelo Tipo II

Sea el sistema adimensionalizado Tipo II:

$$\begin{aligned} \frac{dx}{dt} &= x(1-x) - \frac{x}{a+x}y = f(x, y) \\ \frac{dy}{dt} &= y\left(\delta - \beta\frac{y}{x}\right) = g(x, y), \end{aligned} \tag{14}$$

$x(0) > 0, y(0) > 0$.

El sistema (14) tiene un único punto de equilibrio con coordenadas positivas $E^* = (x^*, y^*)$. Para este punto se cumple el siguiente lema:

Lema 7.3. *El punto de equilibrio E^* de (14) es asintóticamente estable localmente si $P(x^*) > 0$, y E^* es un foco inestable si $P(x^*) < 0$; donde $P(x) = 2x^2 + (a + \delta - 1)x + a\delta$.*

El teorema que se presenta a continuación es importante porque además de dar condiciones para la estabilidad del punto de equilibrio, nos garantiza condiciones para la existencia de un ciclo. Este resultado se puede probar mediante una serie de desigualdades, justo como se puede observar en la bibliografía indicada al iniciar la sección 7.2.

Teorema 7.4.

- i) Supóngase que se cumple $a + \delta \geq 1$ o $a + \delta > 1$ y $(1 - a - \delta)^2 - 8a\delta \leq 0$. Entonces el punto de equilibrio E^* de (14) es globalmente asintóticamente estable en el interior del primer cuadrante.*
- ii) Si se satisface que $a + \delta < 1$ y $(1 - a - \delta)^2 - 8a\delta > 0$, y $\alpha_2 < x^* < 1$, entonces se cumple la conclusión de i).*
- iii) Sea $a + \delta < 1$ y $(1 - a - \delta)^2 - 8a\delta > 0$ cierta. Para $\beta > 0$ suficientemente pequeño $x^* = x^*(\beta)$ es suficientemente cercano a cero y se cumple $0 < x^* < \alpha_1$. Además, se cumple i) para $\beta > 0$ suficientemente pequeño.*
- iv) Supongamos que se cumple $\alpha_1 < x^* < \alpha_2$, entonces existe un ciclo límite para (14).*

Nota 1: α_1 y α_2 se obtienen de expresar $P(x) = 2(x - \alpha_1)(x - \alpha_2)$.

Nota 2: Observaciones relacionadas con el equilibrio.

A continuación se hacen algunos comentarios sobre interpretaciones que tienen relación con el comportamiento de las poblaciones presa y depredador en función del teorema anterior. Estos comentarios se hacen únicamente para este teorema, pero en [9] se tiene una serie de conclusiones completas tanto para este resultado como el de la siguiente sección.

En el modelo (14) interviene tres constantes, a saber: a, δ y β . De acuerdo al cambio de variable, $a = \frac{A}{K}$ es la razón entre la constante de semi-saturación y la capacidad de carga de la población presa. El parámetro β se puede reescribir como

$\beta = \delta \frac{hr}{m}$. Si se supone que el crecimiento de la presa no es semi-limitado en el equilibrio, hr es el número de presas para reemplazar los individuos muertos por un depredador por unidad de tiempo, mientras m es el número máximo de presas consumidas por un depredador por unidad de tiempo. De esta forma se tiene los siguientes casos para el teorema.

Caso I: $a \geq 1$, es decir, $K \leq A$.

Debido a que $a \geq 1$, la isoclina correspondiente a la presa es monótona decreciente, y el punto de equilibrio E^* es asintóticamente estable globalmente. Por lo tanto para capacidad de carga menor de la presa, la presa y el depredador se aproximan a valores constantes, y no hay ningún comportamiento periódico limitado. Esto sucede para valores independientes de α y β .

Caso II: $a < 1$ y $a + \delta \geq 1$.

Del inciso i) del teorema se sigue que el punto de equilibrio E^* es asintóticamente estable globalmente cuando la capacidad de carga de la presa K es mayor que la semi-saturación A , y la razón $\frac{s}{r}$ es mayor que $1 - a$. Entonces la presa y el depredador se acercan al punto de equilibrio y no hay ningún comportamiento periódico limitado. En particular cuando el índice de crecimiento intrínseco s del depredador es mayor que el índice de crecimiento intrínseco r de la presa, no habrá ciclo límite. Esto sucede independientemente del tamaño de β .

Caso III: $a + \delta < 1$, $(1 - a - \delta)^2 - 8a\delta \leq 0$.

Para este caso, se puede probar [9], que una condición necesaria y suficiente para que se cumplan las desigualdades anteriores es que $\delta_1 < \delta < 1 - a$, donde $\delta_1 = 1 + 3a - \sqrt{8a^2 + 8a} > 0$. Del inciso i), E^* es asintóticamente estable globalmente. Por lo tanto, cuando la razón $\frac{s}{r}$ no es demasiado pequeña, la presa y el depredador se aproximan al punto de equilibrio, también independientemente de β .

Caso IV: $a + \delta < 1$, $(1 - a - \delta)^2 - 8a\delta > 0$.

En este caso, la variabilidad de las constantes da pie a tres subcasos, en los dos primeros se concluye que el sistema, aunque es asintóticamente estable globalmente, no hay el surgimiento de algún ciclo límite, mientras que en el tercero sí. Por razones de espacio se deja al lector que consulte la bibliografía indicada para esta sección.

7.4 Estabilidad en el modelo Tipo III

Sea ahora el sistema Tipo III adimensionalizado que se muestra en la sección anterior y se escribe a continuación:

$$\begin{aligned}\frac{dx}{dt} &= x(1-x) - \frac{x^2}{(a+x)(b+x)}y \\ \frac{dy}{dt} &= y\left(\delta - \beta\frac{y}{x}\right),\end{aligned}\tag{15}$$

$x(0) > 0, y(0) > 0$.

Análogamente al sistema anterior, se escribe el siguiente lema para el punto de equilibrio con coordenadas positivas.

Lema 7.5. *El punto de equilibrio $E^* = (x^*, y^*)$ de (15) es asintóticamente estable localmente si $P(x^*) > 0$ y es un foco o un nodo inestable si $P(x^*) < 0$, donde $P(x) = 2x^3 + (a+b-1+\delta)x^2 + (a+b)\delta x + ab(1+\delta)$.*

Para esta misma función y el mismo punto de equilibrio, en analogía al sistema Tipo II, se tiene el siguiente teorema:

Teorema 7.6.

- i) Si $P(x) \geq 0$ para todo $0 \leq x^* \leq 1$, entonces el punto de equilibrio $E^* = (x^*, y^*)$ es asintóticamente estable globalmente en el interior del primer cuadrante.*
- ii) Supóngase que $P(c_2) < 0$ y $\alpha_2 < x^* < \alpha_1$, entonces se cumple la conclusión de i).*
- iii) Supóngase que $P(c_2) < 0$. Para $\beta > 0$ suficientemente pequeño, $x^* = x^*(\beta)$ es suficientemente cercano a cero y se cumple $0 < x^* < \alpha$. Además se cumple la conclusión de i) para $\beta > 0$ suficientemente pequeño.*
- iv) Si $\alpha_1 < x^* < \alpha_2$, entonces existe un ciclo límite de (15).*

Observación: En el teorema anterior, las constantes c_1 y c_2 se supone que son raíces positivas de la ecuación algebraica $P'(x) = 0$. Las correspondientes α_1 y α_2 son dos raíces positivas de $P(x) = 0$ satisfaciendo $0 < c_1 < \alpha_1 < c_2 < 1$ y tal que $P(x)$ se pueda escribir como $P(x) = 2(x + \alpha_0)(x - \alpha_1)(x - \alpha_2)$, $\alpha_0 > 0$.

8 Bifurcación en algunos modelos

De acuerdo a los resultados indicados en los dos últimos teoremas de la sección anterior, particularmente en los incisos iv) de cada uno; nos aseguran que los sistemas presentan un ciclo límite para valores adecuados del parámetro, así que es posible el surgimiento de una bifurcación de Hopf. Además, se puede Verificar en [10] que para el sistema (15), cuando $a, \delta > 0$ las condiciones para las constantes impuestas en el teorema anterior se traducen en términos de la constante β en la siguiente:

$$\beta > \beta_2, \quad 0 < \beta < \beta_1, \quad \beta_1 < \beta < \beta_2,$$

donde

$$\beta_i = \frac{\delta \alpha_i}{(1 - \alpha_i)(a + \alpha_i)}, \quad i = 1, 2,$$

lo que nos garantiza que se puede ocupar a β como el parámetro de bifurcación. En esta sección se presenta un resultado de la bifurcación de Hopf para el sistema Tipo II, extraído de [10].

8.1 Bifurcación para el sistema Tipo II

Ya se explicó que existe un ciclo límite para el sistema (15). Ahora, si se considera $y = ul(x)$ y $l(x) = \left(\frac{1-x}{x}\right)^\delta$, el sistema (15) se transforma en el sistema más general:

$$\begin{aligned} \frac{dx}{dt} &= x(1-x) - \frac{x}{a+x} \frac{u}{l(x)} \\ \frac{du}{dt} &= \frac{u^2 \beta}{xl(x)(1-x)(a+x)} \left(x + \frac{a}{x^*}\right) (x - x^*), \end{aligned} \tag{16}$$

$x(0) > 0, u(0) > 0$.

Si ahora se consideran las isoclinas dadas por la ecuación $h(x) = (1-x)(a+x)l(x)$, entonces el sistema (16) pasa a ser un nuevo sistema:

$$\begin{aligned} \frac{dx}{dt} &= \varphi(x)[h(x) - u] = F(x, y) \\ \frac{dy}{dt} &= \psi(x)u^2 = G(x, y), \end{aligned} \tag{17}$$

$x(0) > 0, u(0) > 0$ y donde

$$\varphi(x) = \frac{x}{a+x} \frac{1}{l(x)}, \quad \psi(x) = \frac{\beta}{xh(x)} (x - x^*) \left(x + \frac{a}{x^*}\right).$$

En el siguiente teorema sobre la bifurcación del sistema (15) que es equivalente al sistema (17) se utiliza al parámetro β como el parámetro de bifurcación. Este resultado considera el fenómeno de la bifurcación para β cercano a β_1 . También notemos que x^* se determina por una función de β ; es decir, que $x^* = x^*(\beta)$, y $\lim_{\beta \rightarrow \beta_1} x^* = \alpha_1$.

Teorema 8.1.

- i) $K > 0$ para $0 < a < 1$ y $0 < \delta < \widehat{\delta}(a)$. La solución periódica de la bifurcación de Hopf de (17) es repelente con $\beta(\varepsilon) < \beta_1$, $|\varepsilon| \leq \varepsilon_0$, $\varepsilon \neq 0$; esto es, la bifurcación de Hopf es subcrítica.
- ii) $K < 0$ para $0 < a < 1$ y $\widehat{\delta}(a) < \delta < \delta^*(a)$. La solución periódica de la bifurcación de Hopf de (17) es atrayente con $\beta(\varepsilon) > \beta_1$, $|\varepsilon| \leq \varepsilon_0$, $\varepsilon \neq 0$; esto es, la bifurcación de Hopf es supercrítica.

En el teorema se ha usado la constante K con la expresión:

$$K = \frac{1}{2\pi} \int_0^{2\pi} C_4(\theta) d\theta - \frac{1}{\sqrt{\rho(\beta_1)}} \frac{1}{2\pi} \int_0^{2\pi} C_3(\theta) D_3(\theta) d\theta$$

donde:

$$\begin{aligned} C_3(\theta) &= \cos(\theta) B_2^1(\cos(\theta), \sin(\theta)) + \sin(\theta) B_2^2(\cos(\theta), \sin(\theta)), \\ D_3(\theta) &= -\sin(\theta) B_2^1(\cos(\theta), \sin(\theta)) + \cos(\theta) B_2^2(\cos(\theta), \sin(\theta)), \\ C_4(\theta) &= \cos(\theta) B_3^1(\cos(\theta), \sin(\theta)) + \sin(\theta) B_3^2(\cos(\theta), \sin(\theta)). \end{aligned}$$

Los demás componentes se obtienen de aplicar el teorema de Andronov Hopf, suponiendo que para el sistema (17) existe una transformación que lo reduce al sistema lineal:

$$X' = A(\mu)X + F(X, \mu),$$

con la matriz y la función como:

$$A(\mu) = \begin{pmatrix} \alpha(\mu) & -\beta(\mu) \\ \beta(\mu) & \alpha(\mu) \end{pmatrix}$$

y

$$F(X, \mu) = \begin{pmatrix} B_2^1(x_1, x_2, \mu) & B_3^1(x_1, x_2, \mu) + O(|x|^4) \\ B_2^2(x_1, x_2, \mu) & B_3^2(x_1, x_2, \mu) + O(|x|^4) \end{pmatrix}.$$

Los demás componentes y constantes que se encuentran presentes en el teorema están explícitos en la bibliografía ya indicada.

8.2 Bifurcación en un modelo con estructura sexual

En esta parte se presenta un sistema de tres ecuaciones diferenciales ordinarias con el fin de describir el modelo depredador-presa con estructura sexual, los correspondientes resultados y sus demostraciones se pueden encontrar en [11]. El modelo que se estudia ahí es:

$$\begin{aligned}\frac{dm}{dt} &= b_1 f - d_1 m - \sigma c_1 m x \\ \frac{df}{dt} &= f(\beta - c_1 x) \\ \frac{dx}{dt} &= x(-a + \sigma c_2 m + c_2 f),\end{aligned}\tag{18}$$

donde σ es el coeficiente de favoritismo sexual, $\sigma > 1$ significa que el depredador prefiere depredar presas masculinas a presas femeninas, $0 < \sigma < 1$ significa que el depredador prefiere presas femeninas a presas masculinas, y $\sigma = 1$ significa que no hay favoritismo sexual.

El sistema tiene puntos de equilibrio de la forma:

$$E^* = (m^*, f^*, x^*) = \left(\frac{ab_1}{c_2(d_1 + \sigma b_1 + \sigma\beta)}, \frac{a(d_1 + \sigma\beta)}{c_2(d_1 + \sigma b_1 + \sigma\beta)}, \frac{\beta}{c_1} \right).$$

A continuación se presenta un teorema que relaciona el sistema indicado.

Teorema 8.2. *El punto de equilibrio positivo E^* del sistema (18) es asintóticamente estable localmente cuando $\sigma > 1$, e inestable cuando $0 < \sigma < 1$. Existe un valor crítico, $\sigma_0 = 1$, tal que ocurre una bifurcación de Hopf en $\sigma = \sigma_0$ para σ decreciendo, esto es, existe una órbita periódica no trivial del sistema anterior si $\sigma \in (\sigma_0 - \varepsilon, \sigma_0)$.*

Por otra parte, el fenómeno de la bifurcación se presenta a menudo en modelos que presentan retardo, este retardo aparece en algunos modelos biológicos. Supongamos que la reproducción del depredador después de depredar la presa no es instantánea y requiere algún tiempo discreto de retardo para su gestación. Un modelo ya estudiado en donde aparece un retardo es:

$$\frac{dm}{dt} = b_1 f - d_1 m - \sigma c_1 m x\tag{19}$$

$$\frac{df}{dt} = f(\beta - c_1 x)\tag{20}$$

$$\frac{dx}{dt} = -xa - \sigma c_2 m(t - \tau)x(t - \tau) + c_2 f(t - \tau)x(t - \tau),$$

donde τ , ($\tau > 0$) es el periodo de gestación de la presa.

La aparición de un retardo en el sistema de ecuaciones, da origen al estudio de un problema más difícil puesto que los espacios en donde se realiza el análisis son de dimensión infinita, y el estudio de estos espacios por su complejidad, quedan fuera del alcance de este escrito.

Bibliografía

- [1] Braun, M. , *Ecuaciones Diferenciales y sus Aplicaciones*, Grupo Editorial Iberoamérica. 1990.
- [2] Berryman, A.A. *The Origins and Evolution of Predator-Prey Theory*, Ecology, **73** (1992), 1530–1535.
- [3] Boyce, W.E. y DiPrima, R.C. *Ecuaciones Diferenciales y Problemas con Valores en la Frontera*, Limusa, Cuarta Edición. 2007.
- [4] Sandy Gómez Pérez. *Estabilidad por Aproximación de Sitemas Leslie-Holling*. Tesis para obtener la Licenciatura en Matemáticas, Facultad de matemáticas, Universidad Veracruzana, 2014.
- [5] Huang, J., Ruan, S., y Song, J. *Bifurcations in a Predator-Prey System of Leslie Type with Generalized Holling type III Functional Response*, Elsevier Journal of Differential Equations, **257** (2013), 1721–1752.
- [6] Huang, J., y Xiao, D. *Analysis of Bifurcations and Stability in a Predator-Prey System with Holling Type-IV Functional Response*, Acta Math. Appl., Sin. Engl., **20**, (2004), 167-178.
- [7] Morris W. Hirsh, Stephen Smale and Robert L. devaney. *Differentia Equations, Dynamical Systems. A. Introduction to Chaos*. Elsevier, Academic Press. USA. 2004.
- [8] Murray J. D. *Mathematical Biology* Interdisciplinary Applied mathematics. Third Edition, Springer Verlag, USA. 2002.
- [9] Sze-Bi Hsu and Tzy-Wei Hwang. *Global Stability for a Class of Predator-Prey Systems*. SIAM J. Appl. Math. **3**, 51 (1995), 763–783.
- [10] Sze-Bi Hsu and Tzy-Wei Hwang. *Hopf Bifurcation Analysis for a Depredator-Prey System of Holling and Leslie Type*. Taiwanese Journal of Mathematics, **3**, **1**, (1999), 35–53.

- [11] Shunyi Li and Zuoliang Xiong *Bifurcation Analysis of a Predator-Prey System with sex-structure and sexual favoritism*. Advances in Difference Equations, a Springer Open Journal. (2013), 1–24.
- [12] T. K. Kar, H. Matsuda; *Global Dynamics and Controllability of a Harvested Prey-Predator System with Holling Type III Functional Response* Nonlinear Analysis: Hybrid Systems. 2007.

Facultad de Matemáticas, Universidad Veracruzana
Lomas del Estadio S/N, Zona Universitaria,
Xalapa Veracruz, México.
evmunoz@uv.mx
fsergios@gmail.com

Filosofía, Divulgación e Historia de las Matemáticas

Capítulo 3

Una aproximación geométrica al cálculo diferencial

José Juan Angoa Amador, Fernando Cocoltzi Adame
FCFM BUAP

Resumen

En este trabajo desarrollamos algunas ideas que formalmente pueden sustituir al concepto de límite; con la ventaja adicional de que son más intuitivas. Cuando una definición matemática apela a protagonistas geométricos, en general se cree que es más intuitiva, ya que la imagen es más sugerente que el concepto. La motivación principal de este trabajo se encuentra en [3], pero la sistematización y organización de los principales resultados es obra de los autores. Este trabajo puede ampliarse hasta constituir una visión integral del cálculo diferencial. Tal vez esto pueda generar algunos discursos alternativos en la enseñanza del cálculo, tal vez sólo sea una información recreativa para los matemáticos; cualquiera que sea el caso, esperamos divierta al lector.

1 Preliminares

Para leer este texto bastan los conocimientos adquiridos tras la lectura del libro *Matemáticas Elementales* [2] o en general un curso de matemáticas básicas universitarias. Es recomendable, a manera de comparación, seguir una presentación clásica de lo que aquí se presenta, por ejemplo en [1].

Una de las ideas fundamentales de las matemáticas es la noción de conjunto. De hecho, el proceso mediante el cual se encuentran *relaciones* entre objetos, y por tanto, se agrupan en colecciones, constituye una de las bases esenciales sin las cuales sería imposible el pensamiento humano.

A lo largo de este escrito, se empleará la notación usual de la teoría elemental de los conjuntos. Así, $A \subseteq B$ significa que el conjunto A es subconjunto del conjunto B , es decir, que cada elemento del conjunto A es también elemento del conjunto B . La igualdad entre dos conjuntos A y B ocurre cuando y solo cuando $A \subseteq B$ y $B \subseteq A$.

Existen métodos elementales de creación de conjuntos a partir de otros. Los más conocidos se muestran sucintamente y sin mayor explicación.

Si A y B son dos conjuntos, se definen:

- i) El conjunto unión de A y B , $A \cup B = \{x \mid x \in A \text{ o } x \in B\}$. En este caso, la disyunción “o” significa que el elemento puede estar en uno de los dos conjuntos, en el otro o en ambos.
- ii) El conjunto intersección de A y B , $A \cap B = \{x \mid x \in A \text{ y } x \in B\}$.
- iii) El conjunto diferencia de A y B , $A \setminus B = \{x \mid x \in A \text{ y } x \notin B\}$.
- iv) El conjunto vacío se denota por \emptyset .

1.1 Producto Cartesiano

Las operaciones elementales con conjuntos, es decir, formas de combinar elementos de algunos conjuntos para formar otros, son usualmente insuficientes para las aplicaciones. Conviene introducir el concepto de *Producto Cartesiano* de dos conjuntos.

Definición 1.1. Dados dos conjuntos no vacíos A y B , se define el **Producto Cartesiano** de A por B como:

$$A \times B = \{(a, b) \mid a \in A \text{ y } b \in B\}.$$

Donde se acepta que

$$(a, b) = (c, d) \text{ si y solo si } a = c \text{ y } b = d.$$

Definición 1.2. Sean X e Y conjuntos no vacíos. Una **relación** de X sobre Y es cualquier subconjunto de $X \times Y$.

Nombres especiales se reservan para cuando estos subconjuntos son el conjunto vacío o $X \times Y$ completo. A la primera relación se le llama nula y a la segunda, universal.

Para referirse a los elementos que pueden figurar en las primeras o segundas componentes de algún par en la relación, se emplean los siguientes conceptos.

Definición 1.3. Sean X e Y conjuntos y R una relación de X sobre Y .

- El **Dominio** de la relación es el conjunto

$$Dom(R) = \{a \in X \mid \exists b \in Y \text{ con } (a, b) \in R\}$$

- La **Imagen** de la relación es el conjunto

$$Im(R) = \{b \in Y \mid \exists a \in X \text{ con } (a, b) \in R\}$$

... Si $a \in \text{Dom}(R)$, se denotará por $R[a] = \{y \in Y \mid (a, y) \in R\}$

Asimismo, si $(a, b) \in R$, se dirá que “ a está relacionado con b mediante la relación R ”.

El concepto de relación es extremadamente fructífero, y aquí se empleará para definir otro concepto importante: el de función.

Definición 1.4. Sean X e Y dos conjuntos no vacíos. Una relación F de X sobre Y se llama una **función de X en Y** si y solo si satisface:

- $\text{Dom}(F) = X$
- .. Siempre que $(x, y) \in F$ y $(x, z) \in F$, entonces $y = z$.

Como es usual, se denotará a la función F junto con los conjuntos X e Y de la definición de la siguiente manera $F : X \rightarrow Y$.

Note el lector que cada que se tome un elemento, $a \in X$, al ser F una función, $a \in \text{Dom}(F)$ y existe un único par $(a, b) \in F$ con a en su primera componente. Puede entonces decirse sin ambigüedad que b es la imagen de a bajo F (o que b es el valor de F en a), escrito mediante la identidad $b = F(a)$.

Note también que para que dos funciones f y g sean iguales, es necesario y suficiente que sean iguales sus dominios y que $\forall a \in \text{Dom}(f) : f(a) = g(a)$, es decir que cada función está determinada por la correspondencia $a \mapsto f(a)$.

La notación para la imagen de un elemento bajo una función es muy útil para definir con simplicidad una operación entre funciones de gran importancia.

Definición 1.5. Sean $f : X \rightarrow Y$ y $g : Y \rightarrow Z$ dos funciones. Se define la **composición** de f con g , $g \circ f : X \rightarrow Z$ por: $\forall a \in X : (g \circ f)(a) = g(f(a))$. El símbolo $g \circ f$ se lee como f seguida de g o también como g enseguida de f .

La definición dada para una función refleja la idea de una regla que asigna a cada elemento de un conjunto otro elemento de otro conjunto (sin excluir la posibilidad de que sean los mismos).

El concepto de función es tan general que no parece posible poder clasificarlas a todas; sin embargo, un primer criterio de interés es el siguiente:

Definición 1.6. Sea $f : X \rightarrow Y$ una función. Se dice que:

- f es **inyectiva** si y solo si elementos distintos de X tienen distintas imágenes bajo f .
- .. f es **suprayectiva** si y solo si $\text{Im}(f) = Y$.
- ... f es **biyectiva** si y solo si es suprayectiva e inyectiva.

1.2 Los números reales

Una buena parte de las matemáticas hace uso de conjuntos dotados de operaciones aritméticas elementales, llamados campos. Un primer ejemplo de campo es el conjunto de números reales, cuyas operaciones de suma (representada por $+$) y producto (representado por \cdot aunque usualmente se omite) satisfacen las siguientes propiedades.

- Se cumple que la suma y el producto son asociativos y conmutativos, pues para cualesquiera elementos $a, b, c \in A$, $(a+b)+c = a+(b+c)$, $(ab)c = a(bc)$, $a+b = b+a$ y $ab = ba$.
- Existen elementos neutros para la suma y el producto, el 0 y el 1 (diferentes entre sí), tales que, para cualquier $a \in A$: $0+a = a$ y $1a = a$
- Para cualquier elemento $b \in A$, existe otro $c \in A$ tal que $b+c = 0$. Además, si $b \neq 0$, también existe otro elemento $d \in A$ tal que $bd = 1$.
- El producto se distribuye respecto a la suma, es decir que para cualesquiera $a, b, c \in A$: $a(b+c) = ab+ac$.

El lector sabe bien que el hecho de poder resolver, conociendo a , la ecuación $a+b = 0$ y $ac = 1$ si a es distinta de 0, permite definir la resta y división. La notación en estos casos es $b = -a$ (b es el opuesto de a) y $c = a^{-1}$ (c es el inverso de a).

El campo de los números reales además de ser un sistema numérico, posee un orden, puesto que existe en él un conjunto de números **positivos** que satisfacen lo siguiente:

- Todo número real es positivo, o su opuesto es positivo, o es el 0. Cumpliéndose solo una de las tres posibilidades.
- El resultado de operar dos números positivos mediante la suma y el producto es también positivo.

Es posible probar que un número es positivo si y solo si es mayor que cero, aceptando la siguiente definición de *ser mayor que*: se dice que a es mayor que b si y solo si $a - b$ es positivo.

El aspecto más significativo del sistema de números reales (en adelante denotado por \mathbb{R}) es que es completo. La afirmación anterior encuentra su formulación precisa a través del siguiente axioma:

Axioma del Supremo. Dado un subconjunto no vacío de los números reales, A , para el que existe un número $a \in \mathbb{R}$ tal que $A \subseteq]-\infty, a]$, existe un número mínimo b con la propiedad de que $A \subseteq]-\infty, b]$, es decir cualquier otro con esta propiedad es mayor o igual a b .

Una forma de exhibir el sentido de la frase “ \mathbb{R} es completo” es refiriéndose a ellos como el *continuo de los números reales*, motivados en la convención familiar de asociar un único número real a cada uno de los puntos de una línea recta, infinita y que se extiende en sus dos sentidos.

Al aceptar esta correspondencia uno a uno entre los números reales y los puntos de una recta, resulta natural tomar prestada la terminología geométrica para referirse a los elementos de \mathbb{R} . Por ejemplo, a cada número real se le suele llamar *punto* en \mathbb{R} . Un número menor que otro se dice que está a su izquierda, mientras que el mayor se dice a la derecha del menor. Si para tres números $a, b, c \in \mathbb{R}$ se cumple que $a < b < c$, se dice que b está entre a y c .

Los segmentos de recta y rayos son subconjuntos de puntos importantes de una recta y una notación especial se reserva para ellos en términos de números reales.

Definición 1.7. Sean $a, b \in \mathbb{R}$ con $a \leq b$.

$$\begin{aligned} [a, b] &= \{x \in \mathbb{R} | a \leq x \leq b\} &]a, b[&= \{x \in \mathbb{R} | a < x < b\} \\]a, b] &= \{x \in \mathbb{R} | a < x \leq b\} & [a, b[&= \{x \in \mathbb{R} | a \leq x < b\} \\]a, \infty[&= \{x \in \mathbb{R} | a < x\} &]-\infty, b] &= \{x \in \mathbb{R} | x \leq b\} \end{aligned}$$

El resto de las posibles combinaciones se definen de manera análoga y a veces al mismo \mathbb{R} se le denota con $] - \infty, \infty [$.

Sin mayores pretensiones que el recuperar una intuición geométrica del lector con respecto a los números reales, se considerará que estos y los puntos de una recta tipo denominada **recta real** son lo mismo.

1.3 Funciones reales de variable real

En adelante se tratará únicamente con un tipo especial de funciones llamadas *funciones reales de variable real*. Se les dice funciones reales porque su imagen es algún subconjunto de \mathbb{R} y de variable real porque el dominio de la función es también un subconjunto de \mathbb{R} .

Una manera usual de representar que una función, f es real de variable real es $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$, indicando que el dominio es el conjunto A y que las imágenes de la función en todos los puntos de A son algunos números reales.

Sabiendo que los resultados de las operaciones aritméticas entre números reales ofrecen resultados únicos, es común encontrar funciones que especifican como obtener la imagen de cada uno de los elementos del dominio mediante operaciones aritméticas simples.

Claro está también que para definir una función no siempre se necesita (y a veces ni es posible construir) una regla aritmética para hallar el valor de la función en cada punto. Por ejemplo, es posible definir una función mediante el siguiente proceso: a cada número, x , se le asigna como pareja al número obtenido por sustituir todas las cifras del desarrollo decimal infinito de x después del primer 4 u 8 por un 3.

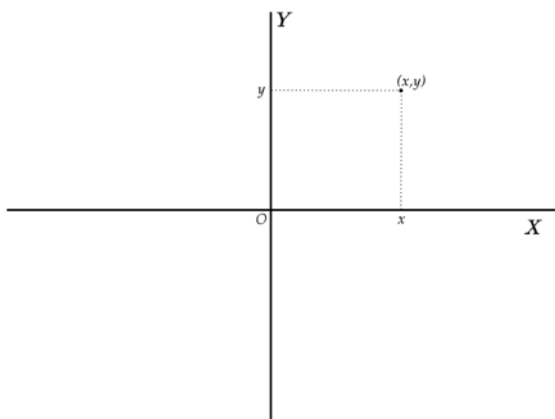


Figura 1: Porción del plano cartesiano.

Así pues, funciones reales de variable real y relaciones de \mathbb{R} sobre \mathbb{R} pueden considerarse como las porciones del plano cartesiano que representan y puede hablarse de ellos indistintamente como conjuntos de puntos en el plano o como las reglas de correspondencia que permiten obtener la segunda componente de un par a partir de la primera. Nuevamente se toma prestado el lenguaje de la geometría analítica y se identifican una función y su gráfica en el plano.

El ejemplo más importante es quizá el de las rectas en el plano (se dedicará toda una sección a su tratamiento), determinadas por un punto (x_0, y_0) y una pendiente m . El conjunto $L = \{(x, y) \in \mathbb{R}^2 \mid y = m(x - x_0) + y_0\}$ se identifica la función $f : \mathbb{R} \rightarrow \mathbb{R}$ definida por $f(x) = m(x - x_0) + y_0$.

La definición de función dada al principio en términos de parejas ordenadas, aunque extraña de inicio, ahora muestra una enorme bondad al considerar a las funciones reales de variable real, puesto que, así como los números reales se identifican con una recta, es usual asociar a los puntos de un plano con parejas de números reales mediante el sistema coordenado rectangular cartesiano.

Más aún, según la disposición usual de los ejes en el plano cartesiano con el eje horizontal *de izquierda a derecha* y el vertical *de abajo a arriba* (ver figura 1), puede cumplirse para dos funciones que el valor de una de ellas en un punto sea mayor que el de la otra y así, se dice que la primera *está por arriba* de la segunda o que la segunda *está por abajo* de la primera en el punto. Esto se extiende a su vez a conjuntos de puntos en \mathbb{R} diciendo por ejemplo que en algún conjunto una función (o relación) *está por arriba* de otra, o entre otras dos, etc.

Definición 1.8. Sean R y P dos relaciones de \mathbb{R} en \mathbb{R} . Sean $A \subseteq \text{Dom}(R) \cap \text{Dom}(P)$

y $x_0 \in A$. Se dice que R *está por arriba o encima* de P en el punto x_0 si cualquier elemento de $R[x_0]$ es mayor que cualquier elemento de $P[x_0]$ (ver la Definición 1.3). Se dice que R *está por arriba o encima* de P en A si R está por arriba de P en todo punto de A .

En el caso de las funciones, la situación es más sencilla:

Observación 1.9. Sean f y g dos funciones reales. Sean $A \subseteq \text{Dom}(f) \cap \text{Dom}(g)$ y $x_0 \in A$. f *está por arriba o encima* de g en el punto x_0 si y solo si $f(x_0) > g(x_0)$. Así, f *está por arriba* de g en A si y solo si para toda x en A , $f(x) > g(x)$.

1.4 Operaciones con funciones reales de variable real

Así como se operan números para obtener nuevos números, se pueden operar funciones reales de variable real (en adelante simplemente funciones o funciones reales) para obtener nuevas funciones.

Definición 1.10. Sean $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ y $g : B \subseteq \mathbb{R} \rightarrow \mathbb{R}$ dos funciones. Se definen:

- i)* La **suma** de f y g como $f + g : A \cap B \rightarrow \mathbb{R}$, $f(x) = f(x) + g(x)$
- ii)* El **producto** de f y g como $fg : A \cap B \rightarrow \mathbb{R}$, $f(x) = f(x) \cdot g(x)$
- iii)* La **resta** de f menos g como $f - g : A \cap B \rightarrow \mathbb{R}$, $f(x) = f(x) - g(x)$
- iv)* Si c es un número real, se define $cf : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = c \cdot f(x)$

Para poder realizar un cociente de funciones, se requiere que el valor que fungirá como denominador no sea 0, así pues, deben retirarse del dominio los puntos, x , tales que $f(x) = 0$.

- i)* Un punto $x_0 \in A$ se llama **cero** de la función (y que la función se anula en x_0) si $f(x_0) = 0$
- ii)* El **cociente** de f entre g como $f/g : A \cap (B \setminus \{x \in B \mid x \text{ es cero de } f\}) \rightarrow \mathbb{R}$, $x \mapsto f(x)/g(x)$.

Funciones especiales son las siguientes:

Definición 1.11.

- La función $I_A : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto x$, se llama función **identidad** en A .
- Sea $c \in \mathbb{R}$ un número fijo. La función $c : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto c$, se llama función **constante**.

2 Funciones lineales

Aún cuando no se tiene una definición precisa de *plano*, se ha aceptado una identificación entre los conjuntos de parejas de números reales y los puntos del mismo. Las rectas jugarán un papel muy importante en el desarrollo de este trabajo, ya que algunas se pueden entender como la gráfica de una función o como una función. Formalmente:

Definición 2.1. Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ una función. Si existen números fijos $m, b \in \mathbb{R}$ tales que $\forall x \in \mathbb{R} : f(x) = mx + b$, entonces a tal función se le llamará una **función lineal** o recta no vertical. Al número m se le llama la **pendiente** de la recta y al número b se le llama **ordenada al origen**.

Los números m y b de la definición caracterizan completamente a las rectas. El siguiente teorema indica en qué sentido.

Teorema 2.2. Sean f y g dos funciones lineales definidas por $f(x) = m_1x + b_1$ y $g(x) = m_2x + b_2$. Entonces, f y g son iguales si y solo si $m_1 = m_2$ y $b_1 = b_2$.

Demostración. Cuando $m_1 = m_2$ y $b_1 = b_2$ la situación es trivial. Suponiendo que las funciones lineales son iguales, entonces $\forall x \in \mathbb{R} : m_1x + b_1 = m_2x + b_2$. En particular $m_1 \cdot 0 + b_1 = m_2 \cdot 0 + b_2$ y así $b_1 = b_2$. Se tiene ahora que $m_1x = m_2x$, y evaluando en $x = 1$, $m_1 = m_2$. \square

En estos términos, la pendiente de una función lineal proporciona información sobre su inclinación con respecto al eje X . Por ejemplo, una recta con pendiente 0 en realidad es una función constante y por tanto una recta horizontal. Cuando las pendientes son diferentes de cero, las funciones lineales ya no son constantes; sin embargo, presentan una *estricta monotonía*, entendiéndose por eso que no ocurre que sus gráficas presenten cimas y valles y que su crecimiento es estricto.

Teorema 2.3. Sea f una función lineal no constante, definida por $f(x) = mx + b$. La pendiente m es positiva si y solo si f es **estrictamente creciente** en \mathbb{R} , es decir, si $x < y$, entonces $f(x) < f(y)$. La pendiente m es negativa si y solo si f es **estrictamente decreciente** en \mathbb{R} , es decir, si $x < y$, entonces $f(x) > f(y)$.

Demostración. Sea f una recta con pendiente m positiva. $x < y$ implica que $mx < my$ y así $mx + b < my + b$, se concluye que $f(x) < f(y)$. Recíprocamente, si $x < y$ implica que $f(x) < f(y)$, entonces $y - x > 0$ y $my + b - mx - b > 0$. De aquí que $m(y - x) > 0$, y por tanto $m > 0$. El caso $m < 0$ es análogo. \square

Las funciones lineales no constantes, a pesar de su sencillez, poseen una propiedad extraordinaria llamada *propiedad de los valores intermedios*. Para ejemplificar, considere a la función identidad. Según la definición se trata de una función lineal con pendiente 1. Dados dos números x, y con $x < y$, $f(x) = x < y = f(y)$, puede tomarse un tercer número z entre $f(x)$ y $f(y)$ y, sin importar cuál sea ese número z , se cumple que $f(z) = z$. Por lo que $f(x) < f(z) < f(y)$ con $x < z < y$. Es decir, la recta *alcanza todos los valores posibles entre dos diferentes*.

Teorema 2.4. *Sea f una función lineal estrictamente creciente. Sean $x, y \in \mathbb{R}$ con $x < y$. Si $z \in]f(x), f(y)[$, entonces existe $x_0 \in]x, y[$ tal que $f(x_0) = z$*

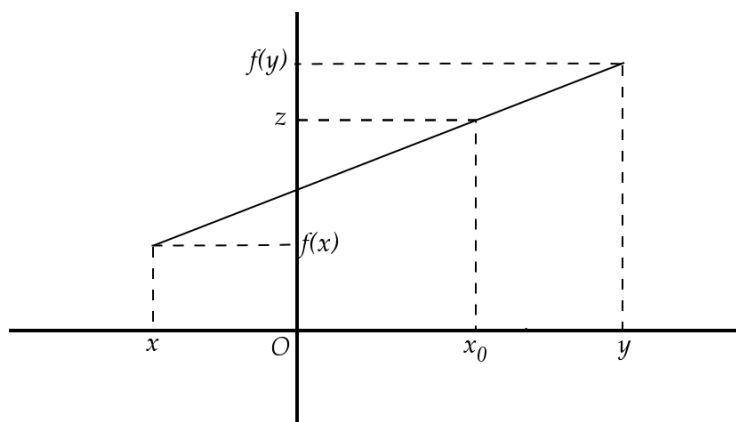


Figura 2: Existe $x_0 \in]x, y[$ con $f(x_0) = z$.

Demostración. La existencia del punto en que se alcanza z se obtiene de un despeje, puesto que $z = f(x_0)$ ocurre cuando $mx_0 + b = z$, que es equivalente a $x_0 = \frac{z-b}{m}$. Más aún, hasta ahora, la posición de z resultó irrelevante y el valor x_0 tal que $f(x_0) = z$ es único.

Como $z \in]f(x), f(y)[$, entonces $\frac{f(x)-b}{m} < \frac{z-b}{m} < \frac{f(y)-b}{m}$, así que $x < x_0 < y$. \square

El análogo del teorema anterior para rectas decrecientes es el siguiente:

Teorema 2.5. *Sea f una función lineal decreciente. Sean $x, y \in \mathbb{R}$ con $x < y$. Si $z \in]f(y), f(x)[$, entonces existe $x_0 \in]x, y[$ tal que $f(x_0) = z$*

2.1 Operaciones con funciones lineales

Al igual que como en el caso de las funciones reales de variable real, las funciones lineales pueden operarse para obtener nuevas funciones lineales.

Proposición 2.6. Sean f y g dos funciones lineales definidas por $f(x) = m_1x + b_1$ y $g(x) = m_2x + b_2$. Sea c un número real.

- i) $f + g$ es una función lineal de pendiente $m_1 + m_2$
- ii) cf es una función lineal de pendiente cm_1
- iii) $f \circ g$ es una función lineal de pendiente m_1m_2 .

Demostración. Sean f y g tales que $f(x) = m_1x + b_1$ y $g(x) = m_2x + b_2$.

$f + g$ manda $x \mapsto m_1x + b_1 + m_2x + b_2 = (m_1 + m_2)x + (b_1 + b_2)$.

cf manda $x \mapsto c(m_1x + b_1) = (cm_1)x + (cb_1)$.

$f \circ g$ manda $x \mapsto m_1(m_2x + b_2) + b_1 = (m_1m_2)x + (m_1b_2 + b_1)$. □

Y al igual que con números reales, pueden plantearse ecuaciones en términos de funciones lineales.

Proposición 2.7. Sean f y g funciones lineales arbitrarias. Existe una única recta no vertical h tal que $f + h = g$, y si f es no constante, existe otra función lineal única t tal que $t \circ f = g$.

Demostración. En el primer caso se busca que $\forall x \in \mathbb{R} : f(x) + h(x) = g(x)$, o equivalentemente $h(x) = g(x) - f(x)$. Es decir que la función buscada es única y es idéntica a $g - f$.

Para la composición, si f, g y t están definidas por $f(x) = m(x) + b$, $g(x) = nx + c$ y $t(x) = qx + p$, entonces $t \circ f = g$ se cumplirá si y solo si $(qm)x + (qb + p) = nx + c$ para cualquier x , si y solo si $qm = n$ y $qb + p = c$. Lo anterior es equivalente a que $q = \frac{n}{m}$ y $p = c - qb$, es decir, la recta es única y definida por $t(x) = \frac{n}{m}x + c - \frac{n}{m}b$. □

2.2 Intersecciones entre rectas

Cuando se tienen dos rectas en el plano, pueden o no existir intersecciones entre rectas. No es difícil convencerse de que el número de intersecciones está condicionado por la dirección de estas, puesto que por ejemplo, si ambas tienen la misma pendiente y son de hecho diferentes, la intersección entre ellas es el conjunto vacío.

Teorema 2.8. Sean f y g dos funciones lineales diferentes. La intersección de ellas contiene a lo más un único elemento. Además, dicha intersección es el conjunto vacío si y solo si las pendientes de las rectas son las mismas.

Demostración. Sean f y g definidas por $f(x) = mx + b$ y $g(x) = nx + c$ funciones lineales distintas. Un par (x_0, y_0) pertenece a la intersección si y solo si satisface ambas ecuaciones, es decir, si $mx_0 + b = y_0 = nx_0 + c$, que es equivalente a que

$(m - n)x_0 = c - b$. Si $m \neq n$ entonces $x_0 = \frac{c-b}{m-n}$ es el único punto que satisface ambas ecuaciones, es decir, es el único punto en la intersección. Si $m = n$, al ser f y g distintas, $c \neq b$ por el Teorema 2.2; pero $0 = c - b$, lo cuál es una contradicción obtenida de suponer que algún punto satisface ambas ecuaciones. \square

Es decir, dadas dos funciones lineales con pendientes diferentes entre sí, puede garantizarse la existencia de un único par (x_0, y_0) que pertenece a ambas. Tomando un punto a a la izquierda y otro b a la derecha de x_0 , las rectas determinan una región del plano en forma de *mariposa*.

Las regiones de esta forma tendrán un papel protagónico en el desarrollo de las ideas siguientes, puesto que fungirán como testigos de la posibilidad de “acercarse” al par (x_0, y_0) mediante una relación, siempre que este contenida dentro de ella.

Para definir estas regiones se requiere conocer a la intersección de ambas rectas, así que un primer paso será dar una expresión para las rectas dado el punto de intersección.

Teorema 2.9. *Sea f una recta que contiene a (x_0, y_0) y que tiene pendiente m , entonces $\forall x \in \mathbb{R} : f(x) = m(x - x_0) + y_0$, y f es única.*

Demostración. Sea f la función lineal que pasa por (x_0, y_0) con pendiente m . Entonces f está definida por la expresión $f(x) = mx + b$. Sabiendo que $y_0 = mx_0 + b$, entonces $b = y_0 - mx_0$, así que $f(x) = mx + y_0 - mx_0$ o equivalentemente $f(x) = m(x - x_0) + y_0$. La unicidad es inmediata. \square

Definición 2.10. Sean f y g dos funciones lineales definidas por $f(x) = m(x - x_0) + y_0$ y $g(x) = n(x - x_0) + y_0$ con $m > n$. Sean $a, b \in \mathbb{R}$ tales que $a < x_0 < b$. Se define el **ala izquierda** de la mariposa desde a entre f y g como el conjunto

$$M_a(f, g) = \{(x, y) \in \mathbb{R}^2 \mid x \in]a, x_0[\text{ y } f(x) < y < g(x)\}.$$

Se define el **ala derecha** de la mariposa hasta b entre f y g como el conjunto

$$M^b(f, g) = \{(x, y) \in \mathbb{R}^2 \mid x \in]x_0, b[\text{ y } g(x) < y < f(x)\}.$$

Por último, la **mariposa de a a b entre f y g** es el conjunto

$$M_a^b(f, g) = M_a(f, g) \cup M^b(f, g).$$

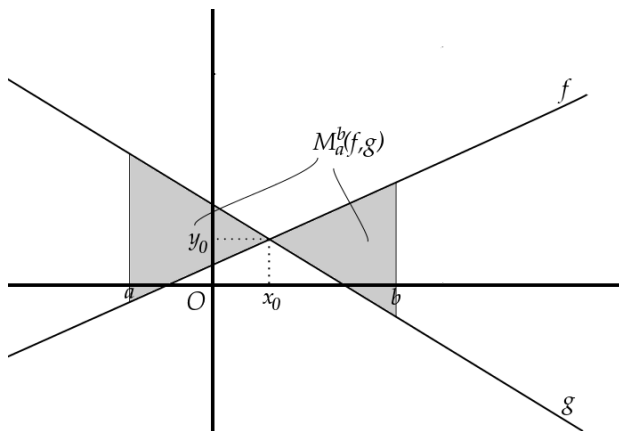


Figura 3: Región de mariposa.

3 Rebase

El hecho de poseer una identificación directa entre funciones y ciertas regiones del plano, provee de una posibilidad teórica de plantear problemas de naturaleza geométrica en términos relacionales. Por ejemplo, considérense a la relación *circunferencia unitaria*: $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ y a su *tangente* en el par $(0, 1)$, definida por $L(x) = 1$. Para obtener una función a partir de la circunferencia unitaria, se puede eliminar la porción de ella que está por debajo del eje de las abscisas, resultando en la función definida por $f(x) = \sqrt{1 - x^2}$, una semicircunferencia.

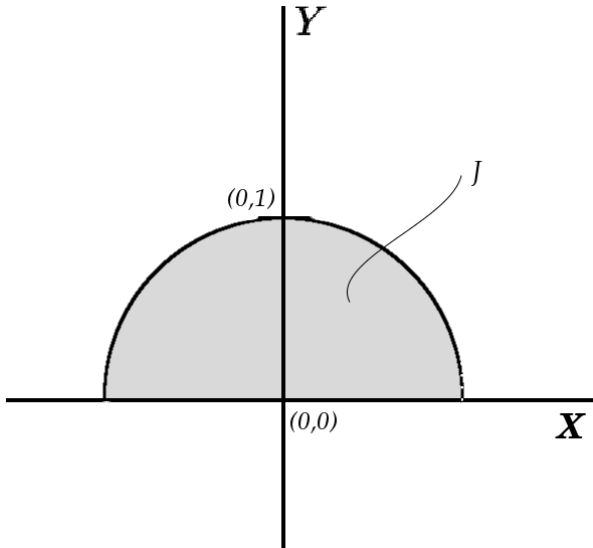
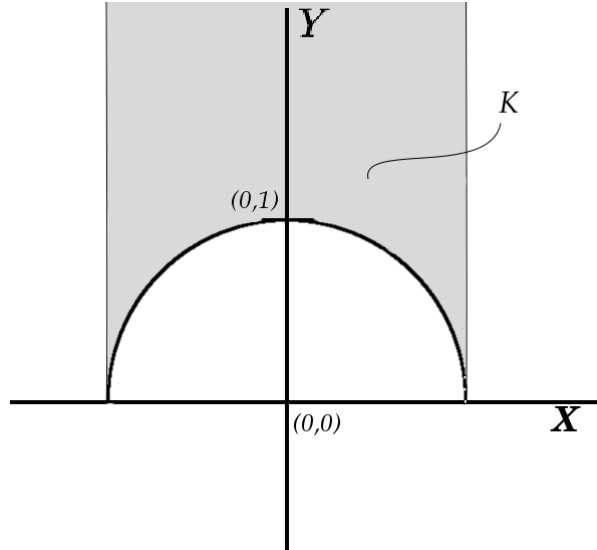
En general, para una circunferencia dada, las rectas que son tangentes a ella en algún par solo comparten con la misma el par de tangencia. Esta imagen es tan familiar que una frase compacta la tradición que se ha formado en torno a ella: “Recta tangente es aquella que *toca* a la curva sin *cortarla*”.

Euclides, desde aproximadamente el 300 a.C. hizo una observación interesante en sus *Elementos* (Libro III, proposición 16) referente a la tangencia a una circunferencia.

La recta trazada por el extremo de un diámetro de un círculo formando ángulos rectos (con el mismo) caerá fuera del círculo, *y no se interpondrá otra recta en el espacio entre la recta y la circunferencia ...*

Si se consideran todas las rectas que contienen al par $(0, 1)$, estas están determinadas por expresiones de la forma $L_m(x) = m \cdot x + 1$ ¹. La tangente en $(0, 1)$ corresponde al caso $L_0(x) = 0 \cdot x + 1 = 1$. Así que, según *Euclides*, deberíamos espe-

¹La gráfica de $L(x)$ contiene a $(0, 1) \Leftrightarrow L(x) = mx + b$ y $L(0) = 1 \Leftrightarrow 1 = b$ y $L(x) = mx + 1 \Leftrightarrow L(x) = L_m(x) = mx + 1$.

Figura 4: Representación de J .Figura 5: Representación de K .

rar que el resto de ellas no se encuentren entre la semicircunferencia, f y la tangente, L_0 . ¿Cuál es entonces su posición relativa con respecto a la circunferencia?

Para dar respuesta, primero considere el lector a la sección del círculo unitario que está por arriba del eje horizontal $J = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1 \text{ e } y \geq 0\}$ (figura 4) y note que se encuentra por debajo de f en $] -1, 1[$. Asimismo, considere la relación $K = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 > 1 \text{ y } -1 < x < 1\}$ (figura 5), observe que está por arriba de f en $] -1, 1[$ como también L_0 está por encima de f en $] -1, 1[\setminus\{0\}$.

Caso I. $m < 0$. Para que una pareja, (x, y) , de la recta definida por $L_m(x) = mx + 1$ se encuentre también en la circunferencia unitaria, debe cumplirse que

$$x^2 + y^2 = 1 \quad y = mx + 1$$

De este modo se obtiene

$$x^2 + (m^2x^2 + 2mx + 1) = 1,$$

resultando en

$$x^2(1 + m^2) + 2mx = 0 \tag{1}$$

que es equivalente a que $x = 0$ o $x = \frac{-2m}{1+m^2} (> 0)$.

Sea $c = \min\{1, \frac{-2m}{1+m^2}\}$. Si $x \in]0, c[$, entonces $x \in]0, \frac{-2m}{1+m^2}[$ y así $0 < x$ y $(1 + m^2)x + 2m < 0$, de donde $(1 + m^2)x^2 + 2mx < 0$ y $(1 + m^2)x^2 + 2mx + 1 < 1$,

por lo que $(x, mx + 1)$ está en J . Según las observaciones anteriores, en $]0, c[$, la recta está por debajo de f .

Por otro lado, siempre que $x \in]-1, 0[$, ocurre que $2mx > 0$ y así $(1 + m^2)x^2 + 2mx + 1 > 1$ y $(x, mx + 1) \in K$. Así se obtiene que en $]0, 1[$ la recta está por arriba de f .

Caso II. $m > 0$. Se puede hacer un análisis similar para concluir que en $] - 1, 0[$ la recta definida por $L_m(x) = mx + 1$ ($m > 0$) está por abajo de f , mientras que existe un número $d > 0$ tal que en $]0, d[$, f está por arriba de L_m . Comience notando que para encontrar las intersecciones de la recta y f , se debe resolver la ecuación (1).

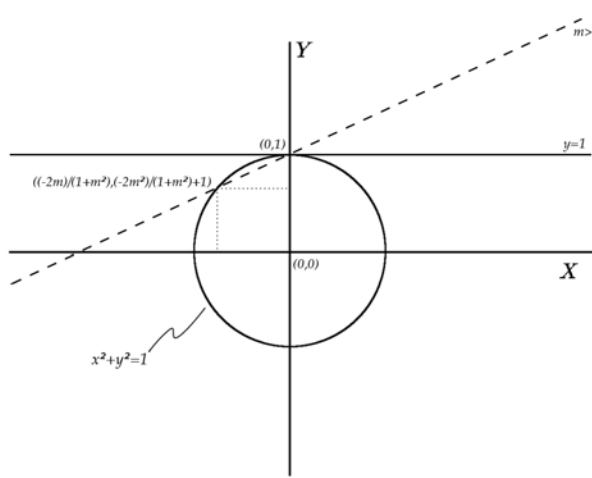


Figura 6: Pendiente positiva.

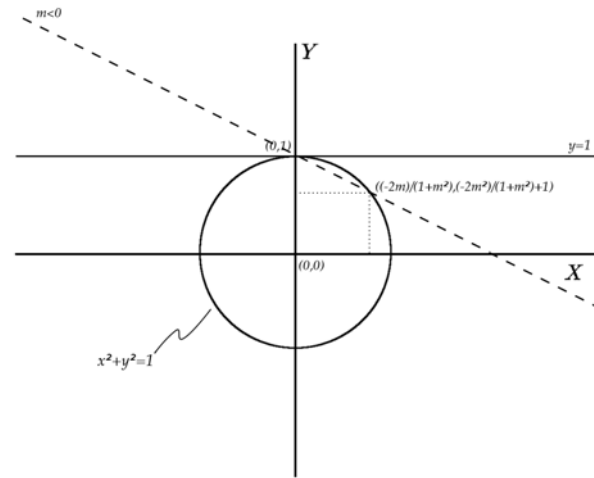


Figura 7: Pendiente negativa.

Se evidencia del análisis y con ayuda de las figuras 6 y 7, que cuando se toman rectas que pasan por un mismo punto de distintas pendientes, existe un cierto *rebase* entre las rectas y la circunferencia.

Para clarificar la noción de rebase, considere a dos rectas con pendientes diferentes y que pasan a través de un mismo punto². Se espera, como muestra la figura 8, que la recta con pendiente mayor rebase a la que tiene pendiente menor.

Dadas las pendientes fijas $m_1, m_2 \in \mathbb{R}$ con $m_2 > m_1$ y el punto de intersección (x_0, y_0) , las rectas correspondientes son $l_1, l_2 : \mathbb{R} \rightarrow \mathbb{R}$ definidas por

$$l_1(x) = m_1(x - x_0) + y_0$$

y

$$l_2(x) = m_2(x - x_0) + y_0.$$

²El de que dos rectas tengan pendiente diferente implica que estas tienen una intersección no vacía que consiste en un único punto (Teorema 2.8)

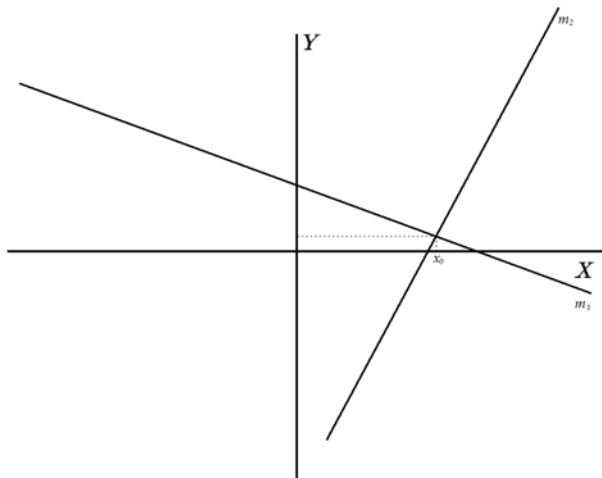


Figura 8: Rectas intersecadas.

Evaluando su diferencia $l_2 - l_1 : \mathbb{R} \rightarrow \mathbb{R}$:

$$[m_1(x - x_0) + y_0] - [m_2(x - x_0) + y_0] = (m_2 - m_1)(x - x_0)$$

Siendo que $m_2 > m_1$, se tiene que $m_2 - m_1 > 0$ y así, el signo del producto al que se redujo $l_2 - l_1$ está completamente determinado por el signo de $x - x_0$:

- $x < x_0 \Rightarrow x - x_0 < 0 \Rightarrow l_2(x) < l_1(x)$
- $x > x_0 \Rightarrow x - x_0 > 0 \Rightarrow l_2(x) > l_1(x)$

Es decir, la recta que se esperó que rebasara toma valores menores que la rebasada a la izquierda de su intersección, y mayores a su derecha (Vea la figura 8).

3.1 Rebase de funciones

Con la motivación anterior se define el rebase de funciones:

Definición 3.1. Sean $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$, $g : B \subseteq \mathbb{R} \rightarrow \mathbb{R}$, dos funciones y $x_0 \in \mathbb{R}$. Se dice que f rebasa a g en x_0 (g es rebasada por f en x_0) si y solo si existen números reales, a y b , con $a < x_0 < b$ tales que $]a, b[\setminus \{x_0\} \subseteq A \cap B$ y:

-) $\forall x \in (a, x_0) : g(x) > f(x)$
-) $\forall x \in (x_0, b) : f(x) > g(x)$.

Ahora hay que realizar unas observaciones urgentes:

- i)* Así como se emplea la palabra rebasar, cualquier sinónimo sirve para referirse a la misma situación, como adelantar o superar. Esto se aplica a la mayoría de los términos empleados a lo largo del texto.
- ii)* La condición de que el conjunto *no vacío* $]a, b[\setminus \{x_0\}$ esté contenido en ambos dominios garantiza que las funciones están definidas en tal conjunto.
- iii)* Los conjuntos de la forma $]a, b[\setminus \{x_0\}$ se llaman **vecindades agujeradas de x_0** y cuando están contenidas en el dominio de alguna función, se dice que esta está **definida para valores cercanos a x_0** .
- iv)* La definición de rebase puede reescribirse entonces de esta forma:

Definición 3.1 Sean $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$, $g : B \subseteq \mathbb{R} \rightarrow \mathbb{R}$, dos funciones y $x_0 \in \mathbb{R}$. Se dice que f rebasa a g en x_0 si ambas están definidas para valores cercanos a x_0 y existe una vecindad agujerada $]a, b[\setminus \{x_0\}$ tal que:

- \cdot) En $]a, x_0[$, f está por abajo de g
- $\cdot\cdot$) En $]x_0, b[$, f está por arriba de g .

Ejemplo 3.2. Considere una carrera de caballos, en la que dos de ellos van a la cabeza. Se trata de un alazán que va por delante de un albino por aproximadamente dos cabezas. De pronto, casi al llegar a la meta, el sofocado alazán experimenta *dolor de caballo* por tragar demasiado aire y reduce un poco su velocidad. Esta situación la aprovecha el jockey del albino, que le exige a su caballo y por casi nada logra ganar la carrera.

Si se construyen las funciones que asocian la posición de los caballos a cualquier instante de tiempo durante la carrera, la función que representa la posición del albino rebasa a la del alazán en el instante en que están cabeza con cabeza. Caballo que alcanza... rebasa y gana.

A diferencia del ejemplo de los caballos, en el que estos tuvieron que estar cabeza con cabeza para que uno pudiera adelantar al otro, según la definición no es necesario que haya intersecciones entre las funciones que se comparan, puesto que el punto en donde ocurre el rebase no tiene siquiera que estar en su dominio.

En cualquier caso, las proposiciones acerca del rebase están en términos de funciones que están por arriba o abajo de otras, que a su vez se enuncian como desigualdades de números reales. Resultados previsibles son los siguientes:

Proposición 3.3. Sean f, g, h , y k funciones reales de variable real tales que un intervalo no vacío $]a, b[$ está contenido en su dominio. Si f está por encima de g y h está por encima de k en $]a, b[$

1. $f + h$ está por encima de $g + k$ en $]a, b[$.
2. $-f$ está por abajo de $-g$ en $]a, b[$.
3. Si c es un número real positivo cf está por arriba de cg en $]a, b[$.
4. Si $]c, d[$ es un subintervalo no vacío de $]a, b[$ en que f y g están por arriba del eje horizontal (para todo $x \in]c, d[, f(x) > 0, g(x) > 0$) $\frac{1}{f}$ está por abajo de $\frac{1}{g}$ en $]c, d[$.
5. Si f está por arriba de g y g está por arriba de h en $]a, b[$, entonces f está por arriba de h en $]a, b[$.
6. Si $]c, d[$ es un subintervalo no vacío de $]a, b[$ en que f, g, h y k están por arriba del eje horizontal, fh está por arriba de gk en $]c, d[$.

Demostración. Todo es prácticamente inmediato de las definiciones de *estar por arriba* y *por abajo*. El único punto que tal vez requiera mayor atención es el cuarto: Como f está por encima de g en $]a, b[, \forall x \in]a, b[: f(x) > g(x)$. Por otro lado, según las hipótesis $\forall x \in]c, d[: f(x) > 0$ y $g(x) > 0$. Al ser $]c, d[\subseteq]a, b[, \forall x \in]c, d[: f(x) > 0, g(x) > 0$ y $f(x) > g(x)$. De ahí que $\forall x \in]c, d[: \frac{1}{f(x)} < \frac{1}{g(x)}$. \square

En la demostración se usó el hecho de que dados dos intervalos en los que se cumplen diferentes propiedades, si la intersección es no vacía, en ella se cumplen todas ellas. Con mayor precisión:

Sean a, b, c, d y x_0 números reales que satisfacen $a < x_0 < b, c < x_0 < d$; para las que son verdaderas ciertas proposiciones $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}$ y \mathfrak{D} , sobre x :

$$\begin{aligned} \forall x \in]a, x_0[: \mathfrak{A}(x) \quad \forall x \in]x_0, b[: \mathfrak{B}(x) \\ \forall x \in]c, x_0[: \mathfrak{C}(x) \quad \forall x \in]x_0, d[: \mathfrak{D}(x). \end{aligned}$$

Al tomarse $\alpha = \max(a, c)$, así como $\beta = \min(b, d)$, se advierte que $]a, x_0[\subseteq]\alpha, x_0[,]\alpha, x_0[\subseteq]c, x_0[,]x_0, \beta[\subseteq]x_0, b[$ y $]x_0, \beta[\subseteq]x_0, d[$.

Por lo anterior, son válidas las proposiciones:

$$\forall x \in]\alpha, x_0[: (\mathfrak{A}(x) \text{ y } \mathfrak{C}(x)) \quad \text{y} \quad \forall x \in]x_0, \beta[: (\mathfrak{B}(x) \text{ y } \mathfrak{D}(x))$$

Por tal motivo, dadas dos vecindades agujeradas en torno a un mismo punto, siempre es posible elegir una vecindad especial y referirse a ella como **una vecindad que atrapa las propiedades de las otras dos**, y respectivamente a sus **parte izquierda** y **parte derecha**.

Dicho esto, una consecuencia de la proposición anterior es el siguiente teorema

Teorema 3.4. Sean $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$, $g : B \subseteq \mathbb{R} \rightarrow \mathbb{R}$, $h : C \subseteq \mathbb{R} \rightarrow \mathbb{R}$, $k : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ funciones con la única condición de que existe x_0 , un número real en el que f rebasa a g y h rebasa a k . Entonces:

1. $f + h$ rebasa a $g + k$ en x_0
2. $(-g)$ rebasa a $(-f)$ en x_0
3. Si c es un número real positivo, entonces cf rebasa a cg en x_0
4. Si existen números $a, b \in A \cap B$ tales que f y g están sobre el eje horizontal en $]a, b[\setminus\{x_0\}$, entonces $\frac{1}{g}$ rebasa a $\frac{1}{f}$ en x_0 .
5. Si existen números $a, b \in A \cap B$ tales que f, g, h y k están sobre el eje horizontal en $]a, b[\setminus\{x_0\}$, entonces fh rebasa a gk en x_0 .

Demostración. (1) Sean f, g, h y k funciones tales que f rebasa a g en x_0 y h rebasa a k en x_0 . Existen entonces $a, b, c, d \in \mathbb{R}$ con $a < x_0 < b$ y $c < x_0 < d$ tales que:

En $]a, x_0[$, f está por abajo de g En $]c, x_0[$, h está por abajo de k
 En $]x_0, b[$, f está por arriba de g En $]x_0, d[$, h está por arriba de k .

Sea $]a, b[\setminus\{x_0\}$ una vecindad que atrapa las propiedades de las otras dos. Entonces

- En $]a, x_0[$, f está por abajo de g y h está por abajo de k
- En $]x_0, b[$, f está por arriba de g y h está por arriba de k .

Por el punto 1 de la Proposición 3.3

- En $]a, x_0[$, $f + h$ está por abajo de $g + k$
- En $]x_0, b[$, $f + h$ está por arriba de $g + k$.

(4) Sean f y g tales que f rebasa a g en x_0 y $c, d \in \mathbb{R}$ con $c < x_0 < d$ que satisfacen que en $]c, d[\setminus\{x_0\}$ f y g están sobre el eje horizontal. Dado el rebase, existen $a, b \in \mathbb{R}$ con $a < x_0 < b$ tales que

En $]a, x_0[$, f está por abajo de g En $]x_0, b[$, f está por arriba de g .

Sea $]a, b[\setminus\{x_0\}$ una vecindad que atrapa las propiedades de las otras dos. Entonces

- En $]a, x_0[$, f está por abajo de g y ambas están por arriba del eje horizontal
- En $]x_0, b[$, f está por arriba de g y ambas están por arriba del eje horizontal.

Por el punto 4 de la Proposición 3.3

- En $] \alpha, x_0[$, $\frac{1}{g}$ está por abajo de $\frac{1}{f}$
- En $] x_0, \beta[$, $\frac{1}{g}$ está por arriba de $\frac{1}{f}$. □

La escritura de las demostraciones de los puntos restantes de la proposición es similar y se aconseja para habituarse al tratamiento de la definición.

Otro resultado de gran importancia es el siguiente:

Teorema 3.5. *Si una función, f , rebasa a otra función, g , en x_0 ; y esta a su vez rebasa a una tercera, h , en el mismo punto, entonces f rebasa a h en x_0 .*

Demostración. Sean f , g , y h funciones tales que f rebasa a g en x_0 y h rebasa a g en x_0 . Existen entonces $a, b, c, d \in \mathbb{R}$ con $a < x_0 < b$ y $c < x_0 < d$ tales que:

- En $] a, x_0[$, f está por abajo de g En $] c, x_0[$, g está por abajo de h
- En $] x_0, b[$, f está por arriba de g En $] x_0, d[$, g está por arriba de h .

Tomando una vecindad agujerada $] e, f[\{x_0\}$ que atrape las propiedades de las otras dos

- En $] e, x_0[$ f está por abajo de g y g está por abajo de h
- En $] x_0, f[$ f está por arriba de g y g está por arriba de h .

Por el punto 5 de la proposición 3.3:

- En $] e, x_0[$ f está por abajo de h
- En $] x_0, f[$ f está por arriba de h . □

A lo largo del párrafo anterior, se analizaron cuestiones relativas al rebase de funciones en puntos. En particular, se probó que existen funciones tales que si sus grafos se intersecan, alguna de ellas rebasa a la otra (caso de las rectas con diferentes pendientes).

Naturalmente, no ocurre que siempre que las gráficas de dos funciones se intersecten, alguna de ellas rebasa a la otra en el punto de intersección: esto se exhibió en la primera parte considerando la circunferencia unitaria y su tangente horizontal en $(0, 1)$.

Un asunto más interesante en que se husmeará aquí es: dada alguna función real definida en un intervalo, para cualquier punto en su dominio, ¿bajo qué condiciones es posible encontrar una recta que la rebasa en ese punto?

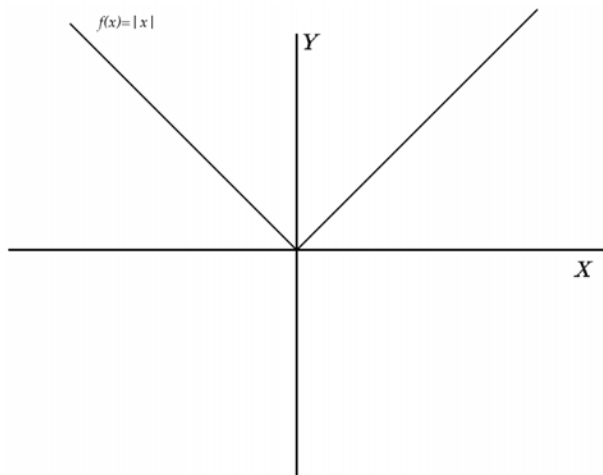


Figura 9: Esbozo de la función valor absoluto.

Ejemplo 3.6.

(1) La función valor absoluto.

Considérese la función $f : \mathbb{R} \rightarrow \mathbb{R}$ definida por:

$$f(x) = \begin{cases} x & \text{si } x \geq 0 \\ -x & \text{si } x < 0 \end{cases}$$

El valor de f en x se denota usualmente por $|x|$.

En cada punto del conjunto $\mathbb{R} \setminus \{0\}$, el análisis de rebase por rectas puede transcribirse, palabra por palabra, del realizado para las rectas mismas, puesto que para cada x que satisface $x < 0$, f coincide en una vecindad agujerada con la función que determina la recta que pasa a través del origen y que tiene pendiente igual a -1 . De forma semejante, para los valores x' que satisfacen $x' > 0$, se observa la coincidencia entre esta función y la función identidad en alguna vecindad. Vea la Figura 9.

Con base en este comentario, considérense las funciones

$$\begin{array}{lll} L : \mathbb{R} \rightarrow \mathbb{R} & L_1 : \mathbb{R} \rightarrow \mathbb{R} & L_2 : \mathbb{R} \rightarrow \mathbb{R} \\ x \mapsto 2x & x \mapsto x & x \mapsto -x \end{array}$$

Las tres funciones contienen al origen $(0, 0)$, así que, empleando el teorema de rebase de rectas: L rebasa a L_1 y a L_2 en 0. Así que existen $a, b \in \mathbb{R}$ con $a < 0 < b$, tales que

$$\text{En }]a, 0[, L \text{ está abajo de } L_1 \quad \text{En }]0, b[, L \text{ está arriba de } L_1.$$

Así mismo, existen $c, d \in \mathbb{R}$ con $c < 0 < d$, tales que

En $]c, 0[$, L está abajo de L_2 En $]0, d[$ L , está arriba de L_2 .

Sea $] \alpha, \beta [\setminus \{x_0\}$ una vecindad agujerada que atrapa las propiedades de las anteriores:

En $] \alpha, 0[$, L está abajo de L_2 En $]0, \beta[$, L está arriba de L_1 .

Por otro lado, cuando $x \in] \alpha, 0[$, $L_2(x) = |x|$; y si $x \in]0, \beta[$, $L_1(x) = |x|$. Así se tiene:

En $] \alpha, 0[$, L está abajo de f En $]0, \beta[$, L está arriba de f .

De este modo, la función L rebasa a la función valor absoluto en 0. El uso de las desigualdades omitidas:

En $] \alpha, 0[$, L está abajo de L_1 En $]0, \beta[$, L está arriba de L_2 .

termina probando de paso que la recta definida por $T(x) = -2x$ es rebasada por la función valor absoluto en 0.

(2) La función de Dirichlet.

Durante los estudios de preparatoria, los estudiantes suelen enfrentarse a funciones usuales de gran interés como las funciones lineales y los polinomios. Saliendo un poco de esa *zona de confort*...

La función de Dirichlet, $D : \mathbb{R} \rightarrow \mathbb{R}$ está definida por la relación:

$$D(x) = \begin{cases} 0 & \text{si } x \in \mathbb{Q} \\ 1 & \text{si } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

Nuevamente, por practicidad se analizará a la función únicamente en el punto $x_0 = 0$. Note que no es posible dibujar con precisión la gráfica de esta función, dada la densidad de los conjuntos \mathbb{Q} y $\mathbb{R} \setminus \mathbb{Q}$.

Para que la función $L_{m,k} : \mathbb{R} \rightarrow \mathbb{R}$ determinada por las constantes m y k - definida por $L_{m,k}(x) = mx + k$ - rebasa a la función de Dirichlet en 0, se requiere que

$$\begin{aligned} mx + k &< 0 && \text{siempre que } x < 0 \\ mx + k &> 1 && \text{siempre que } x > 0 \end{aligned}$$

Si la pendiente m fuese cero, ocurriría para la constante k

$$\begin{aligned} k &< 0 && \text{siempre que } x < 0 \\ k &> 1 && \text{siempre que } x > 0 \end{aligned}$$

Así que $m \neq 0$, y para cada valor, $y \in \mathbb{R}$, existe un punto $x_0 \in \mathbb{R}$ para el que se cumple: $L_{m,k}(x_0) = y$.

Sea y un número arbitrario en el intervalo $]0, 1[$. Entonces en algún punto x_1 , la recta alcanza el valor de y ; pero este número no puede ser positivo, dado que para todo número positivo, el valor de la función recta debe ser mayor que 1, ni tampoco puede ser negativo porque el valor de dicha función para este tipo de cantidades es menor que cero. La única opción que queda es que en cero se alcance el valor de todos los puntos entre 0 y 1.

Las contradicciones anteriores se obtuvieron a partir del hecho de asumir la existencia de una recta $L_{m,k}$ que rebasa a la función de Dirichlet en el punto 0. ¡No existe ninguna recta que rebese a esta función en 0!

Un razonamiento similar permite probar que tampoco existen rectas que sean rebasadas por la función de Dirichlet en 0; y de hecho, que no existe recta alguna que rebese o sea rebasada por esta función en cualquier punto sobre \mathbb{R} .

Vale la pena hacer un comentario: aunque un dibujo o esquema jamás sustituye a una demostración formal, el empleo de algunos provee de una ayuda inigualable para convencerse de la veracidad de los resultados y formular argumentos que posteriormente constituirán la demostración.

(3) La función *máximo entero* es una función real definida para todos los números reales, x de la manera siguiente:

$$f(x) = m \quad \text{donde} \quad m = \max\{n \in \mathbb{Z} | n \leq x\}$$

El valor de la función en x_0 se denota por $[x_0]$.

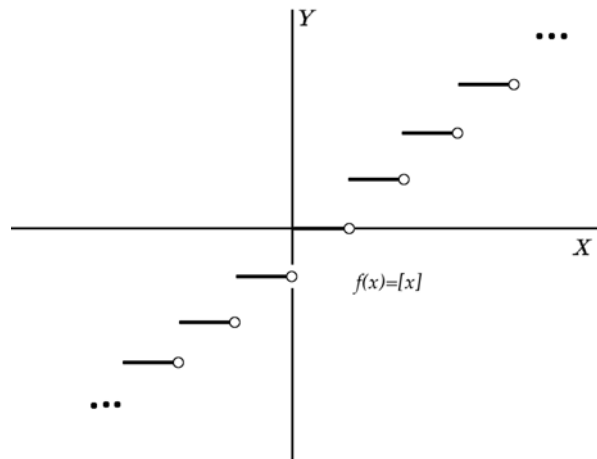


Figura 10: Esbozo de la función máximo entero.

Note que para números no enteros, la función parte entera coincide en alguna vecindad con una recta de pendiente 0. Vea la figura 10. Para cada número entero

z , existe una función lineal que es rebasada por la función parte entera mientras que no existe ninguna que la rebase ahí. [*Sugerencia*: Una recta horizontal debería bastar; pero si no lo ve tan claro, pueden seguirse ideas similares a las empleadas para el análisis de la *función de Dirichlet*].

(4) El hecho de que dadas dos funciones, f y g tales que f rebasa a g en algún punto x_0 implique que $-g$ rebasa $-f$ en x_0 tiene una interpretación geométrica importante. Recuerde el lector que para obtener los puntos simétricos de otros dados respecto al eje X basta tomar los opuestos en las ordenadas. En lo sucesivo remítase a las Figuras 11, 12, 13, 14

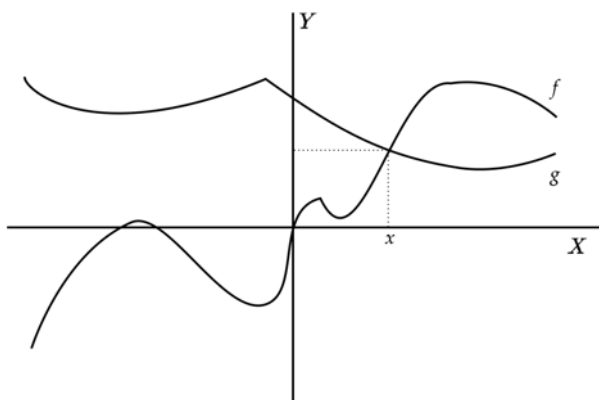


Figura 11: f rebasa a g en x .

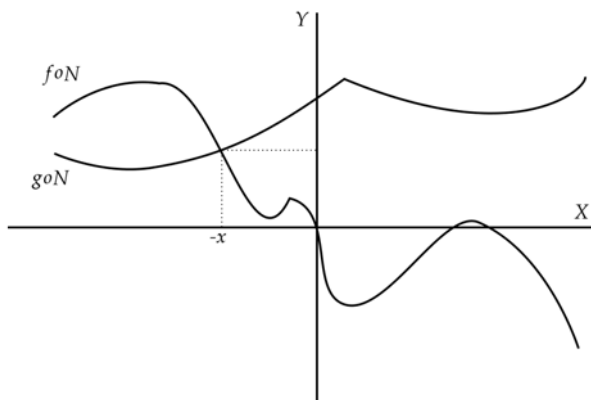


Figura 12: Reflexión sobre eje Y .

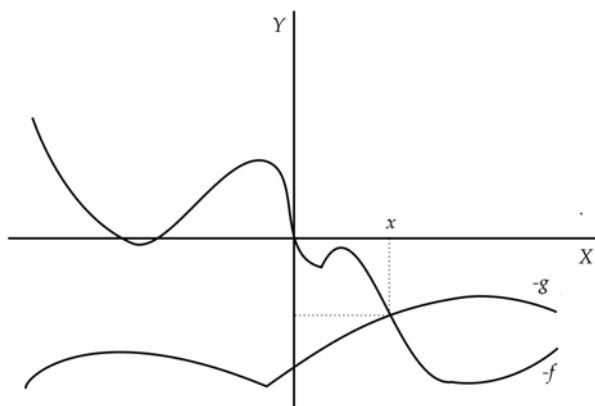


Figura 13: Reflexión sobre eje X .

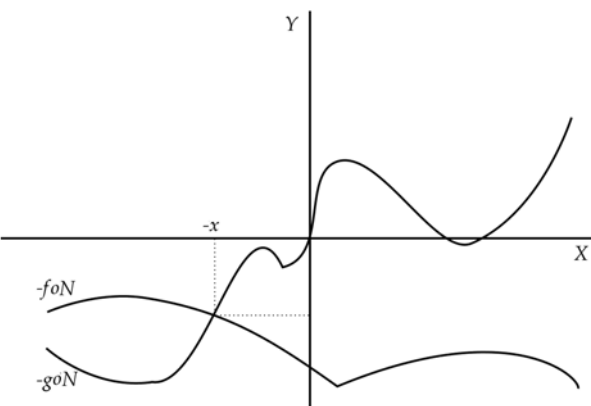


Figura 14: Reflexión respecto al origen.

- i) Considere ahora el reflejo de las gráficas de las funciones respecto al eje Y . Para ello sirve la función $N : \mathbb{R} \rightarrow \mathbb{R}$ definida para todos los números reales x por $N(x) = -x$. Si una función, f , rebasa a otra, g , en un punto x_0 , entonces $g \circ N$ rebasa a $f \circ N$ en $-x_0$.

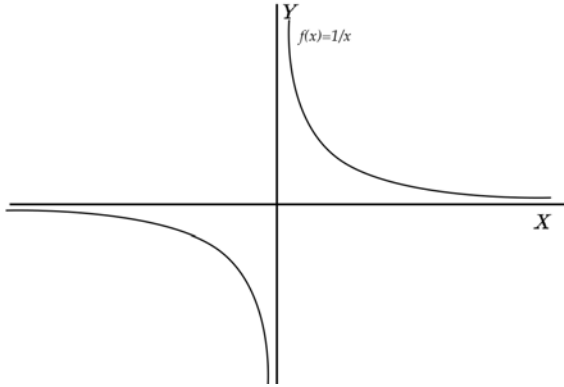


Figura 15: Función definida por $f(x) = \frac{1}{x}$.

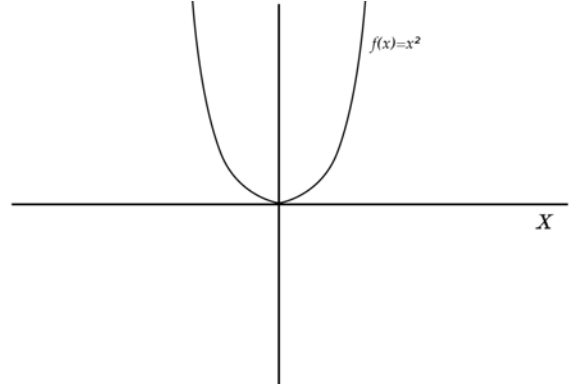


Figura 16: Función definida por $f(x) = x^2$.

- ii) Considere ahora el reflejo de las gráficas respecto al origen. Si una función, f , rebasa a otra, g , en un punto x_0 , entonces $-f \circ N$ rebasa a $-g \circ N$ en $-x_0$.

(5) Sea $L : \mathbb{R} \rightarrow \mathbb{R}$ una función determinada por una recta con pendiente distinta de 0. La función $\frac{1}{L}$ tiene como dominio máximo a todos los números reales salvo el punto en que la función L se anula. Geométricamente representa una hipérbola equilátera (sus asíntotas son perpendiculares entre sí).

Para ver lo anterior, recuerde que si la forma de L está dada por $L(x) = mx + b$ (con $m \neq 0$), se tiene que para cualquier punto distinto de $\frac{-b}{m}$, $\frac{1}{L}(x) = \frac{1}{mx+b}$ y $\frac{1}{m} = ((\frac{1}{L})(x) - 0)(x - (\frac{-b}{m}))$. Se tiene una hipérbola equilátera con centro en el par $(0, \frac{-b}{m})$ y asíntotas paralelas a los ejes coordenados.

Para cualquier punto, x , en el dominio de la función $\frac{1}{L}$, existen rectas que la rebasan y son rebasadas por ella en x .

(Se considerará la hipérbola definida por la ecuación $y = \frac{1}{x}$ y un punto x_0 positivo. Vea la figura 15.

Considere un punto x' mayor que cero; pero menor que x_0 , asimismo los pares $(x_0, \frac{1}{x_0})$ y (x', M) con $M > \frac{1}{x'}$.

Tras generar la recta que los contiene se ve que que ella es rebasada por la función en x_0 . Posteriormente puede hacerse algo similar para un valor $m < \frac{1}{x'}$.

Para probar el resultado para puntos positivos, basta para generalizar el resultado a toda la hipérbola equilátera. Por último note que la demostración para este caso particular es suficiente para los cocientes de todas las funciones lineales no constantes si se recuerda que reemplazar x por $x - a$ en una ecuación es trasladar la gráfica hasta reestablecer el origen de coordenadas en $(a, 0)$.

(6) Sean L_1 y L_2 dos funciones determinadas por dos rectas en el plano. Cuando una de ellas está determinada por una recta con pendiente 0, el producto $L_1 \cdot L_2$ es

una recta, y por tanto existen rectas que rebasan y son rebasadas por este producto en cualquier punto.

Cuando ambas están determinadas por rectas de pendiente distinta de cero, el resultado es una *función cuadrática*. Para toda función cuadrática en cualquier punto, existen rectas que rebasan y son rebasadas por ella. Para esto recuerde que las funciones cuadráticas determinan parábolas en el plano.

(Bastará demostrar para $f : \mathbb{R} \rightarrow \mathbb{R}$ definida por la regla $f(x) = x^2$ de la manera siguiente: tomando la función $h : \mathbb{R} \rightarrow \mathbb{R}$ definida por $h(x) = x^2 - (x_0^2 + m(x - x_0))$; esta función compara a f y las rectas a través de los pares $(x_0, f(x_0))$ para cada x_0 de tal modo que si la función es mayor que la recta, h es positiva y viceversa. h puede reescribirse como $h(x) = (x - x_0)(x + x_0 - m)$. El segundo factor toma el valor de $2x_0 - m$ en x_0 . Aprovechando el cambio de signo cuando se toman $x < x_0$ y $x > x_0$ en el primer factor, se puede saber cuándo f es mayor o menor que la recta cuando m toma valores distintos a $2x_0$.

Esto es suficiente para cualquier función cuadrática; usando a discreción los teoremas de rebase demostrados en esta sección).

4 Continuidad de funciones

En esta sección, se propone una formalización de una intuición geométrica importante: la *continuidad*.

Genéricamente se identifica a un trazo como *continuo* si este puede realizarse *sin levantar el lápiz del papel*, o su equivalente dependiendo de con qué se escriba. . . Dicho de otra forma, es tentador comprender a la continuidad de las funciones como una propiedad presente en intervalos.

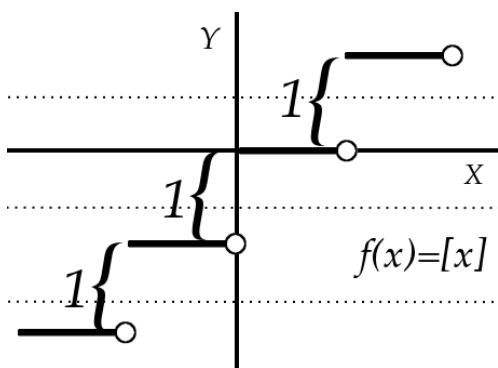


Figura 17: Brechas en $f(x) = [x]$.

Considere por ejemplo a la *función máximo entero*. Para esta función existen rectas que son rebasadas por ella en cada punto entero (marcadas en la Figura 17 con líneas punteadas). Aún cuando no pueda por ahora darse una prueba; parece que cada que se tome un intervalo que contenga a algún entero, el trazo de esta debe interrumpirse en él, pues el valor de la función, *poco antes* difiere en 1 del valor de la función *poco después*.

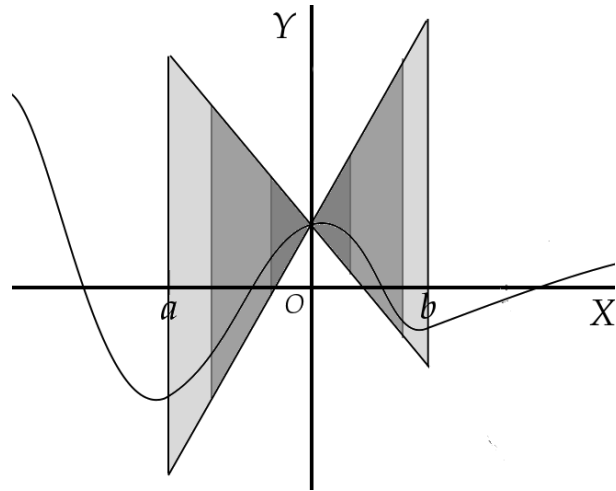
El valor de 1 es completamente irrelevante, puesto que aunque la diferencia fuera cualquier número real, una brecha continuaría existiendo. Es decir, los valores de la función *cerca* de cada número entero no se parecen entre sí comparando puntos a la izquierda

y a la derecha del mismo.

En el caso de la función de Dirichlet, trazar su gráfica es una tarea no menos que imposible, así que en una clasificación de las funciones con respecto a su continuidad, indudablemente debe pertenecer a las funciones no continuas.

Ahora considere el trazo siguiente sobre el plano cartesiano. Este define una función intuitivamente *continua* sobre el intervalo en que esté trazado. Considere el punto O . Existen una recta que es rebasada y otra que rebasa a esta función en O . Si se centra la atención en el intervalo $]a, b[$, entre las rectas verticales a través de a y b y las mencionadas, se genera una *región de mariposa* dentro de la cual está insertada la función para valores cercanos a O .

Mediante sucesivas reducciones de la longitud de $]a, b[$ alrededor de O , la región ocupa una porción del plano más pequeña que “converge” al valor de la función. Es decir, los puntos en la región de mariposa cada vez se parecen más al de la función en O y por ello, no existen brechas de la función en el punto puesto que de haberla, habría diferencia entre los valores de la función cerca de O y el de O ; pero esto no es posible dado que la función está dentro de la mariposa.



Así pues, bajo condiciones especiales, la existencia de rectas que rebasan y son rebasadas por una función en un punto garantiza comportamientos ciertamente en intervalos; pero siempre referidos al punto en que se presentan los rebases.

Con esas bases:

Definición 4.1. Sean $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ una función y x_0 un elemento de su dominio. Se dirá que f es **continua** en x_0 si y solo si existen tanto una recta que rebasa a f en x_0 como otra que es rebasada por f en x_0 y ambas contienen al par $(x_0, f(x_0))$.

Equivalentemente, en términos de regiones de mariposa:

Definición 4.1. Sean $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ una función y x_0 un elemento de su dominio. Se dirá que f es continua en x_0 si y sólo si existe una región de mariposa determinada por rectas l y L que contienen a $(x_0, f(x_0))$ y puntos $a < x_0 < b$ tales que $\forall x \in]a, b[\setminus \{x_0\} : f(x) \in M_a^b(l, L)$ (ver definición 2.10 de mariposa).

Cuando una función no es continua en un punto se llama discontinua ahí. Por otro lado, si A es un conjunto de números reales en que una función es continua, entonces se dice que es continua en A .

Ahora que se ha puesto de relieve la importancia de las funciones lineales para el estudio de las funciones, conviene dotarse de un pequeño arsenal de teoremas acerca de las funciones lineales.

Teorema 4.2. Sean f y g dos funciones lineales tales que $f(x) = m(x - x_0) + y_0$ y $g(x) = n(x - x_0) + y_0$ (notar que si $f(x_0) = g(x_0)$, las funciones lineales son de esta forma). Entonces, f rebasa a g en x_0 si y solo si $m > n$.

Demostración. La prueba de que pendientes diferentes implican el rebase está en el análisis de rectas de la Sección 3. Para el recíproco, note que la negación de $m > n$ es $m \leq n$. Si la igualdad se da, entonces f y g son idénticas; mientras que si $m < n$, entonces g rebasa a f en x_0 , por lo que f no rebasa a g en x_0 . \square

Corolario 4.3. Sean f y g dos funciones lineales tales que $f(x) = m(x - x_0) + y_0$ y $g(x) = n(x - x_0) + y_0$. Existen rectas h y \bar{h} definidas por $h(x) = k(x - x_0) + y_0$ y $\bar{h}(x) = -k(x - x_0) + y_0$ tales que h rebasa a f y a g en x_0 , mientras que estas últimas rebasan a \bar{h} .

Demostración. Sean f y g como en las hipótesis del corolario y sea $k > \max(|m|, |n|)$. Luego, $k > m$, $k > n$, $m > -k$ y $n > -k$. Del Teorema 4.2, definiendo $h(x) = k(x - x_0) + y_0$ y $\bar{h}(x) = -k(x - x_0) + y_0$, se sigue el resultado de inmediato. \square

Corolario 4.4. Sea f una función lineal. Para cualquier número x_0 existen rectas g y h tales que g rebasa a f y f rebasa a h en x_0 . Es decir, f es continua en \mathbb{R} .

Gracias al Corolario 4.4 y a los Ejemplos 3.6, (5) y (6), se tiene el siguiente resultado.

Teorema 4.5. Sean f y g dos funciones lineales, entonces

1. fg es continua en \mathbb{R}
2. $\frac{1}{f}$ es continua en $\mathbb{R} \setminus \{x \in \mathbb{R} \mid x \text{ es cero de } f\}$

Para garantizar la continuidad de una función f en un punto x_0 , se solicita la existencia de una recta que rebasa y una que sea rebasada por la función en el punto, con la particularidad de que ambas contengan a $(x_0, f(x_0))$. Una vez que se sepa que la función es continua, es natural esperar, que cualquier otra recta que rebasa o sea rebasada por f en x_0 también contenga al par $(x_0, f(x_0))$.

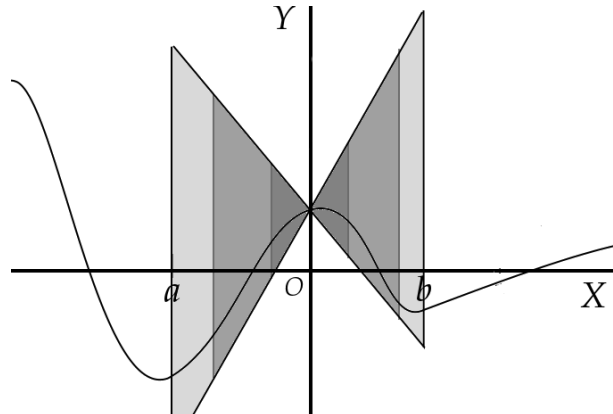
Teorema 4.6. Sea $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ una función continua en x_0 . Cualesquiera rectas que rebasen o sean rebasadas por f en x_0 comparten al par $(x_0, f(x_0))$.

Demostración. Sea L una función lineal que rebasa a f en x_0 . Como f es continua ahí, existe otra función lineal L' tal que f rebasa a L' en x_0 .

Por el Teorema 3.2, el rebase es transitivo, así que L rebasa a L' en x_0 . Sean $a, b \in \mathbb{R}$ tales que $a < x_0 < b$ y $h = L - L'$, se tiene $h(a) = L(a) - L'(a) < 0$ y $h(b) = L(b) - L'(b) > 0$, así que por la propiedad de los valores intermedios para funciones lineales (ver Teorema 2.4), existe un único $z \in]a, b[$ tal que $h(z) = 0$, $z \notin]a, x_0[$ porque en $]a, x_0[$ L está abajo de L' . Tampoco $z \in]x_0, b[$ porque ahí es positiva, así que $z = x_0$ y $0 = h(x_0) = L(x_0) - L'(x_0)$, de modo que $L(x_0) = L'(x_0) = f(x_0)$. En el caso de rectas que son rebasadas por f se procede de manera análoga. \square

Ya se ha hecho énfasis en el hecho de que si una función es continua en un punto, la diferencia entre los valores cerca de él y en él, puede garantizarse menor entre más estrecha sea la vecindad. Para eso sirvió colocar a la función dentro de una región de mariposa.

De manera más precisa: si el valor de la función en el punto de continuidad es mayor o menor que otra cantidad dada, es posible encontrar una vecindad que herede dicha propiedad, puesto que la atención puede centrarse en la porción de la mariposa que está por arriba o por abajo de la recta horizontal que determina el número dado.



Proposición 4.7. Sean $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ una función continua en x_0 y m , un número real, tal que $f(x_0) > m$ ($f(x_0) < m$). Existe una vecindad de x_0 , $] \alpha, \beta [$ en que se preserva dicha desigualdad; es decir:

$$\forall x \in] \alpha, \beta [: f(x) > m (f(x) < m).$$

Demostración. Sean f una función continua en x_0 y m un número real menor que $f(x_0)$. Por los Corolarios 4.4 y 4.3, existen dos funciones lineales, una h que rebasa y otra, \bar{h} que es rebasada por f en x_0 ; siendo la pendiente de h positiva, y la de \bar{h} tal que sumada con la de h se obtiene 0. Así h es creciente y \bar{h} , decreciente.

Como $m < f(x_0)$, existen $x_1 < x_0$ tal que $h(x_1) = m$ y $x_2 > x_0$ tal que $\bar{h}(x_2) = m$, Por esta misma monotonía, $\forall x \in]x_1, x_0[: h(x) > h(x_1) = m$ y $\forall x \in]x_0, x_2[: \bar{h}(x) > \bar{h}(x_2) = m$.

Si $]c, d[\setminus \{x_0\}$ es una vecindad que garantiza el rebase de ambas rectas con f , tómesese una vecindad que atrape estas propiedades. Intersectándola con $]x_1, x_2[\setminus \{x_0\}$ y nombrando a sus extremos α y β , es cierto que:

i) $\forall x \in]\alpha, x_0[: h(x) > m$ y $f(x) > h(x)$

ii) $\forall x \in]x_0, \beta[: \bar{h}(x) > m$ y $f(x) > \bar{h}(x)$.

Considerando que $f(x_0) > m$:

$$\forall x \in]\alpha, \beta[: f(x) > m.$$

El caso en que $m > f(x_0)$ es completamente análogo. \square

Como una consecuencia de este resultado, se tiene el siguiente corolario.

Corolario 4.8. *Sea $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ una función continua en x_0 , entonces f es acotada en alguna vecindad de x_0 .*

4.1 Operaciones con funciones continuas

La siguiente proposición tiene la finalidad de mostrar que el realizar operaciones con las funciones continuas en algún punto ofrece como resultado funciones continuas.

Proposición 4.9. Sean $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ y $g : B \subseteq \mathbb{R} \rightarrow \mathbb{R}$ dos funciones continuas en el punto x_0 . Entonces:

1. $-f$ es continua en el punto x_0 . Si c es un número real, cf es continua en x_0 .
2. $f + g$ es continua en x_0 .
3. Si g es distinta de 0 en x_0 , $\frac{1}{g}$ es continua en x_0 .
4. fg es continua en x_0 .

Demostración. La demostración del primer inciso es sencilla y la omitiremos; pero emplearemos ese resultado para las pruebas de los incisos siguientes.

2.

Considérense dos funciones, f y g , y un punto x_0 para los que son válidas las hipótesis de la afirmación. Como f es continua en x_0 , existen L_1 y L_2 , funciones lineales, tales que L_1 rebasa a f y f rebasa a L_2 en x_0 . De la misma manera, existen L_3 y L_4 , rectas, tales que L_3 rebasa a g y g rebasa a L_4 en x_0 . El inciso 1 del teorema 3.4 garantiza que $(L_1 + L_3)$ rebasa a $f + g$ y $f + g$ rebasa a $(L_2 + L_4)$ en x_0 . Dado que f y g son ambas continuas, se tiene que $L_1(x_0) = f(x_0) = L_2(x_0)$ y $L_3(x_0) = g(x_0) = L_4(x_0)$, por lo que $(L_1 + L_3)(x_0) = (f + g)(x_0) = (L_2 + L_4)(x_0)$. Para finalizar basta notar que la suma de funciones lineales es una función lineal, de modo que $f + g$ es continua en x_0 .

3.

Sea g una función continua en el punto x_0 tal que $g(x_0) \neq 0$. Al ser $g(x_0) \neq 0$ y continua en x_0 , existe una vecindad agüjerada de x_0 en el que el signo de g es el mismo que el de $g(x_0)$ por la Proposición 4.7.

Caso I. $g(x_0) > 0$.

Como g es continua en x_0 , existen una recta, h , que rebasa a g en x_0 y otra, \bar{h} , que es rebasada por g en x_0 , siendo h de pendiente positiva y \bar{h} de pendiente negativa (Corolario 4.3). Por la observación anterior, existe una vecindad en la que tanto L , g y L' son positivas (¡No olvide que también las funciones lineales son continuas!), así que por el inciso 4 del teorema 3.4, $\frac{1}{L}$ rebasa a $\frac{1}{g}$ y $\frac{1}{g}$ rebasa a $\frac{1}{L'}$ en x_0 . El inciso 2 del Teorema 4.5 establece que $\frac{1}{L}$ y $\frac{1}{L'}$ son continuas, así que existen funciones lineales, L_1 y L_2 , tales que L_1 rebasa a $\frac{1}{L'}$ y $\frac{1}{L}$ rebasa a L_2 en x_0 . Por transitividad L_1 rebasa a $\frac{1}{g}$ y $\frac{1}{g}$ rebasa a L_2 en x_0 . Además, dado que $g(x_0) = L(x_0) = L'(x_0)$ y que los cocientes de las funciones lineales son continuos, $L_1(x_0) = L_2(x_0) = (\frac{1}{g})(x_0)$. Se concluye que $\frac{1}{g}$ es continua en x_0 .

Caso II. $g(x_0) < 0$.

Es suficiente tomar a la función $-g$. Se tendría que $\frac{1}{-g}$ es continua en x_0 y al aplicar el inciso 1 de esta proposición, $-\frac{1}{-g} = \frac{1}{g}$ es continua en x_0 .

4.

Caso I. $f(x_0) > 0$ y $g(x_0) > 0$.

Como f y g son continuas en x_0 , existen L_1, L_2, L_3 y L_4 , funciones lineales, tales que L_1 rebasa a f , f rebasa a L_2 , L_3 rebasa a g y g rebasa a L_4 en x_0 . Al ser ambas funciones, f y g , positivas en x_0 , existe una vecindad de x_0 en que se preserve esta característica para las dos. Así, por el inciso 5 del Teorema 3.4, $(L_1 \cdot L_3)$ rebasa a $f \cdot g$ y $f \cdot g$ rebasa a $(L_2 \cdot L_4)$ en x_0 . El apartado 1 del Teorema 4.5 indica que $(L_1 \cdot L_3)$ y $(L_2 \cdot L_4)$ son continuas en x_0 , por lo que existen rectas, L y L' tales que L rebasa a $(L_1 \cdot L_3)$ y $(L_2 \cdot L_4)$ rebasa a L' en x_0 . Por transitividad L rebasa a fg y fg rebasa a L' en x_0 . Además se cumple que $L'(x_0) = (L_1 \cdot L_3)(x_0) = fg(x_0)$ y $L(x_0) = (L_2 \cdot L_4)(x_0) = fg(x_0)$, por lo que fg es continua en x_0 .

Caso II. $f(x_0) < 0$ y $g(x_0) > 0$.

Tomando a la función $-f$ y aplicando el caso anterior a $-f$ y g , $-fg$ es continua en x_0 y por tanto, también fg .

Caso III. $f(x_0) < 0$ y $g(x_0) < 0$.

Ahora debe aplicarse el mismo resultado a las funciones $-f$ y $-g$, ambas continuas en x_0 para lograr la continuidad de $fg = (-f)(-g)$.

Caso IV. $f(x_0) = 0$ y $g(x_0) \neq 0$.

Defínase $\phi : \text{Dom}(f) \rightarrow \mathbb{R}$ por $\phi(x) = M$ con $M > 0$ arbitrario. Se sabe ya que $f + \phi$ es continua en x_0 (al ser f y ϕ continuas en x_0) y $(f + \phi)(x_0) = M$. Aplicando el caso I, $(f + \phi)g$ es continua en x_0 y así $fg + \phi g$ es continua en x_0 , y como ϕg es continua y $fg = (fg + \phi g) - \phi g$, fg es continua en x_0 .

Caso V. $f(x_0) = 0$ y $g(x_0) = 0$.

Hay que aplicar el mismo razonamiento que en el caso anterior para ambas funciones. \square

Mención especial merece la composición de funciones continuas.

Proposición 4.10. Sean f y g dos funciones tales que f es continua en x_0 y g es continua en $f(x_0)$. Entonces $g \circ f$ es continua en x_0 .

Demostración. Con base en el Corolario 4.3, como g es continua en $f(x_0)$, existen funciones lineales h, \bar{h} tales que h tiene pendiente positiva, \bar{h} tiene pendiente negativa, h rebasa a g y g rebasa a \bar{h} en $f(x_0)$. Por otro lado, como f es continua en x_0 , existen funciones lineales t, \bar{t} tales que t tiene pendiente positiva, \bar{t} tiene pendiente negativa, t rebasa a f y f rebasa a \bar{t} en x_0 . Así, existe una vecindad agujerada de $f(x_0)$, $]c, d[\setminus \{f(x_0)\}$ tal que

$$\forall x \in]c, f(x_0)[, h(x) < g(x) < \bar{h}(x) \quad \text{y} \quad \forall x \in]f(x_0), d[, h(x) > g(x) > \bar{h}(x).$$

Como f es continua en x_0 y $c < f(x_0) < d$, entonces, gracias a la Proposición 4.7, existe una vecindad de x_0 tal que para todo elemento x de ella, se cumple la desigualdad $c < f(x) < d$. Así mismo existe una vecindad agujerada de x_0 en la que se garantizan los rebases de f , t y \bar{t} . Sea $]a, b[\setminus \{x_0\}$ una vecindad que atrapa las propiedades de las otras dos. De este modo:

$$\begin{aligned} \forall x \in]a, x_0[, c < t(x) < f(x) < \bar{t}(x) < d \quad \text{y} \\ \forall x \in]x_0, b[, d > t(x) > f(x) > \bar{t}(x) > c. \end{aligned}$$

Ahora se desea encontrar funciones lineales que atrapen dentro de una mariposa a $g \circ f$ para alguna vecindad agujerada de x_0 . Las propuestas naturales son las composiciones $h \circ t$, $h \circ \bar{t}$, $\bar{h} \circ t$ y $\bar{h} \circ \bar{t}$. Note que $h \circ t(x_0) = h(t(x_0)) = h(f(x_0)) = g(f(x_0))$, y que es el mismo valor para $h \circ \bar{t}(x_0)$, $\bar{h} \circ t(x_0)$ y $\bar{h} \circ \bar{t}(x_0)$. Las pendientes de $h \circ t$ y $\bar{h} \circ \bar{t}$ son positivas, mientras que las de $h \circ \bar{t}$ y $\bar{h} \circ t$ negativas.

Como las cuatro funciones contienen a $(x_0, g(f(x_0)))$, sean L una función lineal creciente, que rebasa a las cuatro en x_0 y L' una función lineal decreciente, que es rebasada por las cuatro en x_0 . Podemos denotar a la vecindad agujerada en donde se realiza el rebase como $]a_0, b_0[\setminus \{x_0\} \subseteq]a, b[\setminus \{x_0\}$.

Sea $x \in]a_0, x_0[\subseteq]a, x_0[$, entonces $c < t(x) < f(x) < \bar{t}(x) < d$. Como

$$]c, d[=]c, f(x_0)[\cup \{f(x_0)\} \cup]f(x_0), d[$$

se tienen 3 casos.

- Si $f(x) \in]c, f(x_0)[$, entonces $h(f(x)) < g(f(x)) < \bar{h}(f(x))$. Como h es creciente, $h(t(x)) < h(f(x))$. También \bar{h} es decreciente, de modo que $\bar{h}(f(x)) < \bar{h}(t(x))$. Así $h(t(x)) < g(f(x)) < \bar{h}(t(x))$ y por los rebases con L y L' , $L(x) < g \circ f(x) < L'(x)$.
- Si $f(x) \in]f(x_0), d[$, entonces $\bar{h}(f(x)) < g(f(x)) < h(f(x))$. Como h es creciente, $h(f(x)) < h(\bar{t}(x))$. \bar{h} es decreciente, así que $\bar{h}(\bar{t}(x)) < \bar{h}(f(x))$. Así $\bar{h}(\bar{t}(x)) < f(g(x)) < h(\bar{t}(x))$ y por los rebases con L y L' , $L(x) < g \circ f(x) < L'(x)$.
- Si $f(x) = f(x_0)$, como $x < x_0$, como L es creciente, L' es decreciente y además $L(x_0) = g \circ f(x_0) = L'(x_0)$, entonces $L(x) < L(x_0) = g \circ f(x_0) < L'(x)$. De cualquier forma, si $x \in]a_0, x_0[$: $L(x) < g \circ f(x) < L'(x)$.

Ahora sea $x \in]x_0, b_0[$, como $x \in]x_0, b[$, entonces, $c < \bar{t}(x) < f(x) < t(x) < d$. Nuevamente se tienen los tres casos:

- Si $f(x) \in]c, f(x_0)[$, entonces $h(f(x)) < g(f(x)) < \bar{h}(f(x))$. Como h es creciente, $h(\bar{t}(x)) < h(f(x))$. \bar{h} es decreciente, de modo que $\bar{h}(f(x)) < \bar{h}(\bar{t}(x))$. Así $h(\bar{t}(x)) < g(f(x)) < \bar{h}(\bar{t}(x))$ y por los rebases con L y L' , $L'(x) < g \circ f(x) < L(x)$.
- Si $f(x) \in]f(x_0), d[$, entonces $\bar{h}(f(x)) < g(f(x)) < h(f(x))$. Como h es creciente, $h(f(x)) < h(t(x))$. \bar{h} es decreciente, así que $\bar{h}(t(x)) < \bar{h}(f(x))$. Así $\bar{h}(t(x)) < f(g(x)) < h(t(x))$ y por los rebases con L y L' , $L'(x) < g \circ f(x) < L(x)$.
- Si $f(x) = f(x_0)$, entonces, como L es creciente, L' decreciente y $L(x_0) = g \circ f(x_0) = L'(x_0)$: $L'(x) < g \circ f(x) < L(x)$. De cualquier forma, si $x \in]x_0, b_0[$: $L'(x) < g \circ f < L(x)$.

En resumen, si $x \in]a_0, x_0[$: $L(x) < g \circ f(x) < L'(x)$ y si $x \in]x_0, b_0[$: $L'(x) < g \circ f(x) < L(x)$ con $L(x_0) = g \circ f(x_0) = L'(x_0)$. Por ende, $g \circ f$ es continua en x_0 . \square

4.2 Continuidad en intervalos

En la primera discusión sobre la continuidad de las funciones, se mencionó que existen características *deseables* en ellas siempre que la continuidad ocurra en intervalos. Esta sección tiene como meta el mostrar que dichas particularidades están efectivamente presentes en las funciones continuas en el sentido de nuestra definición (ver Definición 4.1). No resulta muy complicado convencerse de que la continuidad de las funciones en intervalos abiertos está relacionada con la *continuidad* (o completitud) misma de la recta real. De este modo, las pruebas requerirán el uso del Axioma del Supremo en números reales.

En lo que sigue, el hecho de que si una función es continua en un punto, y el valor ahí presenta desigualdad con otro número real, existe una vecindad del punto en que dicha desigualdad se preserva (Proposición 4.7), jugará un papel muy importante, y es que en ella puede hallarse en cierta forma la esencia de la continuidad de las funciones.

Para el desarrollo del Teorema 4.13, es necesario comentar la continuidad a derecha y a izquierda, para poder definir la continuidad en un intervalo cerrado; las propiedades que se cumplen para puntos de continuidad se cumplen también cuando sólo existe continuidad lateral (ver Teorema 4.12).

Definición 4.11. Sea $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ y $x_0 \in A$, tal que existe $a \in \mathbb{R}$, tal que $]x_0, a[\subseteq A$. Diremos que f es continua en x_0 por la derecha, si existen L y T funciones lineales tales que:

1. $L(x_0) = f(x_0) = T(x_0)$.
2. Para todo $x \in]x_0, a[$, se cumple $L(x) < f(x) < T(x)$.

Sea $x_0 \in A$ tal que existe $b \in \mathbb{R}$ tal que $]b, x_0] \subseteq A$. Diremos que f es continua en x_0 por la izquierda, si existen L y T funciones lineales tales que:

1. $L(x_0) = f(x_0) = T(x_0)$.
2. Para todo $x \in]b, x_0[$, se cumple $L(x) > f(x) > T(x)$.

Diremos que f es continua en $[a, b]$ si f es continua en cada punto de (a, b) y f es continua en a por la derecha y en b por la izquierda.

Teorema 4.12. Sean $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ una función continua a derecha (izquierda) en x_0 y m , un número real, tal que $f(x_0) > m$ ($f(x_0) < m$). Existe una vecindad, $]x_0, \beta[$ ($] \alpha, x_0[$), en que se preserva dicha desigualdad; es decir:

$$\forall x \in]x_0, \beta[(] \alpha, x_0[) : f(x) > m \text{ (} f(x) < m \text{)}.$$

Demostración. Ver la demostración del Teorema 4.7. \square

Teorema 4.13 (de Bolzano). *Sea f una función real continua en $[a, b]$ con $f(a) < 0 < f(b)$. Entonces existe (al menos) un punto c del intervalo (a, b) con $f(c) = 0$.*

Demostración. Considerando una función, f , con las características de las hipótesis, considérese al conjunto:

$$A := \{x \in [a, b] \mid \forall y \in [a, x] : f(y) < 0\}.$$

Dado que $A \subseteq [a, b]$, A es acotado superiormente, además $a \in [a, b]$ y $f(a) < 0$, así que A es no vacío. El axioma del supremo garantiza que existe un número real, M , tal que $M = \sup(A)$.

Como f es continua en b y $f(b) > 0$, existe un número c con $c < b$ y $\forall x \in]c, b[$: $f(x) > 0$ (Teorema 4.12). En particular, si se toma el extremo izquierdo del intervalo intersección de $]c, b[$ y $[a, b]$ y se le llama d , se tiene que $d < b$ y que es cota superior de A . Así $M \leq d < b$ y $M < b$.

Como f es continua en a y $f(a) < 0$, existe un número e con $a < e$ y $\forall x \in]a, e[$: $f(x) < 0$ (Teorema 4.12). En particular, si se toma el extremo derecho del intervalo intersección de $]a, e[$ y $[a, b]$ y se le llama g , se tiene que $a < g$ y que $g \in A$. Así $a < g \leq M$. En resumen $a < M < b$. Ahora se procederá por reducción al absurdo.

Caso I. $f(M) < 0$.

Como $M \in]a, b[$, $f(M) < 0$ y f es continua en M , existe un número real α tal que $\alpha > M$ y $\forall x \in]M, \alpha[$: $f(x) < 0$. Como para todo $x \in [a, M[$: $f(x) < 0$ (M es supremo de A), por hipótesis $f(M) < 0$ y $\forall x \in]M, \alpha[$: $f(x) < 0$, se tiene que $\forall x \in [a, \alpha[$: $f(x) < 0$, y así M no es cota superior de A .

Caso II. $f(M) > 0$.

Si $f(M) > 0$, como f es continua en M , existe $h < M$ tal que $\forall x \in]h, M[$: $f(x) > 0$. Por otro lado, tomando $B =]h, M] \cap [a, M]$, esta intersección es no vacía y $\forall x \in B$: $0 > f(x) > 0$.

Agotados así los casos, se concluye que $f(M) = 0$. \square

Para los resultados que restan en la sección, se pueden obtener demostraciones como las que se encuentran en [1], y que se basan en los Teoremas 4.7 y 4.12, así que omitimos tales demostraciones. Los siguientes corolarios se llaman *teoremas de los valores intermedios*.

Corolario 4.14. *Sea f una función real continua en $[a, b]$ con $f(a) < f(b)$. Entonces para cualquier número c con $c \in]f(a), f(b)[$, existe un punto, x_0 , en $]a, b[$ que cumple $f(x_0) = c$*

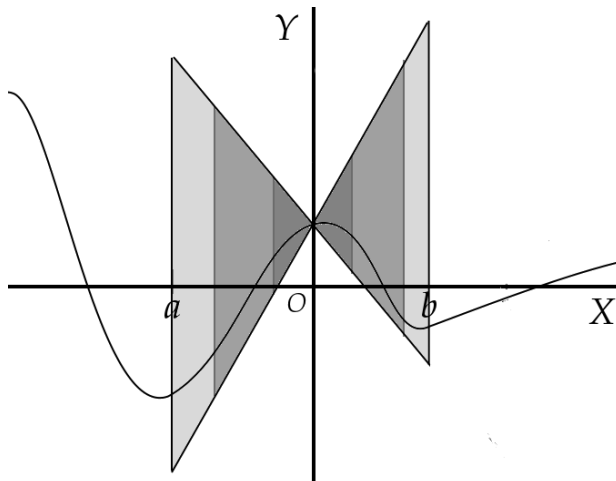


Figura 18: La función está dentro de la región sombreada y alcanza máximos y mínimos.

Corolario 4.15. *Sea f una función real continua en $[a, b]$ con $f(a) > f(b)$. Entonces para cualquier número c con $c \in]f(b), f(a)[$, existe un punto, x_0 , en $]a, b[$ que cumple $f(x_0) = c$.*

Teorema 4.16. *Sea f una función real continua en el intervalo $[a, b]$. Entonces f es acotada en $[a, b]$.*

Teorema 4.17 (De Weierstrass). *Sea f una función real continua en el intervalo $[a, b]$. Entonces existen puntos x_0 y x_1 en el intervalo $[a, b]$ (ver Figura 18), tales que para cualquier otro punto en el intervalo, x , se cumple*

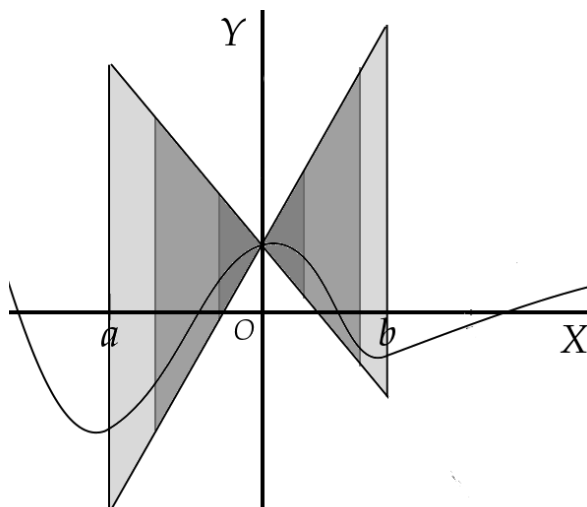
$$f(x_0) \leq f(x) \leq f(x_1).$$

5 Tangentes y derivadas

Ahora que ya se ha explorado el rebase de funciones, y en particular, el rebase de rectas no verticales ha ofrecido una singular potencia, es momento de regresar al punto de origen de esta exposición: el problema de las rectas tangentes.

Note que para una función continua en un punto, son infinitas las funciones lineales que rebasan o son adelantadas por ella, puesto que se sabe que al menos existen una recta que rebasa y otra que es rebasada por la función en el punto y posteriormente pueden tomarse funciones lineales con pendientes aún mayores o menores.

Otra forma de ver esa misma situación es mediante regiones de mariposa, pues siempre puede construirse una sucesión de regiones de mariposa donde cada una contiene propiamente a la anterior y por ende, la función queda dentro.



Si por el contrario, se quieren observar regiones de mariposa cada vez más angostas que sigan garantizando la continuidad de la función, puede procederse *cerrando* la mariposa, y esperar la existencia de un conjunto de puntos que es subconjunto de cualquier mariposa. De hecho, mediante la construcción de la región de mariposa se ha aproximado a la función cerca de un punto.

El *ancho* de la mariposa depende de la diferencia que existe entre las pendientes de las rectas que la delimitan.

Definición 5.1. Sea f una función continua en un punto x_0 . Se definen a los conjuntos:

$$S(f, x_0) = \{m \in \mathbb{R} \mid m \text{ es pendiente de una recta que rebasa a } f \text{ en } x_0\},$$

$$I(f, x_0) = \{m \in \mathbb{R} \mid m \text{ es pendiente de una recta que es rebasada por } f \text{ en } x_0\}.$$

Gracias al Teorema 3.4, se tiene que todas las funciones que rebasan o son rebasadas por f en un punto donde esta es continua contienen al par $(x_0, f(x_0))$, de modo que puede darse una definición alternativa de esos conjuntos.

$$S(f, x_0) = \{m \in \mathbb{R} \mid L_m(x) = m(x - x_0) + f(x_0) \text{ rebasa a } f \text{ en } x_0\},$$

$$I(f, x_0) = \{m \in \mathbb{R} \mid L_m(x) = m(x - x_0) + f(x_0) \text{ es rebasada por } f \text{ en } x_0\}.$$

Proposición 5.2. Sea f una función continua en un punto x_0 . Sean $m_1 \in S(f, x_0)$ y $m_2 \in I(f, x_0)$ elementos arbitrarios de sus respectivos conjuntos. Entonces se tiene que $m_1 > m_2$.

Demostración. Es una consecuencia directa del Teorema 4.2 (de rebase de rectas) y la transitividad del rebase. \square

Así pues, para una función continua f en un punto x_0 se tendrá que:

- $S(f, x_0)$ no es acotado superiormente; pero sí lo es inferiormente.
- $I(f, x_0)$ no es acotado inferiormente; pero sí lo es superiormente.

Corolario 5.3. *Sea f una función continua en un punto x_0 . Entonces $S(f, x_0) \cap I(f, x_0) = \emptyset$.*

Así, se tiene que para una función f que es continua en un punto x_0 :

- i) $S(f, x_0) \neq \emptyset$ e $I(f, x_0) \neq \emptyset$
- ii) $S(f, x_0)$ es acotado inferiormente e $I(f, x_0)$ es acotado superiormente
- iii) $\sup I(f, x_0) \leq \inf S(f, x_0)$

Cuando ocurre la igualdad $\sup I(f, x_0) = \inf S(f, x_0)$, la aproximación de la función f en x_0 mediante rectas es la máxima posible, y más aún, ese valor común determina la pendiente de una recta que caracteriza la *dirección* de f en x_0 .

Definición 5.4. Sea f una función continua en un punto x_0 . Si ocurre que

$$\sup I(f, x_0) = \inf S(f, x_0),$$

entonces se dirá que f es una función **derivable** en x_0 y que la **derivada** de f en x_0 es el valor común anterior. Este último se denotará por $f'(x_0)$. A la recta con pendiente $f'(x_0)$ que pasa por el par $(x_0, f(x_0))$ se le llama la **recta tangente** a f en x_0 .

Ahora hay que apresurar un resultado que muestra la relación entre este valor común y el resto: Todo valor mayor a $f'(x_0)$ está en $S(f, x_0)$, pues $f'(x_0) = \inf S(f, x_0)$. Por otro lado, todo valor menor a $f'(x_0)$ está en $I(f, x_0)$, pues $f'(x_0) = \sup I(f, x_0)$

Teorema 5.5. *Sea f una función continua en el punto x_0 . f es derivable en x_0 y $f'(x_0) = m_0$ si y solo si se cumplen las siguientes dos condiciones:*

- $\forall m \in \mathbb{R} : (m > m_0 \Rightarrow \text{existe una función lineal con pendiente } m \text{ que rebasa a } f \text{ en } x_0)$.
- $\forall m \in \mathbb{R} : (m < m_0 \Rightarrow \text{existe una función lineal con pendiente } m \text{ que es rebasada por } f \text{ en } x_0)$.

De hecho tales funciones son de la forma $m(x - x_0) + f(x_0)$ en ambos casos.

Con base en el Teorema 5.5 y los resultados sobre el rebase de rectas, el siguiente resultado se verifica:

Corolario 5.6. *Sea f una recta con pendiente m . La derivada de f en cualquier punto es igual a m .*

5.1 Operaciones con derivadas

Por ahora se han calculado la no muy impresionante derivada de una función lineal y (a través del análisis de la sección sobre rebases) de la función definida por $f(x) = \sqrt{1-x^2}$ en el punto 0. Más funciones derivables y el valor de sus derivadas pueden calcularse con la siguiente proposición, note que aunque el opuesto negativo y la suma de rectas ofrece nuevamente rectas, este resultado facilitará el cálculo de derivadas de otro tipo de funciones:

Proposición 5.7. Sean f y g dos funciones derivables en el punto x_0 .

1. $-f$ es derivable en x_0 y $(-f)'(x_0) = -(f'(x_0))$.
2. $f + g$ es derivable en x_0 y $(f + g)'(x_0) = f'(x_0) + g'(x_0)$.

Demostración. (2.) Como f y g son derivables en x_0 , son continuas en x_0 y así $f + g$ es continua en x_0 .

Si $m > f'(x_0) + g'(x_0)$, existe un número positivo d tal que $m = f'(x_0) + g'(x_0) + d$. La recta definida por $x \mapsto (f'(x_0) + \frac{d}{2})(x - x_0) + f(x_0)$ rebasa a f en x_0 , y la definida por $x \mapsto (g'(x_0) + \frac{d}{2})(x - x_0) + g(x_0)$ rebasa a g en x_0 . Así la suma rebasa a $f + g$, de donde la recta definida por $x \mapsto m(x - x_0) + (f + g)(x_0)$ rebasa a $f + g$ en x_0 .

Si $m < f'(x_0) + g'(x_0)$, existe un número negativo d tal que $m = f'(x_0) + g'(x_0) + d$. La recta definida por $x \mapsto (f'(x_0) + \frac{d}{2})(x - x_0) + f(x_0)$ es adelantada por f en x_0 , y la definida por $x \mapsto (g'(x_0) + \frac{d}{2})(x - x_0) + g(x_0)$ es adelantada por g en x_0 . Así la suma es adelantada por $f + g$, de donde la recta definida por $x \mapsto m(x - x_0) + (f + g)(x_0)$ es adelantada por $f + g$ en x_0 .

Así, por el Teorema 5.5, $(f + g)'(x_0) = f'(x_0) + g'(x_0)$. □

Con la Proposición 5.7 en la mano, es posible dar una condición necesaria y suficiente para que una función sea derivable en un punto basados en la aproximación que realiza su tangente.

Corolario 5.8. Una función f es derivable en x_0 y $f'(x_0) = m_0$ si y solo si la función r definida por $r(x) = f(x) - [m_0(x - x_0) + f(x_0)]$ cumple que $r'(x_0) = 0$.

Note el lector que para una función f , la función r definida en el Corolario 5.8, cumple que $r(x_0) = 0$. Un ejemplo de una función que cumple que en x_0 se anula es la definida por $x - x_0$. Un ejemplo de una función que cumple que se anula en x_0 y que su derivada también se anula en x_0 es la definida por $h(x) = (x - x_0)^2$. Para probarlo basta exhibir que dado un número positivo arbitrario, es cierto que las funciones t, \bar{t} definidas por $t(x) = m(x - x_0)$ y $\bar{t}(x) = -m(x - x_0)$ rebasa y es rebasada respectivamente por h en x_0 .

- Si $x \in]x_0 - m, x_0[$: entonces $x_0 - m < x < x_0$ y así $-m < x - x_0 < 0 < m$, por lo que $m(x - x_0) < (x - x_0)^2 < -m(x - x_0)$.
- Si $x \in]x_0, x_0 + m[$: entonces $x_0 < x < x_0 + m$ y así $-m < 0 < x - x_0 < m$, por lo que $-m(x - x_0) < (x - x_0)^2 < m(x - x_0)$.

Así, t rebasa a h y h rebasa a \bar{t} en x_0 .

Visto esto, un resultado importante que va a utilizarse es el siguiente.

Lema 5.9. *Sea r una función tal que $r(x_0) = 0$ y $r'(x_0) = 0$. Sea f una función acotada en una vecindad de x_0 . Entonces $(rf)(x_0) = 0$ y $(rf)'(x_0) = 0$.*

Demostración. El hecho de que $(rf)(x_0) = 0$ es obvio. Por otro lado, si f es una función acotada, existen $a, b, c \in \mathbb{R}$ tales que $a < x_0 < b$ y $\forall x \in]a, b[$: $|f(x)| < c$. Como $r'(x_0) = 0$, si $m > 0$, entonces existe una vecindad agujerada de x_0 en la que $\frac{m}{c}(x - x_0)$ rebasa a r y esta rebasa a $-\frac{m}{c}(x - x_0)$. Tomando una vecindad que atrapa las propiedades de las otras dos y llamándola $]j, k[\setminus\{x_0\}$, se tiene que:

- $\forall x \in]j, x_0[$: $\frac{m}{c}(x - x_0) < r(x) < -\frac{m}{c}(x - x_0)$ o lo que es lo mismo $|r(x)| < -\frac{m}{c}(x - x_0)$. Entonces $|r(x)||f(x)| < -m(x - x_0)$ y $m(x - x_0) < (rf)(x) < -m(x - x_0)$.
- $\forall x \in]x_0, k[$: $-\frac{m}{c}(x - x_0) < r(x) < \frac{m}{c}(x - x_0)$ o lo que es lo mismo $|r(x)| < \frac{m}{c}(x - x_0)$. Entonces $|r(x)||f(x)| < m(x - x_0)$ y $-m(x - x_0) < (rf)(x) < m(x - x_0)$.

Así $(rf)'(x_0) = 0$. □

Ahora pueden plantearse los problemas de la derivada de el producto y cociente de funciones derivables.

Proposición 5.10. Sean f y g dos funciones derivables en el punto x_0 .

1. fg es derivable en x_0 y $(fg)'(x_0) = f'(x_0)g(x_0) + g'(x_0)f(x_0)$.
2. Si $g(x_0) \neq 0$, entonces $\frac{1}{g}$ es derivable en x_0 y $(\frac{1}{g})'(x_0) = \frac{-g'(x_0)}{[g(x_0)]^2}$.

Demostración. (1) Va a emplearse el resultado del Corolario 5.8. Sean r_1 y r_2 dos funciones definidas por $r_1(x) = f(x) - [f'(x_0)(x - x_0) + f(x_0)]$ y $r_2(x) = g(x) - [g'(x_0)(x - x_0) + g(x_0)]$. Por el Corolario 5.8, $r_1(x_0) = 0 = r_2(x_0)$ y $r_1'(x_0) = 0 = r_2'(x_0)$. Para probar que fg es derivable en x_0 , se requiere un número m tal que

$$r(x) = (fg)(x) - [m(x - x_0) + (fg)(x_0)]$$

cumpla $r'(x_0) = 0$. Sustituyendo las expresiones para f y g en términos de r_1 y r_2 , eso ocurrirá si y solo si:

$$[r_1(x) + f'(x_0)(x - x_0) + f(x_0)][r_2(x) + g'(x_0)(x - x_0) + g(x_0)] - m(x - x_0) + (fg)(x_0)$$

cumple que su derivada se anula en x_0 . Agrupando los términos, eso ocurrirá si y solo si:

$$(r_1 r_2)(x) + f'(x_0)g'(x_0)(x - x_0)^2 + (f'(x_0)g(x_0) + g'(x_0)f(x_0) - m)(x - x_0) + r_1(x)g'(x_0)(x - x_0) + r_2(x)f'(x_0)(x - x_0) + r_1(x)g(x_0) + r_2(x)f(x_0)$$

cumple que su derivada se anula. Luego, como $r'_1(x_0) = 0 = r'_2(x_0)$ y la función definida por $h(x) = (x - x_0)$ cumple $(h^2)'(x_0) = 0$ (Lema 5.9), se tiene lo siguiente:

- $(r_1 r_2)'(x_0) = 0$ pues $r'_1(x_0) = 0$ y r_2 es derivable, continua y por tanto acotada en una vecindad de x_0 (Lema 5.9).
- $(f'(x_0)g'(x_0)h^2)'(x_0) = 0$ pues $(h^2)'(x_0) = 0$ y las funciones constantes $f'(x_0)$ y $g'(x_0)$ son acotadas.
- $(g'(x_0)r_1 h)'(x_0) = 0$ pues $r'_1(x_0) = 0$ y tanto $g'(x_0)$ como h son acotadas en una vecindad de x_0 .
- Análogamente, $(f'(x_0)r_2 h)'(x_0) = 0$, $(g(x_0)r_1)'(x_0) = 0$ y $(f(x_0)r_2)'(x_0) = 0$.

Así, para que $r'(x_0) = 0$ es necesario y suficiente que la función definida por $(f'(x_0)g(x_0) + g'(x_0)f(x_0) - m)(x - x_0)$ anule su derivada en x_0 . Esta función es lineal, de modo que su derivada es $(f'(x_0)g(x_0) + g'(x_0)f(x_0) - m)$ y esta se anula si y solo si $m = f'(x_0)g(x_0) + g'(x_0)f(x_0)$.

(2) Se busca un número m tal que la función r definida por

$$r(x) = \frac{1}{g}(x) - [m(x - x_0) + \frac{1}{g}(x_0)]$$

cumpla que $r'(x_0) = 0$. La función r puede reescribirse como:

$$r(x) = \frac{g(x_0) - m(x - x_0)g(x)g(x_0) - g(x)}{g(x_0)g(x)}.$$

La función $\frac{1}{g(x_0)g(x)}$ es acotada en una vecindad de x_0 , pues $g(x_0) \neq 0$ y así, únicamente es necesario que la función definida por el numerador en la expresión anterior anule su derivada. Note el lector que la función definida por el numerador ya es derivable en x_0 pues es suma y producto de funciones derivables y su derivada coincide con $-[mg(x_0)g(x_0) + g'(x_0)(0)] - g'(x_0)$ y esta será cero si y solo si $m = \frac{-g'(x_0)}{(g(x_0))^2}$. \square

5.2 La regla de la cadena

Nuevamente se hace mención especial para la regla de derivación de la composición de funciones. Esta regla se conoce como *regla de la cadena*. Para poder probarla, se hará uso del siguiente resultado:

Lema 5.11. *Sean f una función derivable en x_0 y r una función tal que $r(f(x_0)) = 0$ y $r'(f(x_0)) = 0$. Entonces $(r \circ f)(x_0) = 0$ y $(r \circ f)'(x_0) = 0$*

Demostración. Para comenzar, por hipótesis $(r \circ f)(x_0) = 0$. Por otro lado, sea k un número tal que $k > |f'(x_0)|$. Como $r'(f(x_0)) = 0$, entonces, si m es un número positivo, las funciones definidas por $h(y) = \frac{m}{k}(y - f(x_0))$ y $\bar{h}(y) = -\frac{m}{k}(y - f(x_0))$ cumplen que h rebasa a r y r rebasa a \bar{h} en $f(x_0)$.

Como $k > |f'(x_0)|$, entonces, las funciones definidas por $t(x) = k(x - x_0) + f(x_0)$ y $\bar{t}(x) = -k(x - x_0) + f(x_0)$ cumplen que t rebasa a f y f rebasa a \bar{t} en x_0 . Se tiene así que existen $c, d \in \mathbb{R}$ tales que $c < f(x_0) < d$ y:

- $\forall y \in]c, f(x_0)[: \frac{m}{k}(y - f(x_0)) < r(y) < -\frac{m}{k}(y - f(x_0))$
- $\forall y \in]f(x_0), d[: -\frac{m}{k}(y - f(x_0)) < r(y) < \frac{m}{k}(y - f(x_0))$

o equivalentemente

$$\forall y \in]c, d[\setminus \{x_0\} : |r(y)| < \frac{m}{k}|y - f(x_0)|$$

Como f es derivable en x_0 , en particular es continua en x_0 , de modo que existe, gracias a la Proposición 4.7, un intervalo alrededor de x_0 en el que $f(x) \in]c, d[$. Tomando una vecindad que atrape esta propiedad junto con los rebases por las funciones t y \bar{t} , y llamándola $]a, b[\setminus \{x_0\}$, se tiene que:

- $\forall x \in]a, x_0[: k(x - x_0) + f(x_0) < f(x) < -k(x - x_0) + f(x_0)$
- $\forall x \in]x_0, b[: -k(x - x_0) + f(x_0) < f(x) < k(x - x_0) + f(x_0)$
- $\forall x \in]a, b[: f(x) \in]c, d[$

o equivalentemente

- $\forall x \in]a, b[: |f(x) - f(x_0)| < k|x - x_0|$
- $\forall x \in]a, b[: f(x) \in]c, d[$.

Ahora, sea $x \in]a, b[\setminus\{x_0\}$. Se tienen dos casos:

Caso I: $f(x) = f(x_0)$. Entonces $|r(f(x))| = 0 < m|x - x_0|$.

Caso II: $f(x) \neq f(x_0)$. Entonces $f(x) \in]c, d[\setminus\{x_0\}$ por lo que $|r(f(x))| < \frac{m}{k}|f(x) - f(x_0)|$ y $|f(x) - f(x_0)| < k|x - x_0|$ de donde $|r(f(x))| < m|x - x_0|$.

De cualquier forma $\forall x \in]a, b[\setminus\{x_0\} : |(r \circ f)(x)| < m|x - x_0|$.

Si $x \in]a, x_0[$, entonces $x - x_0 < 0$ y así $m(x - x_0) < (r \circ f)(x) < -m(x - x_0)$. Si $x \in]x_0, b[$, entonces $0 < x - x_0$ y así $-m(x - x_0) < (r \circ f)(x) < m(x - x_0)$.

De este modo, si $m > 0$, se tiene que $m(x - x_0)$ rebasa a $r \circ f$ en x_0 . Ahora si $m < 0$, entonces $-m > 0$ y por lo anterior $r \circ f$ rebasa a $-(-m)(x - X - 0) = m(x - X - 0)$, se concluye que $(r \circ f)'(x_0) = 0$. \square

Ahora ya se está en posición de probar la regla para la derivación de la composición de funciones.

Proposición 5.12 (Regla de la cadena). Sean f y g dos funciones tales que f es derivable en x_0 y g es derivable en $f(x_0)$, entonces $(g \circ f)$ es derivable en x_0 y $(g \circ f)'(x_0) = g'(f(x_0))f'(x_0)$.

Demostración. Se busca un número m que force a la función r definida por $r(x) = (g \circ f)(x) - [m(x - x_0) + g(f(x_0))]$ a cumplir que $r'(x_0) = 0$. Note que siempre se cumple que $r(x_0) = 0$. Para ello se sabe que existen dos funciones r_1, r_2 definidas por

$$r_1(x) = f(x) - [f'(x_0)(x - x_0) + f(x_0)]$$

y

$$r_2(y) = g(y) - [g'(f(x_0))(y - f(x_0)) + g(f(x_0))]$$

y que satisfacen $r_1(x_0) = r_2(f(x_0)) = 0$ y $r_1'(x_0) = r_2'(f(x_0)) = 0$. Despejando g en la expresión para r_2 , y sustituyéndola en la expresión de $r(x)$, luego de evaluarla en $f(x)$, se tiene que

$$r(x) = r_2(f(x)) + g'(f(x_0))(f(x) - f(x_0)) - m(x - x_0),$$

ahora, despejando f en la expresión de r_1 y sustituyéndola en la expresión anterior, obtenemos

$$r(x) = r_2(f(x)) + g'(f(x_0))r_1(x) + g'(f(x_0))f'(x_0)(x - x_0) - m(x - x_0). \quad (2)$$

Empleando los Lemas 5.9 y 5.11, se tiene que el miembro izquierdo de la igualdad 2 es derivable y su derivada es $g'(f(x_0))f'(x_0) - m$ y esta será igual a cero si y solo si $m = g'(f(x_0))f'(x_0)$. \square

6 Teorema del valor medio

El propósito de esta corta sección es mostrar un teorema que probablemente es el más importante del cálculo diferencial y ver que es asequible a través del enfoque utilizado hasta ahora. Discusión suficiente sobre sus consecuencias y significado puede encontrarse en textos estándares de cálculo. Para ello se requieren algunos resultados auxiliares.

Lema 6.1. *Sea f una función derivable en un punto x_0 . Si $f'(x_0) > 0$, para cualesquiera $a, b \in \mathbb{R}$ con $a < x_0 < b$, existen $x_1, x_2 \in \mathbb{R}$ tales que $a < x_1 < x_0 < x_2 < b$ y $f(x_1) < f(x_0) < f(x_2)$.*

Demostración. Sean f y x_0 como en las hipótesis.

Como $f'(x_0) > 0$, $f'(x_0) > \frac{f'(x_0)}{2} > 0$. Entonces la recta definida por $L(x) = \frac{f'(x_0)}{2}(x - x_0) + f(x_0)$ es rebasada por f en x_0 , de modo que existen $c, d \in \mathbb{R}$ tales que si $c < x < x_0$, $f(x) < L(x)$ y si $x_0 < x < d$, entonces $f(x) > L(x)$. L es creciente al ser que $\frac{f'(x_0)}{2} > 0$. Sean x_1, x_2 con $x_1 \in]a, x_0[\cap]c, x_0[$ y $x_2 \in]x_0, b[\cap]x_0, d[$. Entonces $f(x_1) < L(x_1) < L(x_0) < L(x_2) < f(x_2)$, además $f(x_0) = L(x_0)$, de modo que $f(x_1) < f(x_0) < f(x_2)$. \square

Se tiene un resultado completamente análogo.

Lema 6.2. *Sea f una función derivable en un punto x_0 . Si $f'(x_0) < 0$, para cualesquiera $a, b \in \mathbb{R}$ con $a < x_0 < b$, existen $x_1, x_2 \in \mathbb{R}$ tales que $a < x_1 < x_0 < x_2 < b$ y $f(x_2) < f(x_0) < f(x_1)$.*

Estos lemas, expresan una idea geométrica simple: los puntos de una función derivable con derivada distinta de 0 no pueden ser cimas ni valles de las gráficas de las funciones, pues siempre pueden encontrarse tan cerca de ellos como se desee, otros puntos en que la función toma valores menores y mayores. Para precisar sobre estas afirmaciones:

Definición 6.3. Sea $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ una función y sea $x_0 \in A$. Decimos que $f(x_0)$ es un máximo local de f si existen $a, b \in \mathbb{R}$ tales que $a < x_0 < b$ y $\forall x \in]a, b[: f(x_0) \geq f(x)$. Decimos que $f(x_0)$ es un mínimo local de f si existen $a, b \in \mathbb{R}$ tales que $a < x_0 < b$ y $\forall x \in]a, b[: f(x_0) \leq f(x)$. Del mismo modo, si $f(x_0)$ es un máximo o mínimo local de f , se dice que f alcanza un valor máximo local o mínimo local en x_0 .

Ahora, con esa nueva terminología, el resultado siguiente es una consecuencia directa de los lemas 6.1 y 6.2.

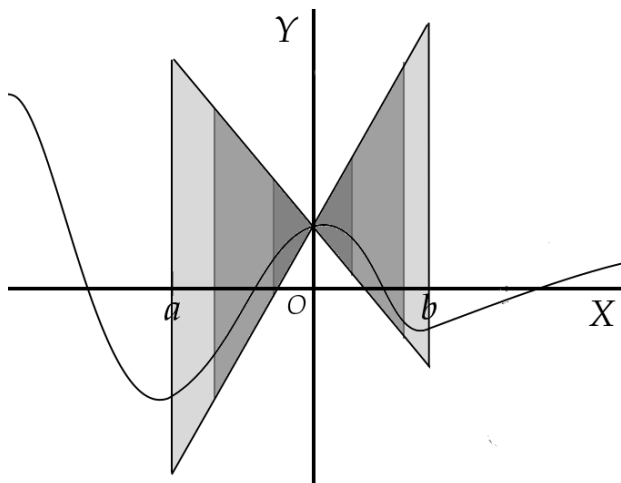


Figura 19: Los máximos y mínimos locales tienen derivada asociada nula.

Teorema 6.4 (de Fermat). *Sea f una función. Si f es derivable en x_0 y $f(x_0)$ es máximo o mínimo local de f , entonces $f'(x_0) = 0$.*

Gráficamente puede verse que en una función derivable, la recta tangente a puntos máximos o mínimos locales es horizontal (paralela al eje de las abscisas), ver Figura 19.

En el caso de las funciones continuas, es posible garantizar la existencia tanto de máximos como de mínimos, sin embargo, *la continuidad de una función no implica su derivabilidad*. Así, para poder obtener nuevos resultados a partir de lo conocido para funciones continuas, esto debe incluirse como una fuerte hipótesis. Incluimos los siguientes resultados, que se siguen del Teorema 6.4. Pueden verse las demostraciones respectivas en [1].

Teorema 6.5 (de Rolle). *Sea f una función continua en un intervalo $[a, b]$ y derivable en el correspondiente $]a, b[$ tal que $f(a) = f(b)$. Entonces existe $x_0 \in \mathbb{R}$ tal que $a < x_0 < b$ y $f'(x_0) = 0$.*

Teorema 6.6 (del valor medio de Cauchy). *Sean f y g funciones continuas en un intervalo $[a, b]$ y derivables en el correspondiente $]a, b[$. Entonces existe $x_0 \in \mathbb{R}$ tal que $a < x_0 < b$ y $f'(x_0)(g(b) - g(a)) = g'(x_0)(f(b) - f(a))$.*

Teorema 6.7 (del valor medio de Lagrange). *Sea f una función continua en un intervalo $[a, b]$ y derivable en el correspondiente $]a, b[$. Entonces existe $x_0 \in \mathbb{R}$ tal que $a < x_0 < b$ y $f'(x_0) = \frac{f(b) - f(a)}{b - a}$.*

7 Límites

En esta última sección del escrito se demuestra que el concepto de derivada mostrado aquí y la usual en los cursos de cálculo diferencial son completamente equivalentes. Para esto no se presupone un conocimiento de los conceptos de límite y derivada en los sentidos usuales.

Como se hizo notar en la sección de continuidad, para que una función sea continua en un punto es necesario que los valores que esta toma *cerca* del mismo se *parezcan* al que toma la función en él. Para esto se emplearon las regiones de mariposa, que mediante la sucesiva reducción de la longitud de sus alas permiten realizar acercamientos a dicho valor.

Suponiendo que a una función f que es continua en un punto x_0 se le quita este punto de su dominio, entonces la función resultante, h no es más continua en x_0 ; sin embargo, eso no modifica de ninguna forma que esta pueda quedar encerrada en una región de mariposa en una vecindad agujerada de x_0 . Es decir, que los valores de h se *parecen* al de $f(x_0)$ *cerca* de x_0 .

Definición 7.1. Sean $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ una función y x_0 un número real. Se dirá que **un límite de f cuando la variable tiende a x_0 es L** si y solo si existen tanto una recta que rebasa a f en x_0 como otra que es rebasada por f en x_0 y ambas contienen al par (x_0, L) .

Equivalentemente, en términos de regiones de mariposa:

Definición 7.1. Sean $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ una función y x_0 un número real. Se dirá que **un límite de f cuando la variable tiende a x_0 es L** si y sólo si existe una región de mariposa determinada por rectas l y l' que contienen a (x_0, L) y puntos $a < x_0 < b$ tales que $\forall x \in]a, b[\setminus \{x_0\} : f(x) \in M_a^b(l, l')$.

Otra manera de aproximar el valor de f *cerca* de x_0 , es: dada una función k definida en una vecindad agujerada de x_0 , puede verse que los valores de k *cerca* de x_0 se *parecen* a un cierto número L si ocurre que dado un intervalo abierto arbitrario que contenga a L , puede garantizarse que los valores de la función están también contenidos en el intervalo; aunque esta vecindad agujerada tuviera que ser muy pequeña. Nada impide, por comodidad, tomar intervalos *simétricos* alrededor de L y de x_0 , es decir que se extiendan igualmente tanto a izquierda como a derecha de cada uno de esos números. Bajo esos términos, puede imaginarse que cada que se nombre un intervalo alrededor de L (por llamarlo, V) se responda con un intervalo alrededor de x_0 (digamos H) en el que la porción correspondiente de la función quede atrapada en el *cuadro* $H \times V$. Como haciendo simétricos los intervalos, estos quedan determinados por la longitud que tienen a izquierda o a derecha del centro, hay que gestar una correspondencia entre esos números.

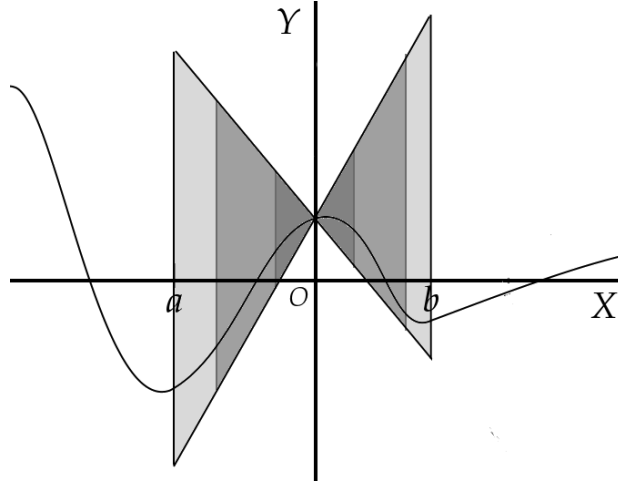


Figura 20: Para puntos cercanos a x_0 , f está contenida en $H \times V$

Proposición 7.2. Sean $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ una función y x_0 un número real. L es un límite de f cuando la variable tiende a x_0 si y solo si f está definida para valores cercanos a x_0 y existe una función $\delta :]0, \infty[\rightarrow]0, \infty[$ que tiene la siguiente propiedad: Si ocurre que $0 < |x - x_0| < \delta(\varepsilon)$, entonces $|f(x) - L| < \varepsilon$.

Demostración. Como toda proposición de equivalencia, la prueba se dividirá en dos partes.

(Primera parte) Suponiendo que:

- $\exists m \in \mathbb{R}$ tal que la recta definida por $L + m(x - x_0)$ es rebasada por f en x_0 y
- $\exists M \in \mathbb{R}$ tal que la recta definida por $L + M(x - x_0)$ rebasa a f en x_0 ,

como $m \neq M$, al menos uno de ambos números es distinto de cero. Sea α un número real tal que $\alpha > \max\{|m|, |M|\}$; de este modo $\alpha > M$ y $-\alpha < m$. Definiendo:

$$\begin{array}{llll} L_1 : \mathbb{R} \rightarrow \mathbb{R} & T : \mathbb{R} \rightarrow \mathbb{R} & \bar{T} : \mathbb{R} \rightarrow \mathbb{R} & L_2 : \mathbb{R} \rightarrow \mathbb{R} \\ x \mapsto L + m(x - x_0) & x \mapsto L + \alpha(x - x_0) & x \mapsto L - \alpha(x - x_0) & x \mapsto L + M(x - x_0) \end{array}$$

se tiene que T rebasa a L_1 y L_2 rebasa a \bar{T} en x_0 y, por transitividad, T rebasa a f y f rebasa a \bar{T} en x_0 . Eligiendo una vecindad que atrape ambos rebases y llamando a sus extremos inferior y superior a y b respectivamente:

- $\forall x \in]a, x_0[: L + \alpha(x - x_0) < f(x) < L - \alpha(x - x_0)$ y
- $\forall x \in]x_0, b[: L + \alpha(x - x_0) > f(x) > L - \alpha(x - x_0)$.

De ahí que

- $\forall x \in]a, x_0[: \alpha(x - x_0) < f(x) - L < -\alpha(x - x_0)$ y
- $\forall x \in]x_0, b[: \alpha(x - x_0) > f(x) - L > -\alpha(x - x_0)$.

Con el uso de las propiedades elementales de las desigualdades:

- $\forall x \in]a, x_0[: |f(x) - L| < -\alpha(x - x_0)$ y
- $\forall x \in]x_0, b[: |f(x) - L| < \alpha(x - x_0)$.

Y en suma $\forall x \in]a, b[\setminus\{x_0\} : |f(x) - L| < |\alpha(x - x_0)|$.

Sean ε un número real positivo y se define $\delta(\varepsilon) = \min\{b - x_0, x_0 - a, \frac{\varepsilon}{\alpha}\}$. Si x es tal que $x_0 - \delta(\varepsilon) < x < x_0 + \delta(\varepsilon)$, entonces $x \in]a, b[\setminus\{x_0\}$ y por tanto $|f(x) - L| < \alpha|x - x_0|$. Por otro lado $\alpha|x - x_0| < \alpha\delta(\varepsilon) < \alpha\frac{\varepsilon}{\alpha}$. Finalmente si $0 < |x - x_0| < \delta(\varepsilon)$, entonces $|f(x) - L| < \varepsilon$.

(*Segunda parte*) Desmostraremos la contrarrecíproca de la proposición a demostrar. O sea, vamos a probar que para todo $L \in \mathbb{R}$, que si no es cierto que existan rectas que contienen a (x_0, L) , que rebasen a f y rectas que contienen a (x_0, L) , que sean rebasadas por la función f en x_0 , entonces ningún número real será límite de la función.

Dado $L \in \mathbb{R}$, sean

$$A_L = \{M \in \mathbb{R} : T(x) = L + M(x - x_0), \text{ rebasa a } f \text{ en } x_0\} \text{ y}$$

$$B_L = \{N \in \mathbb{R} : H(x) = L + N(x - x_0), \text{ es rebasada por } f \text{ en } x_0\}.$$

La hipótesis entonces es equivalente a que para cualquier número real L , alguno de los conjuntos es vacío, y va a probarse que $\forall L \in \mathbb{R} : \exists \varepsilon > 0$ tal que $\forall \delta > 0 : \exists x \in \mathbb{R}$ para el que se cumplen $0 < |x - x_0| < \delta$ y $|f(x) - L| \geq \varepsilon$.

Se analizará el caso, en que para $L \in \mathbb{R}$ tenemos que $B_L = \emptyset$, puesto que el otro es completamente análogo. Como $B_L = \emptyset$, para $M < 0$ y $\delta > 0$

- $\exists x_1 \in]x_0 - \delta, x_0[: f(x_1) \geq L + M(x_1 - x_0)$ o
- $\exists x_2 \in]x_0, x_0 + \delta[: f(x_2) \leq L + M(x_2 - x_0)$.

Sea $\varepsilon = \min\{M(x_1 - x_0), -M(x_2 - x_0)\}$.

Caso I: $\exists x_1 \in]x_0 - \delta, x_0[: f(x_1) - L \geq M(x_1 - x_0) > 0$. Entonces $|x_1 - x_0| < \delta$ y $|f(x_1) - L| = f(x_1) - L \geq M(x_1 - x_0) \geq \varepsilon$.

Caso II: $\exists x_2 \in]x_0, x_0 + \delta[: f(x_2) - L \leq M(x_2 - x_0) < 0$. Entonces $f(x_2) - L \leq M(x_2 - x_0) < 0$. Así, $|x_2 - x_0| < \delta$ y $|f(x_2) - L| = L - f(x_2) \geq -M(x_2 - x_0) \geq \varepsilon$. \square

En la Proposición 7.2 se encuentra el enunciado usual de la definición de límite de una función en un punto.

Gracias a la completa analogía entre las definiciones 7.1 de límite y 4.1 de continuidad, un equivalente de la proposición 4.7 es válida para el caso de límites, es decir.

Proposición 7.3. Sean $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ una función, x_0 un punto en el que existe el límite de f cuando la variable tiende a x_0 , L , y m , un número real, tal que $L > m$ ($L < m$). Existe una vecindad agujerada de x_0 , $] \alpha, \beta [\setminus \{x_0\}$ en que se preserva dicha desigualdad; es decir:

$$\forall x \in] \alpha, \beta [\setminus \{x_0\} : f(x) > m (f(x) < m).$$

Con ayuda de la proposición 7.3, es fácil probar que de existir un límite de una función en un punto, este es único:

Teorema 7.4. Sea f una función y x_0 un punto en el que existe un límite de f cuando la variable tiende a x_0 . Si este límite existe es único.

Demostración. Ver [1]. □

Observación 7.5. Dado que el límite de una función en un punto es único, se emplea la notación

$$\lim_{x \rightarrow x_0} f(x) = L$$

para denotar que el límite de f cuando la variable tiende a x_0 es L . La letra x no es determinante, de modo que lo mismo significará esa notación que, por ejemplo:

$$\lim_{t \rightarrow x_0} f(t) = L \quad \lim_{y \rightarrow x_0} f(y) = L.$$

Observación 7.6. Por último, note el lector que una nueva caracterización de la continuidad viene dada por la definición 7.1 y la observación 7.5: Una función f es continua en un punto x_0 si y solo si $\lim_{x \rightarrow x_0} f(x) = f(x_0)$.

Para una función continua f en un punto x_0 , se definieron los conjuntos $S(f, x_0)$ e $I(f, x_0)$ y si ocurría que $I(f, x_0) = \inf S(f, x_0)$ entonces se llamaba a ese valor común la derivada de f en x_0 , que gráficamente representa la pendiente de la recta tangente a f en x_0 . Otra forma de intentar encontrar el valor de dicha pendiente es mediante la noción de límite, puesto que como se mencionó, la tangente a una curva es una buena aproximación a la misma cerca del punto de tangencia. Es decir, si m es la pendiente de la recta tangente, se cumple que para puntos x cercanos a x_0 , $f(x) \approx m(x - x_0) + f(x_0)$ o lo que es lo mismo

$$m \approx \frac{f(x) - f(x_0)}{x - x_0}.$$

Teorema 7.7. *Sea f una función definida en una vecindad de un punto. La función f es derivable en x_0 y $f'(x_0) = m_0$, si y solo si*

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = m_0$$

Demostración. (Primera parte:) Supóngase que $\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = m_0$. De este modo se define una función h por

$$h(x) = \begin{cases} \frac{f(x) - f(x_0)}{x - x_0} & \text{si } x \in \text{Dom}(f), x \neq x_0 \\ m_0 & \text{si } x = x_0 \end{cases}$$

Como $\lim_{x \rightarrow x_0} h(x) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = m_0 = h(x_0)$, h es continua en x_0 .

Si $m > m_0$, entonces, por los resultados de continuidad de la Proposición 4.7, existen $a, b \in \mathbb{R}$ con $a < x_0 < b$ tales que $\forall x \in]a, b[\setminus \{x_0\} : h(x) < m$. Entonces $\forall x \in]a, b[\setminus \{x_0\} : \frac{f(x) - f(x_0)}{x - x_0} < m$.

- Si $x \in]a, x_0[$, $x - x_0 < 0$, entonces $f(x) - f(x_0) > m(x - x_0)$ y $f(x) > m(x - x_0) + f(x_0)$.
- Si $x \in]x_0, b[$, $x - x_0 > 0$, entonces $f(x) - f(x_0) < m(x - x_0)$ y $f(x) < m(x - x_0) + f(x_0)$.

Si se define $L_1 : \mathbb{R} \rightarrow \mathbb{R}$ por $L_1(x) = m(x - x_0) + f(x_0)$, se ve que L_1 contiene a $(x_0, f(x_0))$ y que L_1 rebasa a f en x_0 .

Del mismo modo, si $m < m_0$, existen $c, d \in \mathbb{R}$ con $c < x_0 < d$ tales que $\forall x \in]c, d[\setminus \{x_0\} : h(x) > m$. Entonces $\forall x \in]c, d[\setminus \{x_0\} : \frac{f(x) - f(x_0)}{x - x_0} > m$, y la recta L_2 definida por $L_2(x) = m(x - x_0) + f(x_0)$ contiene a $(x_0, f(x_0))$ y es rebasada por f en x_0 . De este modo, f es continua, derivable y $m_0 = f'(x_0)$.

(Segunda parte:) Para esta parte se empleará el Teorema 7.2. Si f es derivable en x_0 con $f'(x_0) = m_0$ y $\varepsilon > 0$; como $m_0 - \varepsilon < m_0 < m_0 + \varepsilon$, existen $a, b \in \mathbb{R}$ con $a < x_0 < b$ tales que (ver Teorema 5.5):

- si $x \in]a, x_0[$: $(m_0 + \varepsilon)(x - x_0) + f(x_0) < f(x) < (m_0 - \varepsilon)(x - x_0) + f(x_0)$;
- si $x \in]x_0, b[$: $(m_0 - \varepsilon)(x - x_0) + f(x_0) < f(x) < (m_0 + \varepsilon)(x - x_0) + f(x_0)$.

Entonces:

- si $x \in]a, x_0[$: $(m_0 + \varepsilon)(x - x_0) < f(x) - f(x_0) < (m_0 - \varepsilon)(x - x_0)$;
- si $x \in]x_0, b[$: $(m_0 - \varepsilon)(x - x_0) < f(x) - f(x_0) < (m_0 + \varepsilon)(x - x_0)$.

Entonces:

- si $x \in]a, x_0[$: $-\varepsilon < \frac{f(x)-f(x_0)}{x-x_0} - m_0 < \varepsilon$;
- si $x \in]x_0, b[$: $-\varepsilon < \frac{f(x)-f(x_0)}{x-x_0} - m_0 < \varepsilon$.

De cualquier forma, si $x \in]a, b[\setminus \{x_0\}$, entonces:

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - m_0 \right| < \varepsilon .$$

Se define $\delta(\varepsilon) = \min(b - x_0, x_0 - a)$. Si $0 < |x - x_0| < \delta(\varepsilon)$, entonces

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - m_0 \right| < \varepsilon . \quad \square$$

Agradecimientos

Queremos agradecer al anónimo arbitro que leyó con cuidado nuestro trabajo, proponiendo cambios que lo mejoraron en mucho.

Bibliografía

- [1] Juan Angoa et al., *Cálculo Diferencial en una variable*. Textos científicos. 2005. 81–87, 111–121, 125–130.
- [2] Juan Angoa et al., *Matemáticas Elementales*. Textos científicos. 2010. 15–22, 73–76, 121–129.
- [3] Jerrold Marsden, *Calculus Unlimited*. The Benjamin/Cummings Publishing Company Inc. 1981. 1–31.
- [4] Michael Spivak, *Calculus*. Reverté. 2012. 90–99.

Facultad de Ciencias Físico Matemáticas, BUAP
 Avenida San Claudio y 18 Sur, Colonia San Manuel,
 Puebla, Pue. C.P. 72570
 jangoa@fcfm.buap.mx
 fernando.cocolletzi.a@gmail.com

Capítulo 4

Los números perfectos y los números primos de Mersenne, a través del tiempo

Aarón Aparicio Hernández

Universidad Autónoma de la Ciudad de México;
Facultad de Ciencias, UNAM.

Resumen

Los números primos de Mersenne son llamados así a partir del siglo XVII, en honor al monje Francés Marín Mersenne (1588-1648). Mersenne fue amigo de Descartes, Fermat y Pascal; a él se le da el crédito porque investigó los números de la forma $M_p = 2^p - 1$ en 1644. Mersenne demostró que M_p es primo para $p = 2, 3, 5, 7, 13, 17, 19$ y conjeturó que M_p también es primo para $p = 31, 67, 127, 257$. Sin embargo para $p = 67, 257$ M_p son números compuestos. En este trabajo revisamos como se han encontrado durante los últimos dos mil años.

*A don Alberto Barajas en el centenario,
de su nacimiento.*

1 Números perfectos y números primos de Mersenne

Definición 1.1. Si $M_n = 2^n - 1$ es un número primo, entonces M_n se llama *primo de Mersenne*.

Proposición 1.1. Si $M_n = 2^n - 1$ es primo, entonces n es primo.

Demostración. Supongamos que n no es primo, entonces $n = pq$ con $p > 1$ y $q > 1$, por lo tanto $M_n = 2^n - 1 = 2^{pq} - 1 = (2^p)^q - 1 = (2^p - 1)[(2^p)^{q-1} + (2^p)^{q-2} + \dots + 1]$, así M_n es compuesto y por lo tanto M_n no es primo. \square

Definición 1.2. Un número n se llama *número perfecto* si al sumar todos los divisores positivos y menores que n es igual a n .

Ejemplo 1.1. Consideremos el número 6 y sus divisores positivos menores que 6, ellos son 1, 2, 3 y por consiguiente $1 + 2 + 3 = 6$, por lo que 6 resulta ser un número perfecto. Por otra parte, si consideramos solamente los divisores positivos de 6 obtenemos 1, 2, 3, 6; por lo tanto $1 + 2 + 3 + 6 = 12 = 2(6)$. Por lo que podemos enunciar una definición alternativa de número perfecto. Dado un número n , $\sigma(n)$ denota la suma de los divisores positivos de n . Por consiguiente, si $\sigma(n) = 2n$ entonces n se llama *número perfecto*.

Teorema 1.2 (Euclides). Si $M_p = 2^p - 1$ es primo, entonces $P_p = 2^{p-1}(2^p - 1)$ es un número perfecto.

Teorema 1.3 (Euler). Si n es un número par y perfecto, entonces $n = 2^{p-1}(2^p - 1)$ donde $2^p - 1$ es primo.

Observación 1.2. La importancia de los teoremas de Euclides y de Euler es fundamental porque caracterizan a los números perfectos pares, esto es, cada vez que se encuentre un nuevo número primo de Mersenne automáticamente obtenemos un número perfecto.

1, 2, 3, 4

Los primeros cuatro números perfectos pares son conocidos desde 100 a.c. Theón de Smyrna (130 d.c.) y Nicomachus de Gerasa (100 d.c.) ambos dan la misma definición de número perfecto, así como las propiedades que tienen, las cuales Euclides probó en el libro 9, proposición 36 de los Elementos. Posteriormente los matemáticos continuaron buscando y encontrando números perfectos. Hrotsvit (935-1002) y la primera poeta alemana, enlistaron los primeros cuatro números perfectos en la *Tocata Sapientia* en el siglo X [3].

5, 6, 7

El quinto número perfecto es citado por el año 1461 en un manuscrito cuyo autor es desconocido. Hudalricus Regius obtiene el quinto número perfecto en su epitomo en 1536 y demuestra que $M_{11} = 2047 = 23(89)$ no es primo, además empieza a estudiar los números deficientes, esto es, los números para los cuales $\sigma(n) < 2n$. En 1603, Pietro Cataldi de Bologna un entusiasta acerca de los números perfectos publica *tratato de numeri perfetti* (tratado de números perfectos). En este trabajo indica que escribió en 1588, el quinto, sexto y séptimo número perfecto; también verificó que M_{13} , M_{17} y M_{19} son primos, accidentalmente demostró que M_{11} no es primo.



(a) Cataldi



(b) Euler

Figura 1

8

Leonhard Euler el más prolífico de todos los matemáticos, tiene el crédito al haber encontrado el octavo número perfecto M_{31} en 1772. En una carta a Christian Goldbach en 1752, Euler lista correctamente los primeros 7 números perfectos, pero al mismo tiempo afirma que M_{31} es primo. En una carta a Johann Bernoulli en 1772, Euler empieza a demostrar que M_{31} es necesariamente un número primo. Peter Barlow en su libro *Theory of numbers* (Teoría de los números) publicado en 1811 escribe acerca del octavo número perfecto: «Es el más grande, eso alguna vez se descubrirá pero, como ellos es meramente curioso sin ser útil, no es probable que cualquier persona intentará encontrar uno más allá de él».

Observación 1.3. Como veremos a lo largo del artículo Peter Barlow estaba equivocado, pues se siguen investigando en la actualidad.

9

El noveno número de Mersenne M_{61} fue hallado por J. Pervushin en 1883, pero el resultado no se conoció hasta 1887. En 1886 P. Seelhoff de manera independiente estableció el mismo resultado. Jules Hudelot usó un método diferente en 1887 para determinar que M_{61} es primo, este cálculo le tomó 54 horas utilizando el algoritmo que había encontrado Edouard Lucas en 1883. Marín Mersenne hizo una atrevida afirmación de que M_{67} era primo, esta afirmación no fue discutida por más de 250 años; hasta que en 1903 el profesor Frank Nelson Cole, de la Universidad de Columbia dio una conferencia bajo el título discreto *sobre la factorización de números grandes* en una reunión de la American Mathematical Society. Cole que siempre fue un hombre de pocas palabras caminó hasta el pizarrón y sin decir nada, tomó el gis y comenzó a calcular 2^{67} y posteriormente le restó 1, obteniendo un número de 21 dígitos a saber 147, 573, 952, 589, 676, 412, 927. Sin decir una palabra,

pasó a un espacio en blanco en el pizarrón y multiplicó a mano, 193, 707, 721 por 761, 838, 257, 287, las dos operaciones coincidieron. Por primera vez que se recuerde, los asistentes aplaudieron vigorosamente al autor. Cole volvió a su asiento sin haber pronunciado una sola palabra, nadie le hizo ninguna pregunta. Posteriormente, ya en el pasillo comentó que tal hazaña le había tomado las tardes de los domingos durante veinte años para encontrar la factorización.



(a) Lucas



(b) Cole

Figura 2

10, 11

R. E. Powers verificó que M_{89} es primo en 1911 usando el algoritmo de E. Lucas [6], E. Fauquemberg también probó que M_{89} es primo usando una técnica que consiste en examinar residuos en base dos. Ambos matemáticos demostraron que M_{107} es primo en 1914 de manera independiente.

12

E. Lucas en 1876 estableció que M_{127} es primo y en 1877 demostró tal afirmación. En 1914 E. Fauquemberg también demostró que M_{127} es primo. Además el número perfecto P_{127} asociado a M_{127} en notación decimal tiene 77 dígitos y este número perfecto fue hallado sin la ayuda de las computadoras. Durante más de 2000 años solo se conocieron 12 números perfectos y los 12 números perfectos siguientes, se encontraron en tan solo 100 años. En 1930 Derrick Henry Lehmer publicó un artículo extendiendo el algoritmo de Lucas, este resultado se le conoce como el algoritmo de Lucas-Lehmer.

Teorema 1.4 (Algoritmo de Lucas-Lehmer). $M_p = 2^p - 1$ es primo si y solo si $S_{p-1} \equiv 0 \pmod{M_p}$, donde S_k es el mínimo residuo $\pmod{M_p}$ definido recursiva-

mente como sigue:

$$S_1 = 4, \quad S_k \equiv S_{k-1}^2 - 2 \pmod{M_p} \quad \forall k \geq 2.$$

Ejemplo 1.4. Usando el criterio de Lucas-Lehmer demuestre que M_{13} es un número primo.

Demostración. Como $p = 13$ entonces $M_{13} = 2^{13} - 1 = 8191$. Ahora apliquemos el criterio de Lucas-Lehmer:

$$\begin{array}{llll} S_1 & = & 4 & \\ S_2 & \equiv & 4^2 - 2 & \equiv 16 - 2 \equiv 14 \pmod{M_{13}} \\ S_3 & \equiv & 14^2 - 2 & \equiv 196 - 2 \equiv 194 \pmod{M_{13}} \\ S_4 & \equiv & 194^2 - 2 & \equiv 37636 - 2 \equiv 4870 \pmod{M_{13}} \\ S_5 & \equiv & 4870^2 - 2 & \equiv 23716900 - 2 \equiv 3953 \pmod{M_{13}} \\ S_6 & \equiv & 3953^2 - 2 & \equiv 15626209 - 2 \equiv -2221 \pmod{M_{13}} \\ S_7 & \equiv & (-2221)^2 - 2 & \equiv 4932841 - 2 \equiv 1857 \pmod{M_{13}} \\ S_8 & \equiv & 1857^2 - 2 & \equiv 3448449 - 2 \equiv 36 \pmod{M_{13}} \\ S_9 & \equiv & 36^2 - 2 & \equiv 1296 - 2 \equiv 1294 \pmod{M_{13}} \\ S_{10} & \equiv & 1294^2 - 2 & \equiv 1674436 - 2 \equiv 3470 \pmod{M_{13}} \\ S_{11} & \equiv & 3470^2 - 2 & \equiv 12040900 - 2 \equiv 128 \pmod{M_{13}} \\ S_{12} & \equiv & 128^2 - 2 & \equiv 16384 - 2 \equiv 0 \pmod{M_{13}} \end{array}$$

Por lo tanto, M_{13} es un número primo. □

Ejemplo 1.5. Usando el criterio de Lucas-Lehmer demuestre que M_{11} no es un número primo.

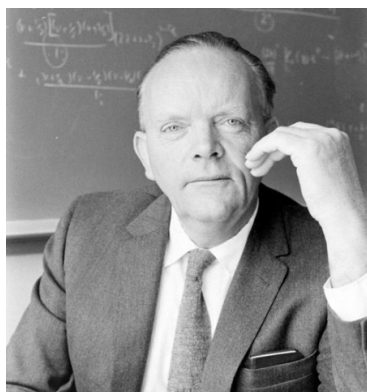
Demostración. Como $p = 11$ entonces $M_{11} = 2^{11} - 1 = 2047$. Ahora apliquemos el criterio de Lucas-Lehmer:

$$\begin{array}{llll} S_1 & = & 4 & \\ S_2 & \equiv & 4^2 - 2 & \equiv 16 - 2 \equiv 14 \pmod{M_{11}} \\ S_3 & \equiv & 14^2 - 2 & \equiv 196 - 2 \equiv 194 \pmod{M_{11}} \\ S_4 & \equiv & 194^2 - 2 & \equiv 37636 - 2 \equiv 788 \pmod{M_{11}} \\ S_5 & \equiv & 788^2 - 2 & \equiv 620944 - 2 \equiv 701 \pmod{M_{11}} \\ S_6 & \equiv & 701^2 - 2 & \equiv 491401 - 2 \equiv 119 \pmod{M_{11}} \\ S_7 & \equiv & 119^2 - 2 & \equiv 14161 - 2 \equiv -170 \pmod{M_{11}} \end{array}$$

$$\begin{aligned}
 S_8 &\equiv (-170)^2 - 2 \equiv 28900 - 2 \equiv 240 \pmod{M_{11}} \\
 S_9 &\equiv 240^2 - 2 \equiv 57600 - 2 \equiv 282 \pmod{M_{11}} \\
 S_{10} &\equiv 282^2 - 2 \equiv 79524 - 2 \equiv 1736 \not\equiv 0 \pmod{M_{11}}
 \end{aligned}$$

Por lo tanto, M_{11} no es un número primo; de hecho, $M_{11} = 2^{11} - 1 = 2047 = 23(89)$. \square

13, 14, 15, 16, 17



(a) Lehmer



(b) Robinson

Figura 3

En 1952 la Western Automatic Computer (SWAC) del Buró Nacional de Estándares en los Angeles California, fue programado por Raphael Mitchel Robinson para buscar nuevos primos de Mersenne. Cinco nuevos números perfectos fueron descubiertos como resultado de este esfuerzo. SAWC demostró que los números M_{521} y M_{607} son primos en la noche del 30 de enero de 1952 y M_{1279} el 25 de junio de 1952. El décimo sexto número perfecto fue encontrado el 7 de octubre de 1952 y el siguiente fue descubierto sólo dos días después.

18

Hans Riesel usando la computadora electrónica digital BESK encontró el décimo octavo número perfecto el 8 de septiembre de 1957.

19, 20

En noviembre de 1961, Alexander Hurwitz anunció que había encontrado los números de Mersenne 19 y 20 de la lista. Usando una computadora IBM7090 de la

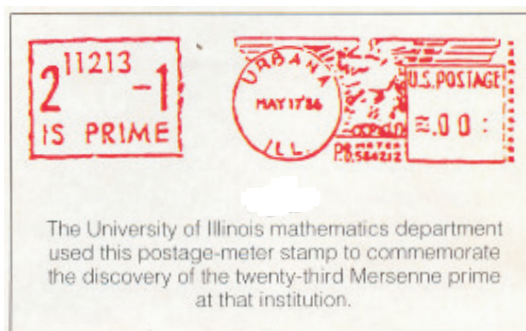
UCLA y requirió sólo de 50 minutos para determinar M_{4423} .

21, 22, 23

Los números de Mersenne 21,22 y 23 de la lista fueron descubiertos por Donald B. Gillies en 1963, con ayuda de la computadora ILLIAC de la Universidad de Illinois. En el primero se tardó 1 hora con 32 minutos, en el segundo 1 hora con 30 minutos y en el tercero 2 horas y 15 minutos. Para justificar la hazaña, los matemáticos del Departamento de Matemáticas de la Universidad de Illinois usaron un sello postal para proclamar el primo de Mersenne más grande que se conocía (ver figura 4).



(a) Gillies



(b) Sello postal

Figura 4

24

Este primo de Mersenne fue encontrado por Bryant Tuckerman, usando el algoritmo de Lucas-Lehmer el 4 de marzo de 1971. Tuckerman hizo este cálculo en una computadora IBM 360/91 del centro de Investigación Thomas J. Watson de IBM en Nueva York, tardándose 39 minutos y 44 segundos. Para conmemorar este acontecimiento IBM utilizó como membrete la siguiente figura.

25, 26

El vigésimo quinto primo de Mersenne fue descubierto por dos jóvenes de 18 años, Laura Nickel y Landon Curt Noll; después de 3 años de investigación el 30 de octubre de 1978, usando una computadora Cyber-174 de la Universidad de California en Hayward. Este logro fue proclamado al público americano en la noticias de la tarde de la CBS por Walter Cronkite. En febrero de 1979 Noll encontró el vigésimo sexto primo de Mersenne, usando la misma computadora. Dos semanas después, el 23 de

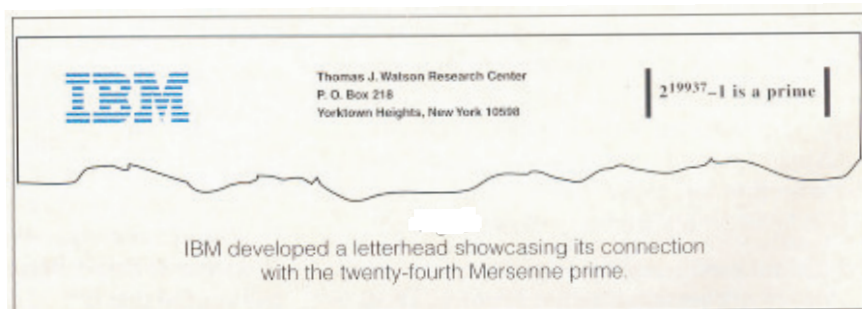


Figura 5

febrero David Slowinski del Centro de Investigación Cray, encontró el mismo número de manera independiente. La búsqueda de Noll le tomó ochos horas y cuarenta minutos en la Cyber-174, mientras que Slowinski lo hizo en menos de siete minutos en una supercomputadora Cray-1 con mayor velocidad.



Figura 6: Noll

27

El vigésimo séptimo primo de Mersenne ($2^{44497} - 1$) fue encontrado el 8 de abril de 1979 por Harry Nelson y David Slowinski. El programa fue corrido en una supercomputadora Cray-1 y para encontrar este primo en particular, se aplicó el algoritmo de Lucas-Lehmer a casi mil números.

28

En Julio de 1983 el vigésimo octavo primo de Mersenne ($2^{86243} - 1$) fue hallado por David Slowinski, usando una supercomputadora Cray-1 tomándole 600 horas de cálculos computacionales. En esa época David Slowinski era un experto analista en sistemas en el centro de investigación Cray en Chippewa Falls, Wisconsin.



Figura 7: Slowinski

29

El 23 de septiembre de 1983, David Slowinski anunció en el periódico Los Angeles Times, haber encontrado el vigésimo noveno primo de Mersenne ($2^{132049} - 1$), este primo fue hallado con una computadora experimental en el Centro de Investigación Cray. Esta computadora la tenía trabajando durante la mañana y sorprendentemente encontró el primo de Mersenne a la hora del desayuno. En el orden de importancia, este número se convertiría en el trigésimo primo de Mersenne, mientras que el primo de Mersenne ($2^{110503} - 1$) sería encontrado en 1988 por Walter N. Colquitt y Luther Welsh Jr.

30

El 18 de septiembre de 1985, científicos de la Compañía de Geociencias de Chevron en Houston Texas, anunciaron haber encontrado el trigésimo primo de Mersenne ($2^{216091} - 1$), utilizando un programa de Slowinski en una supercomputadora Cray X-MP. La prueba para comprobar la primalidad le tomó tres horas en un rango de 400,000,000 de cálculos por segundo. Sin embargo, varios meses de trabajo se necesitaron para determinar, que este número era un candidato razonable para la prueba. Por orden de tamaño, este número sería el trigésimo primero primo de Mersenne.

31

Walter N. Colquitt del Centro de Investigación del Área de Houston y el consultor en sistemas Luther Welsh Jr. de El Toro California, tienen el crédito por haber encontrado el trigésimo primero primo de Mersenne ($2^{110503} - 1$), como fue anunciado en la Revista Science News publicada el 6 de febrero de 1988. En una supercomputadora NEC SX-2 se usó un programa completamente escrito en Fortran y le tomó cerca de once minutos para confirmar que ($2^{110503} - 1$) es primo. Colquitt había estado verificando algunos cálculos que no habían sido investigados exhaustivamente en el pasado. El trigésimo primero primo de Mersenne apareció en el intervalo

comprendido entre los dos primos de Mersenne previamente conocidos, a saber el vigésimo octavo ($2^{86243} - 1$) y el vigésimo noveno ($2^{132049} - 1$).

32

En junio de 1992 el Centro de Investigación Cray, anunció el trigésimo segundo primo de Mersenne ($2^{756839} - 1$), el cual fue encontrado por David Slowinski y Paul Gage del Laboratorio Harwell que era una autoridad en Energía atómica del Reino Unido, utilizando una supercomputadora Cray-2 y el algoritmo de Lucas-Lehmer. La primalidad de este número fue verificada después de diecinueve horas de cálculos computacionales.

33

El cuatro de enero de 1994 David Slowinski anunció haber encontrado junto con Paul Gage, el trigésimo tercero primo de Mersenne ($2^{859433} - 1$). El algoritmo requirió siete horas y 12 minutos de cálculos en una supercomputadora Cray C90. Posteriormente, Richard Crandall verificó la primalidad de manera independiente.

34

En septiembre de 1996, el Centro de Investigación Cray, que era una subsidiaria de Silicon Graphics, anunció el descubrimiento del trigésimo cuarto primo de Mersenne ($2^{1257787} - 1$) por David Slowinski y Paul Gage mientras realizaban unas pruebas en una supercomputadora Cray T94, en Chippewa Falls, Wisconsin. Este número tiene 378,632 dígitos. Si quisiéramos imprimir este número en un periódico, llenaría aproximadamente doce páginas. La prueba de primalidad tomó alrededor de seis horas en una unidad de procesamiento central de la Cray T94. El trigésimo cuarto primo de Mersenne fue descubierto en el verano de 1996, pero el anuncio se retrasó hasta que hubo otra comprobación. El programa para encontrar números primos fue desarrollado por Slowinski y Gage, y es utilizado por el Centro de Investigación Cray, como una prueba de garantía de la calidad rigurosa sobre todos los elementos de sus sistemas de supercomputadoras. No era claro si este número primo era el trigésimo cuarto de la lista de primos de Mersenne, ya que faltaba comprobar si no había otros primos de Mersenne menores.

35

Internet fue la responsable para determinar el trigésimo quinto número perfecto. Al inicio de 1996, ante la insistencia de George Woltman de Orlando Florida por

The screenshot shows the homepage of the Great Internet Mersenne Prime Search (GIMPS). At the top, there is a navigation bar with the GIMPS logo and the text "Great Internet Mersenne Prime Search" and "Finding World Record Primes Since 1996". A Google Translate widget is visible in the top right corner. Below the header, there is a prominent announcement: "Largest Known Prime, 48th Known Mersenne Prime Found!!". The main text of the announcement states that on January 25th, 2013, Dr. Curtis Cooper discovered the 48th known Mersenne prime, $2^{57,885,161} - 1$, a 17,425,170 digit number. The page also includes a sidebar with navigation links such as "Login", "Donate", "Settings", "Download Software", and "Manual Testing".

Figura 8: Fundación GIMPS

el Gran Buscador de Números Primos de Mersenne de la Internet (Great Internet Mersenne Prime Search (GIMPS)) fue iniciado. El 13 de noviembre, Joel Armengaud de París, Francia, fue el afortunado que descubrió el trigésimo quinto primo de Mersenne ($2^{1398269} - 1$), el cual consta de 420,991 dígitos, utilizando un programa escrito por Woltman. En realidad el hallazgo era el resultado de los esfuerzos de colaboración de más de setecientas personas que se auxiliaron con sus computadoras personales. El tiempo que utilizó Armengaud para probar que era un número primo, fue de ochenta y ocho horas, usando una computadora Pentium 90 MHz. Así como sucedió con el vigésimo cuarto primo de Mersenne, este nuevo primo, podría no ser el trigésimo quinto debido a que otras potencias de dos menos uno, con exponentes menores tenían que ser evaluados.

36

El 24 de agosto de 1997, Gordon Spence de Hampshire, Inglaterra; utilizando el programa de George Woltman, descubrió el primo más grande conocido ($2^{2976221} - 1$), con 895,932 dígitos. Cabe mencionar que este número llenaría un libro de cuatrocientas páginas. El hallazgo fue el resultado del proyecto GIMPS. A Spence le tomó quince días para probar que el número era primo en una computadora Pentium de 100 MHz. Este número primo de Mersenne fue verificado de manera independiente por David Slowinski en una supercomputadora Cray T90 el 29 de agosto. Este número primo de Mersenne dio lugar al trigésimo sexto número perfecto que consta de 1,791,864 dígitos.



(a) Spence



(b) Clarkson

Figura 9

37

El 27 de enero de 1998, Ronald Clarkson un estudiante de diecinueve años de la Universidad Dominguez Hills del estado de California, encontró el trigésimo séptimo primo de Mersenne ($2^{3021377} - 1$). Dos años atrás, el fundador del GIMPS, George Woltman escribió un programa muy eficiente para encontrar primos de Mersenne y estableció una base de datos para facilitar la búsqueda, la cual combina el trabajo de muchos investigadores. Scott Kurowski y otros desarrollaron el proyecto PrimeNet, que permite que el programa de Woltman se comuniquen directamente con la base de datos sin la intervención humana. Para entonces, cerca de 4,000 personas usaron este programa consumiendo aproximadamente, el equivalente a una década con una computadora de escritorio cada día y aunque el nuevo primo parecía virtualmente asegurado, nadie podía saber cuando sería encontrado. Se dice «virtualmente asegurado» porque aunque se conjetura que hay una infinidad de primos de Mersenne, esto no ha sido probado todavía. El resultado de Clarkson se obtuvo después de varios días de cálculos en computadora, de hecho Clarkson halló el registro del primo de Mersenne en su computadora personal, cuando el servidor de PrimeNet eligió el exponente para que Clarkson lo probara, él no quiso hacerlo. Clarkson comentó «nunca me habría imaginado que dos primos de Mersenne se encontraran tan cercanos». De cualquier manera, Clarkson continuó adelante y encontró un pequeño hueco entre dos primos de Mersenne. La prueba real, tomó alrededor de cuarenta y seis días con una computadora Pentium 200 MHz trabajando la mitad del tiempo, mientras que trabajando tiempo completo, le habría tomado alrededor de una semana. El 30 de enero de 1998 David Slowinski terminó de verificar la primalidad, dando origen al trigésimo sexto número perfecto $2^{3021376}(2^{3021377} - 1)$ que consta de 1,819,050 dígitos.

38



Figura 10

El primero de junio de 1999, Nayan Hajratwala quien pertenece al GIMPS, descubrió el trigésimo octavo primo de Mersenne ($2^{6972593} - 1$), el cual contiene 2,098,960 dígitos, calificando para los \$50,000 dólares que ofrecía la Fundación Electronic Frontier (EFF). Nayan Hajratwala de Plymouth Michigan, era un empleado de Price Waterhouse Coopers. De hecho, Nayan quien utilizaba el sistema PrimeNet no sabía que él había encontrado este nuevo primo en su máquina casera, hasta que PrimeNet lo envió porque él, como la mayoría de los usuarios la dejó funcionando mientras no la utilizaba, *este tiene que ser el mejor mensaje de e-mail que he leído hasta ahora*, dijo Nayan. Hajratwala, utilizó una computadora Pentium II IBM activa de 350 MHz, que trabajando parcialmente durante 111 días verificó que este número era primo, de hecho, se pudo haber verificado en sólo tres semanas trabajando tiempo completo. Cabe mencionar que PrimeNet coordina a más de 21,500 computadoras, de las cuales 12,600 son usuarios caseros de internet y otras pertenecen a escuelas y negocios alrededor del mundo, realizando 720 billones de cálculos por segundo.

39

El trigésimo noveno primo de Mersenne fue encontrado el 14 de noviembre de 2001 por Michael Cameron, un estudiante de 20 años de Ontario Canadá, y miembro del GIMPS; este número ($2^{13466917} - 1$) contiene 4,053,946 dígitos. Cameron utilizó una PC con procesador AMD 800 MHz, que trabajando parcialmente durante 45 días logró probar que este número era primo, así mismo, se auxilió de un software generado por George Woltman y Entropía Inc., como parte de una agrupación internacional conformada por más de 205,000 computadoras interconectadas y operadas por dicha compañía. El nuevo primo de Mersenne fue verificado de manera independiente por Ernst Mayer y Paul Victor Novarese, con ayuda de una computadora Alpha workstation de 667 MHz durante tres semanas. Este descubrimiento representa el quinto número de Mersenne encontrado por el GIMPS.

40



(a) Shafer



(b) Ballester

Figura 11

El cuadragésimo primo de Mersenne fue encontrado el 17 de noviembre de 2003 por Michael Shafer, un ingeniero químico de 26 años de la Universidad de Michigan y miembro del GIMPS. Éste número primo ($2^{20996011} - 1$) consta de 6, 320, 430 dígitos. De hecho, tiene más de dos millones de dígitos que el anterior primo de Mersenne. Shafer, utilizó una PC de laboratorio de la Universidad de Michigan y un software libre de George Woltman y Scott Kurowski. Michael comenta «había terminado una reunión con mi consejero cuando observé en la computadora que había encontrado un nuevo número primo de Mersenne. Después de un baile corto en señal de victoria, llamé a mi esposa y amigos del GIMPS para comentarles la gran noticia». Shafer, utilizó una computadora Dell Pentium 4 de 2 GHz, durante 19 días para probar que el número es primo. Cabe señalar que, el proyecto PrimeNet ha jugado un papel muy importante, ya que aglutina cientos de millones de computadoras en paralelo, para crear un superordenador virtual que realiza 9 teraflops de cálculos por segundo; lo que permite a los miembros del GIMPS encontrar un primo en solamente dos años, en vez de 25,000 años trabajando sólo una PC. Este nuevo primo de Mersenne, fue verificado de manera independiente por Guillermo Ballester de Granada España, en un servidor Itanium II de 1.4 GHz durante doce días en el Centro de Prueba de HP; y por Ernts Mayer de Cupertino California, usando una Alpha workstation HP de 1 GHz durante tres semanas.

41

El cuadragésimo primero primo de Mersenne, fue encontrado el 15 de mayo de 2004 por Josh Findley del GIMPS, ($2^{24036583} - 1$) consta de 7, 235, 733 dígitos. De hecho, tiene casi un millón de dígitos más que el anterior. Éste número primo, abarca el 72 % del tamaño de los diez millones de dígitos, por el que la Fundación Electronic

Frontier ofrecía \$100,000 dólares. Findley, un consultor de la Administración Oceánica y Atmosférica Nacional de La Jolla California, describió el hallazgo: *todavía estoy sorprendido del descubrimiento. Incluso después de cinco años de ejecutar el GIMPS en mis computadoras, no esperaba encontrar un nuevo primo de Mersenne.* Findley utilizó una computadora Pentium IV de 2.4 GHz con Windows XP, durante catorce días para probar que este número es primo. El nuevo primo fue verificado de manera independiente por Tony Reix de Grenoble Francia, usando la mitad de una computadora Bull NovaScale 5000 HPC, que corría Linux en una 16 Itanium II de 1.3 GHz, durante cinco días con ayuda del programa de Guillermo Ballester y por Jeff Gilchrist de Elytra Enterprises Inc. en Ottawa Canadá, utilizando por once días una Itanium II de HP rx5670 de 1.5 GHz.

42



Figura 12: Nowak

El cuadragésimo segundo primo de Mersenne, fue encontrado el 18 de febrero de 2005 por el Dr. Martin Nowak, cirujano oculista de Michelfeld Alemania, y miembro del GIMPS, $(2^{25964951} - 1)$ consta de 7,816,230 dígitos. De hecho, tiene casi medio millón de dígitos más que el anterior. Nowak utilizó una computadora Pentium 4 de 2.4 GHz durante más de 50 días. Este número primo fue verificado de manera independiente por Tony Reix de Grenoble Francia, usando una computadora 16 Itanium de Bull NovaScale 5000 HPC, que corría el programa de Guillermo Ballester. Una segunda verificación fue realizada por Jeff Gilchrist de Elytra Enterprises Inc. en Ottawa Canadá, utilizando quince días de tiempo con 12 computadoras alfa Compaq GS160 de 1.2 GHz.

43

El cuadragésimo tercero primo de Mersenne, fue encontrado el 15 de diciembre de 2005 por el Dr. Curtis Cooper y el Dr. Steven Boone de la Universidad Central del estado de Missouri. Éste número $(2^{30402457} - 1)$ consta de 9,152,052 dígitos. De



(a) Cooper



(b) Boone

Figura 13

hecho, tiene un millón trescientos mil dígitos más que el anterior. Este nuevo primo se obtuvo en una computadora del laboratorio de Comunicación de la Universidad, después de correr el programa de George Woltman en intervalos de tiempo por cerca de 50 días. Posteriormente fue verificado de manera independiente durante cinco días por Tony Reix de Bull S.A. en Grenoble Francia; usando una computadora 16 Itanium II de 1.5 GHz y el programa de Guillermo Ballester. Una segunda verificación fue hecha por Jeff Gilchrist.

44

El cuadragésimo cuarto primo de Mersenne fue descubierto el 4 de septiembre de 2006, en el mismo edificio a pocos metros de distancia el Dr. Curtis Cooper y el Dr. Steven Boone del equipo CMSU rompió su propio récord mundial, el número descubierto ($2^{32,582,657} - 1$) está formado por 9,808,358 dígitos y tiene más de 650,000 dígitos que el número primo anterior. Sin embargo, el número no rebasa los 10 millones de dígitos que son necesarios para reclamar el premio que ofrecían al proyecto GIMPS. Cabe destacar que en menos de tres años el GIMPS ha descubierto cinco primos de Mersenne. Por otra parte Tony Reix de Bull SA en Grenoble Francia, verificó la primalidad del cuadragésimo cuarto primo de Mersenne de forma independientemente en 6 días, usando 16 computadoras Itanium2 de 1.5 GHz de la Bull NovaScale 6160 HPC, del centro de investigación Bull en Grenoble, ejecutando el programa Glucas de Guillermo Ballester de Granada España. Es importante señalar que el Dr. Cooper y el Dr. Boone, no podrían haber hecho este descubrimiento por sí solos. En reconocimiento a las contribuciones hechas por los coordinadores del proyecto y las decenas de miles de voluntarios, el GIMPS dio el crédito de este descubrimiento a «Cooper, Boone, Woltman, Kurowski et al.». El descubrimiento es el décimo número obtenido por el proyecto GIMPS.

45, 46



(a) Elvenich



(b) Kurowski

Figura 14

El 15 de septiembre de 2008, se encontró el cuadragésimo sexto y el cuadragésimo quinto primo de Mersenne, cada uno consta de 12,978,189 y 11,185,272 dígitos, como parte del proyecto GIMPS en 12 años. Los números primos son $(2^{43112609} - 1)$ y $(2^{37156667} - 1)$. El número más grande calificaba para un premio de investigación de \$100,000 dólares, la mayor parte del premio el GIMPS lo donó a la Universidad de California (UCLA) y otro porcentaje a una institución de beneficencia.

Como reconocimiento de la contribución de cada participante del GIMPS, el crédito para la primera fase de clasificación fue para «Edson Smith, George Woltman, Scott Kurowski, et al.» y el otro para «Hans-Michael Elvenich, George Woltman, Scott Kurowski, et al.» Un concurso de casi una década de duración, para un premio de \$100,000 dólares de la Fundación Electronic Frontier, terminó cuando el primo mayor surgió en un equipo de la UCLA dirigido por Edson Smith, apenas dos semanas antes de que el segundo primo fuera encontrado por el equipo de Hans-Michael Elvenich, en Langenfeld Alemania, cerca de Colonia. Ambos primos estaban entre los 100,000 ordenadores del GIMPS, el cual ha estado funcionando continuamente desde 1996 y realizando alrededor de 29 billones de cálculos por segundo. «Estamos orgullosos de ser participantes en el GIMPS y agradecidos con el Departamento de Matemáticas de la UCLA, para el suministro de recursos computacionales para el proyecto», dijo Edson Smith, Gerente de Informática. Por su parte, Hans-Michael Elvenich un ingeniero eléctrico alemán y un entusiasta de los números primos, agregaba: «Después de cuatro años de búsqueda en el GIMPS, obtuvimos un gran éxito». «El desarrollo de las tecnologías y los métodos para aplicar el increíblemente vasto poder de la computación en la investigación cooperativa, es la razón por la que la Fundación Electronic Frontier creó los premios. Es una investigación seria, pero es divertido y educativo también».

El Fundador del GIMPS George Woltman en Orlando Florida, dijo: «Además de felicitar y agradecer las enormes contribuciones de nuestros cientos de miles de

participantes a lo largo de los años, nos hemos comprometido a dar \$25,000 dólares para la caridad; \$50,000 dólares a la UCLA por su parte en el descubrimiento» y la mayor parte de lo que queda para otros buscadores de primos en el GIMPS y agregaba: «Nuestro proyecto de investigación pronto ofrecerá la oportunidad de lograr el próximo reto, el premio de \$150,000 dólares para un primo de 100 millones de dígitos, lo que parece más difícil. Todo lo que necesitas para participar, es nuestro programa de descarga gratuita y mucha paciencia».

En la siguiente tabla, presentamos los números primos de Mersenne que se conocen y la cantidad de dígitos que tienen, así como la cantidad de dígitos que forman los números perfectos pares, inducidos por los números primos de Mersenne.

No.	p	Dígitos en M_p	Dígitos en P_p	Año	Autor
1	2	1	1
2	3	1	2
3	5	2	3
4	7	3	4
5	13	4	8	1456	Anónimo
6	17	6	10	1588	Cataldi
7	19	6	12	1588	Cataldi
8	31	10	19	1772	Euler
9	61	19	37	1883	Pervushin
10	89	27	54	1911	Powers
11	107	33	65	1914	Powers
12	127	39	77	1876	Lucas
13	521	157	314	1952	Robinson
14	607	183	366	1952	Robinson
15	1279	386	770	1952	Robinson
16	2203	664	1327	1952	Robinson
17	2281	687	1373	1952	Robinson
18	3217	969	1937	1957	Riesel
19	4253	1281	2561	1961	Hurwitz
20	4423	1332	2663	1961	Hurwitz
21	9689	2917	5834	1963	Gillies

No.	p	Dígitos en M_p	Dígitos en P_p	Año	Autor
22	9941	2993	5985	1963	Gillies
23	11213	3376	6751	1963	Gillies
24	19937	6002	12003	1971	Tuckerman
25	21701	6533	13066	1978	Noll y Nickel
26	23209	6987	13973	1979	Noll
27	44497	13395	26790	1979	Nelson y Slowinski
28	86243	25962	51924	1982	Slowinski
29	110503	33265	66530	1988	Colquitt y Welsh
30	132049	39751	79502	1983	Slowinski
31	216091	65050	130100	1985	Slowinski
32	756839	227832	455663	1992	Slowinski y Gage
33	859433	258716	517430	1994	Slowinski y Gage
34	1257787	378632	757263	1996	Slowinski y Gage
35	1398269	420921	841842	1996	Armengaud, Woltman et al. (GIMPS)
36	2976221	895932	1791864	1997	Spence, Woltman et al.
37	3021377	909526	1819050	1998	Clarkson, Woltman, Kurowski et al.
38	6972593	2098960	4197919	1999	Hajratwala, Woltman, Kurowski et al.
39	13466917	4053496	8107892	2001	Cameron, Woltman, Kurowski et al.
40	20996011	6320430	12640858	2003	Shafer, Woltman, Kurowski et al.
41	24036583	7235733	14471465	2004	Findley, Woltman, Kurowski et al.
42	25964951	7816230	15632458	2005	Nowak, Woltman, Kurowski et al.
43	30402457	9152052	18304103	2005	Cooper, Boone, Woltman, Kurowski et al.

No.	p	Dígitos en M_p	Dígitos en P_p	Año	Autor
44	32582657	9808358	19616714	2006	Cooper, Boone, Woltman, Kurowski et al.
45	13466917	4053496	8107892	2008	Elvenich, Woltman, Kurowski et al.
46	42643801	12837064	25674127	2009	Strindmo, Woltman, Kurowski et al.
47	43112609	12978189	25956377	2008	Smith, Woltman, Kurowski et al.
48	57885161	17425170	34850339	2013	Cooper, Woltman, Kurowski et al.

47

El 12 de abril de 2009 fue encontrado el 47 número primo de Mersenne a saber $(2^{42643801} - 1)$, este número consta de 12, 837, 064 dígitos y fue encontrado por Odd Magnar Strindmo de Melhus Noruega. Este es el penúltimo primo de Mersenne conocido y tiene 141,125 dígitos menos que el primo de Mersenne encontrado en agosto pasado. Odd es un profesional de TI cuyos equipos han estado trabajando en el GIMPS. El cálculo le tomó 29 días con un procesador de 3 GHz Intel Core 2. Éste número primo, también conocido como $M_{42643801}$, fue encontrado el 12 de abril, pero debido a un fallo del servidor pasó desapercibido hasta el 4 de junio. También este número, fue verificado en forma independiente por Tony Reix el 12 de junio con el programa Glucas, que se ejecuta en los servidores HPC de Bull NovaScale, uno con una computadora Itanium2 y otro con una computadora Nehalem EX. Odd es voluntario del proyecto GIMPS, de hecho, el GIMPS ha encontrado los últimos 13 números primos de Mersenne, de los cuales 11 fueron récords mundiales. En reconocimiento a la contribución de cada participante del GIMPS, el crédito para este número irá a «O. Strindmo, G. Woltman, S. Kurowski, et al.»

48

El cuadragésimo octavo primo de Mersenne fue descubierto el 25 de enero de 2013 a las 23:30:26, el mayor número primo conocido es $(2^{57885161} - 1)$ y fue descubierto por el Dr. Curtis Cooper un voluntario del GIMPS. El número primo tiene 17, 425, 170 dígitos y se utilizaron 360, 000 computadoras en horas pico realizando 150 billones de

cálculos por segundo, durante 17 años corriendo de manera continua en el GIMPS. Este es el tercer registro para el Dr. Cooper y su Universidad. La UCLA tenía el récord hasta que la Universidad Central de Missouri, recuperó el récord mundial con este descubrimiento. La nueva prueba de primalidad, tomó 39 días de cálculos computacionales sin parar en uno de los ordenadores de la universidad. El Dr. Cooper y su Universidad son los mayores contribuyentes individuales al proyecto. El descubrimiento fue elegible para un premio de investigación del GIMPS de \$3,000 dólares. Es el último primo de Mersenne descubierto hasta ahora y cada vez parece más difícil de encontrar.

Chris Caldwell tiene un sitio web autorizado <https://primes.utm.edu/> sobre los números primos conocidos. Para demostrar que no había errores en el proceso de descubrimiento, el primo fue verificado de forma independiente con diferentes programas que se ejecutan en un hardware diferente, Serge Batalov de la fundación Ernst Mayer's, corrió el algoritmo de Lucas-Lehmer en un servidor de 32 núcleos donado por Novartis en 6 días y Jerry Hallett también lo verificó en 3.6 días. Por último, el Dr. Jeff Gilchrist verificó el hallazgo mediante el software del GIMPS en una computadora Intel i7 en 4.5 días y con el programa CUDALucas en una NVidia GTX 560 Ti en 7.7 días.

Cabe señalar, que el crédito para los descubrimientos del GIMPS no va solo al Dr. Cooper por ejecutar el software en los ordenadores de su Universidad, sino también para Woltman y Kurowski por la creación del software y ejecutar el proyecto, además para los miles de voluntarios del GIMPS. Por lo tanto, el crédito oficial para este descubrimiento se otorgó a «C. Cooper, G. Woltman, S. Kurowski, et al.»

La historia sobre la búsqueda de números primos de Mersenne y números perfectos continúa y nada despreciables \$150,000 dólares les están esperando...

Ahora nos podemos preguntar ¿existen una infinidad de primos de Mersenne?, se cree que si pero nadie ha podido demostrarlo. Una consecuencia natural es que si hay una infinidad de primos de Mersenne, entonces por el teorema de Euclides, habría una infinidad de números perfectos pares.

Hay otras preguntas abiertas que surgen en torno a los números primos, como por ejemplo:

- 1) ¿Existen infinitos primos de la forma $n^2 + 1$?
- 2) ¿Existe un número primo entre n^2 y $(n + 1)^2$? Cabe señalar que Chebychev probó el llamado postulado de Bertrand: «Para todo número natural $n \geq 2$, siempre existe un número primo p mayor que n y menor que $2n$ ».

Agradecimientos

Agradezco al árbitro sus valiosos comentarios y sugerencias para el enriquecimiento de este trabajo.

Bibliografía

- [1] P. Ribenboim, *The Book of Prime Numbers Records*, Springer-Verlag, New York, 1989.
- [2] R. E. Crandall, *Projects in Scientific Computation*, Springer-Verlag, Santa Clara, CA; 1989.
- [3] K. Wilson, *The Plays of Hrotsvit of Gandersheim*, Garland Publishing, New York, 1989.
- [4] S.J. Bezuska and M. J. Kenney, *Even perfect numbers: Update*, The mathematics Teacher, V.90, N.8; 1997.
- [5] P.Hoffman, *El hombre que sólo amaba a los números*, Granica, 2000.
- [6] T. Koshy, *Elementary number theory with applications*, Harcourt Academic Press, 2002.
- [7] www.mersenne.org

Universidad Autónoma de la Ciudad de México, UACM
Fray Servando Teresa de Mier 92 y 99, Colonia Centro,
México, D. F. C.P. 06080.

Facultad de Ciencias, UNAM
Circuito Exterior S/N, Cd. Universitaria,
México, D. F. C.P. 04510.
amersen@yahoo.com.mx

Geometría

Capítulo 5

La magia de los cuadriláteros cíclicos

Aarón Aparicio Hernández

Universidad Autónoma de la Ciudad de México;
Facultad de Ciencias, UNAM.

Resumen

Es bien conocido que cualesquiera tres puntos pertenecen a una misma circunferencia, sin embargo, para cualesquiera cuatro puntos eso ya no ocurre. En el presente artículo analizamos algunos resultados que determinan cuándo un cuadrilátero es cíclico, para ello, utilizamos conceptos básicos tales como congruencia y semejanza de triángulos; el ángulo central y el teorema de Ptolomeo.

*A don Alberto Barajas en el centenario,
de su nacimiento.*



Alberto Barajas Celis nació el 17 de julio de 1913 en la Ciudad de México y murió el 3 de julio de 2004. Realizó diversos trabajos en matemáticas teóricas junto a colegas y compañeros suyos como Carlos Graef Fernández, tales como *Invariantes proyectivos de las transformaciones circulares* (1942). Obtuvo la beca Guggenheim por un año que lo llevó a estudiar en la Universidad de Harvard (1944-1945) con el matemático George David Birkhoff hasta antes de su muerte, estudio que le serviría para conseguir su doctorado en ciencias matemáticas con la tesis *Teoría de la gravitación* (1944) basado en la Teoría relativista de la gravitación realizado por los científicos rusos Logunov, Mestvirishvili y Petrov, trabajo que aparece cita-

do en la Enciclopedia Británica, por lo que fue invitado a dar conferencias en las universidades de Brown, Harvard y Princeton, en donde tuvo la oportunidad de discutir su trabajo con Albert Einstein, ya publicado por el propio Barajas algún tiempo antes y asesorado por Birkhoff. Desempeñó los siguientes puestos a partir

de 1934: investigador de carrera del Instituto de Matemáticas; profesor, director y coordinador en la Facultad de Ciencias (Universidad Nacional Autónoma de México); miembro de la Comisión de Nuevos Métodos de Enseñanza y miembro de la Junta de Gobierno 1970-1979. Desde 1976 fue profesor emérito de la Universidad Nacional Autónoma de México, y en 1985 le fue otorgado el grado de Doctor honoris causa en Matemáticas por la Universidad Nacional Autónoma de México. El doctor Alberto Barajas fue un entusiasta impulsor para que se construyera Ciudad Universitaria en la década de 1950. Asimismo, fue coordinador de la Investigación Científica de la UNAM y presidente del Consejo Consultivo de la CNEN (Comisión Nacional de Energía Nuclear (México)), entre otros cargos.

Con respecto a la enseñanza, Alberto Barajas impartió por casi 70 años varios cursos de matemáticas en la Facultad de Ciencias, se dice fácil pero muy pocos tienen una carrera tan larga, brillante y lúcida hasta el final. Barajas fue un profesor fuera de serie, tenía una extraordinaria habilidad como expositor y orador, que desplegaba clase por clase, era muy cuidadoso en sus exposiciones, trazaba círculos perfectos, manejaba los tiempos y las pausas a la perfección y acompañaba sus clases con anécdotas históricas que contaba como si hubiera sido amigo de todos los grandes creadores de las matemáticas. En los últimos años don Alberto tuvo la costumbre de alternar sus cursos: impartía un año las materias de Geometría Moderna I y II, y otro año las materias de Teoría de los números I y II. Fue en el año de 1987 cuando tuve el privilegio de asistir como estudiante a sus cursos y los cuadriláteros cíclicos fue uno de los temas que más me impresionaron cuando lo expuso con gran elegancia en la clase.

*Les recuerdo que matemático no es el nombre de un talento,
sino de una pasión.
A. Barajas*

1 Algunas curiosidades de los cuadriláteros cíclicos

Es muy conocido que cualesquiera tres puntos pertenecen a una misma circunferencia, sin embargo, para cualesquiera cuatro puntos eso ya no ocurre. Por consiguiente, decimos que un cuadrilátero es *cíclico* si existe una circunferencia que pasa por los vértices del cuadrilátero.

El segmento que va del punto A al punto B se denotará con \overline{AB} , el segmento \overline{AB} es paralelo al segmento \overline{CD} con $\overline{AB} \parallel \overline{CD}$, el segmento \overline{AB} es perpendicular al segmento \overline{CD} con $\overline{AB} \perp \overline{CD}$, el arco que va del punto A al punto B con \widehat{AB} ; el triángulo que pasa por los puntos A, B, C con $\triangle ABC$, el cuadrilátero que pasa por los puntos A, B, C, D con $\square ABCD$, al ángulo formado por los puntos A, B, C

con $\angle ABC$. Los triángulos ABC y DEF son congruentes con $\triangle ABC \cong \triangle DEF$, los triángulos ABC y DEF son semejantes con $\triangle ABC \approx \triangle DEF$.

Empecemos con el teorema del ángulo central, es un resultado básico con consecuencias importantes, como podrá verse a lo largo del capítulo.

Teorema 1.1 (del ángulo central). *Si el ángulo β es un ángulo inscrito en una circunferencia \mathcal{C} y el ángulo α tiene vértice en el centro y abarca el mismo arco (ver figura 1), entonces $\beta = \frac{1}{2}\alpha$.*

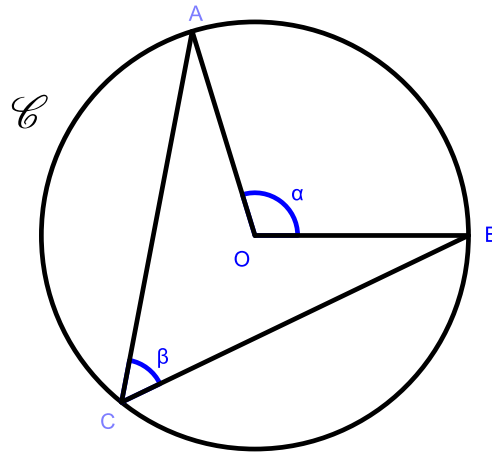


Figura 1

Demostración. Como veremos, los tres casos ilustrados en la figura 2 se reducen al primero.

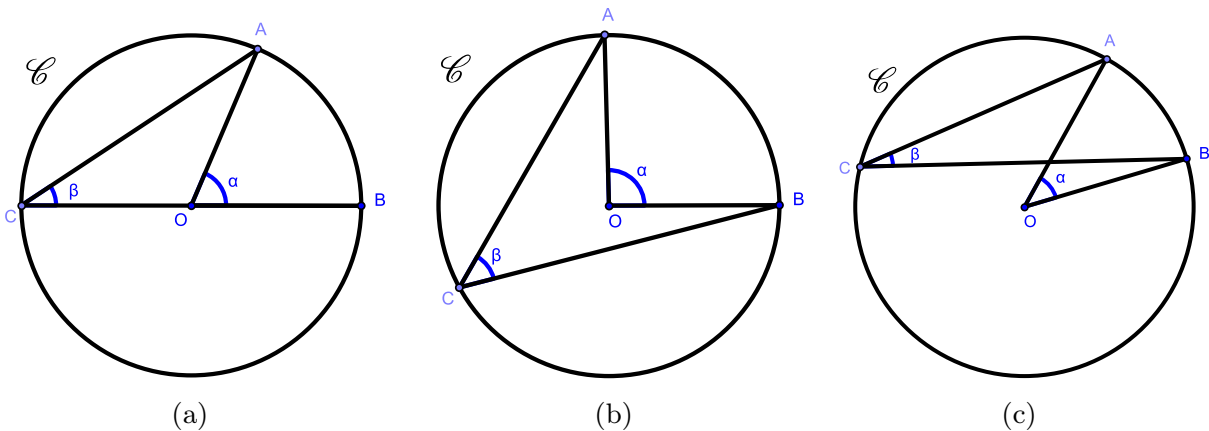


Figura 2

i) Cuando \overline{CB} es un diámetro (ver figura 2(a)).

El $\triangle OAC$ es isósceles pues $\overline{OA} = \overline{OC}$ son radios, así $\beta = \angle OAC$. Además $\beta + \angle OAC + \angle COA = \beta + \beta + \angle COA = 180^\circ$ y $\angle COA + \alpha = 180^\circ$, luego $2\beta = \alpha$. Por lo tanto $\beta = \frac{1}{2}\alpha$.

ii) Consideremos el caso de la figura 2(b).

Este caso lo reducimos al caso anterior, para ello trazamos un diámetro \overline{CD} que pase por C y O (ver figura 3(a)). Por el caso i) sabemos que $\angle AOD = 2\angle ACO$ y

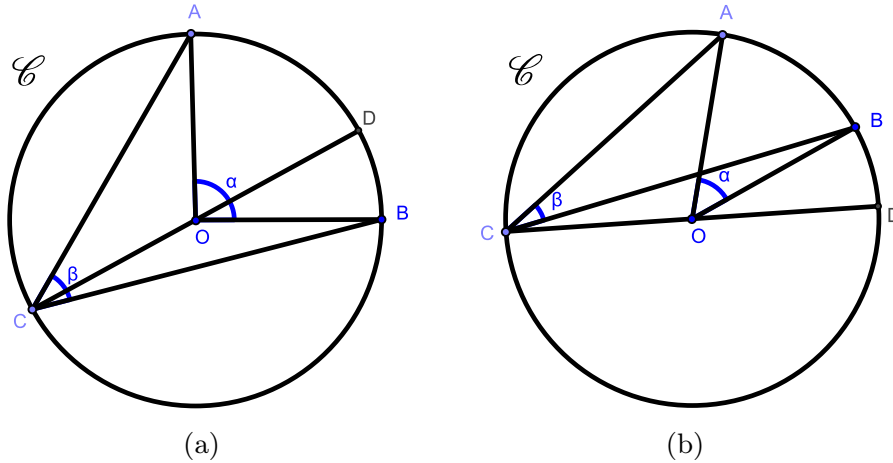


Figura 3

$\angle DOB = 2\angle OCB$, por lo tanto

$$\begin{aligned}\alpha &= \angle AOD + \angle DOB = 2\angle ACO + 2\angle OCB \\ &= 2(\angle ACO + \angle OCB) = 2\beta,\end{aligned}$$

así $\beta = \frac{1}{2}\alpha$.

iii) Ahora consideremos el caso de la figura 2(c).

Primero trazamos un diámetro \overline{CD} que pase por C y O (ver figura 3(b)). Aplicando el caso i) tenemos que $\alpha + \angle BOD = 2(\beta + \angle BCO) = 2\beta + 2\angle BCO$, pero $\angle BOD = 2\angle BCO$. Por consiguiente $\alpha = 2\beta$, luego $\beta = \frac{1}{2}\alpha$. ■

Teorema 1.2 (del cuadrilátero cíclico). Sean E y D puntos en el arco \widehat{AC} y B en el arco complementario (ver figura 4(a)), entonces

- (1) $\square ACDE$ es cíclico si y solo si $\angle CEA = \angle CDA$,
- (2) $\square ABCD$ es cíclico si y solo si $\angle CDA + \angle ABC = 180^\circ$.

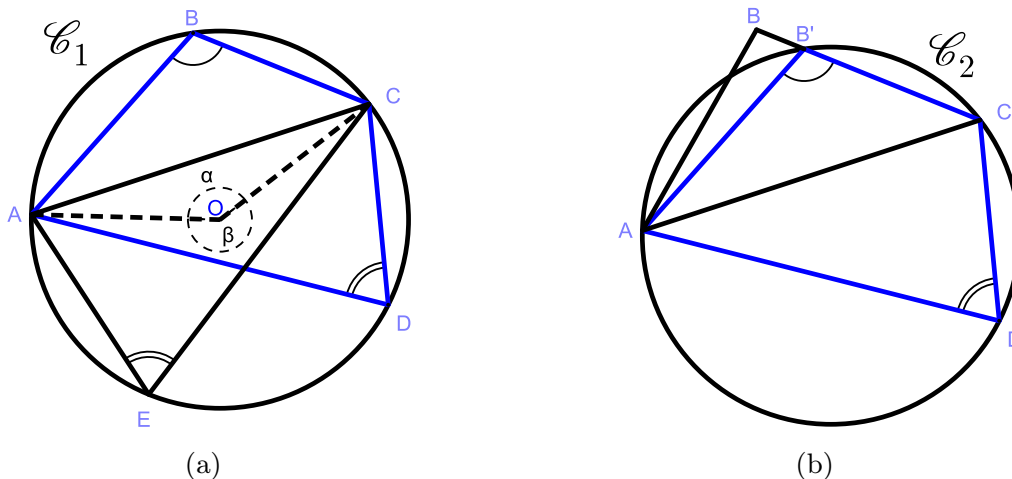


Figura 4

Demostración.

(1) $\square ACDE$ es cíclico si y solo si $\angle CEA = \angle CDA$, pues $\angle CEA = \frac{1}{2}\alpha$ y $\angle CDA = \frac{1}{2}\alpha$ por el teorema del ángulo central.

(2) Supongamos que \mathcal{C}_1 es un círculo con centro en O (ver figura 4(a)). Si $\square ABCD$ es cíclico, entonces $\angle CDA = \frac{1}{2}\alpha$ y $\angle ABC = \frac{1}{2}\beta$ por el teorema del ángulo central, pero $\alpha + \beta = 360^\circ$ y por lo tanto $\angle CDA + \angle ABC = \frac{1}{2}(\alpha + \beta) = 180^\circ$.

Inversamente, supongamos que $\angle CDA + \angle ABC = 180^\circ$ y \mathcal{C}_2 el círculo que pasa por A, D y C (ver figura 4(b)), basta probar que B pertenece a \mathcal{C}_2 . Supongamos lo contrario, entonces $\overline{CB} \cap \mathcal{C}_2 = B'$ y por consiguiente A, B', C y D son concíclicos, por lo tanto $\angle CDA + \angle AB'C = 180^\circ$, así $\angle ABC = \angle AB'C = 180^\circ - \angle CDA$ lo cual es imposible, pues $\angle AB'C > \angle ABC$. Por lo tanto, $B = B'$; es decir, $\square ABCD$ es cíclico. ■

Teorema 1.3. *Si el cuadrilátero $CDEF$ es cíclico y que no sea un cuadrado (ver figura 5(a)) y se trazan las bisectrices interiores, entonces se obtiene un cuadrilátero $GHIJ$ que es cíclico.*

Demostración. En la figura 5(b) $\angle DGC = \theta$ y $\angle FIE = \rho$ porque son opuestos por el vértice, por otra parte $2\alpha + 2\beta + 2\eta + 2\varphi = 360^\circ$, luego $\alpha + \beta + \eta + \varphi = 180^\circ$; además $\alpha + \beta + \theta = 180^\circ$ y $\rho + \theta + \varphi = 180^\circ$, por lo tanto $\theta + \rho = (180^\circ - \alpha - \beta) + (180^\circ - \eta - \varphi) = 360^\circ - (\alpha + \beta + \eta + \varphi) = 180^\circ$, así $\theta + \rho = 180^\circ$ y por consiguiente el cuadrilátero $GHIJ$ cíclico. ■

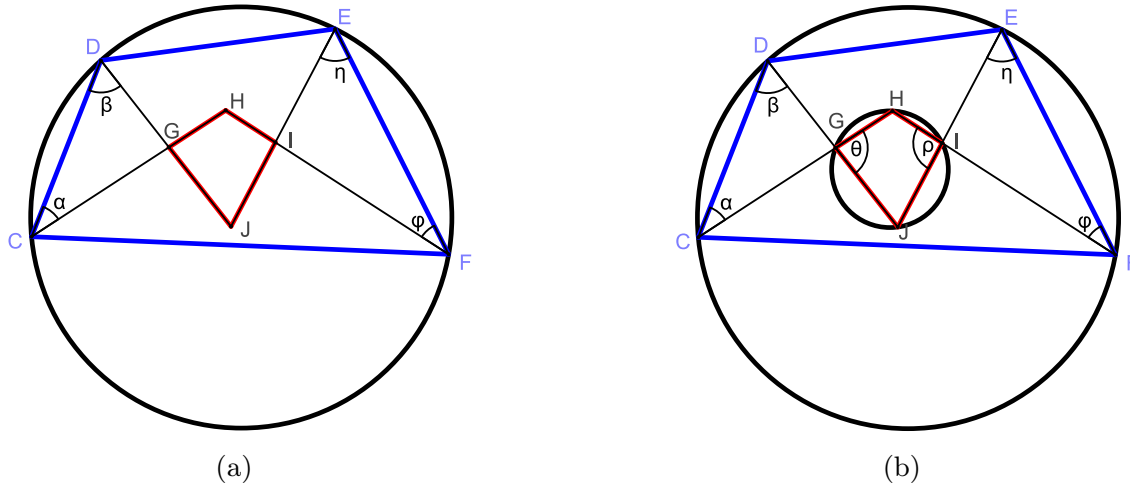


Figura 5

Observación 1.4. El teorema anterior se cumple para todos los cuadriláteros convexos no necesariamente cíclicos, la misma demostración sirve para ello. El siguiente teorema lo cumple cualquier cuadrilátero, en particular los cuadriláteros que son cíclicos.

Teorema 1.5. Si L, M, N, R son los puntos medios de los lados del $\square ABCD$ (ver figura 6), entonces el $\square LMNR$ es un paralelogramo.

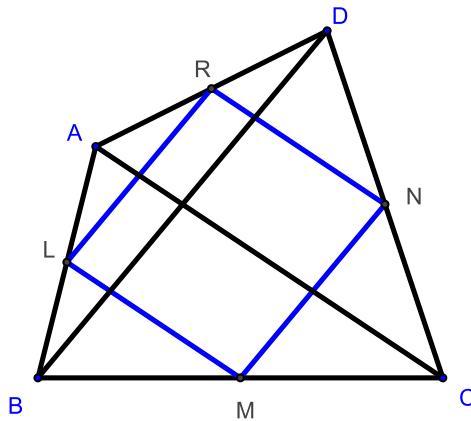


Figura 6

Demostración. En los triángulos BCD y BDA los segmentos \overline{MN} y \overline{LR} son paralelos a \overline{BD} y $\overline{MN} = \frac{1}{2}\overline{BD} = \overline{RL}$. Análogamente los segmentos \overline{LM} y \overline{NR} son paralelos a \overline{AC} y $\overline{LM} = \frac{1}{2}\overline{AC} = \overline{NR}$, por lo tanto $\square LMNR$ es un paralelogramo. ■

Definición 1.6. Si $\square ABCD$ es cíclico con L, M, N y R los puntos medios de los segmentos $\overline{AB}, \overline{BC}, \overline{CD}$ y \overline{DA} respectivamente, entonces \overline{LN} y \overline{MR} se llaman *bimedias* y $\overline{LN} \cap \overline{MR} = P$ se llama *centroide* del $\square ABCD$ (ver figura 7).

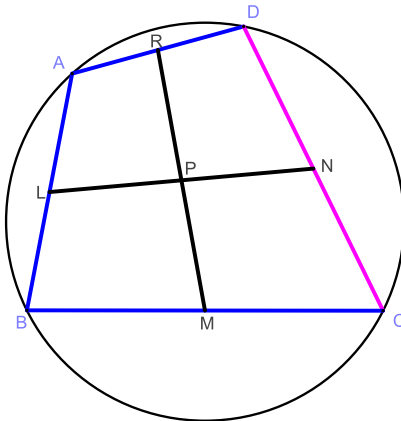


Figura 7

Un concepto interesante surge cuando reflejamos el circuncentro de un cuadrilátero cíclico respecto de su centroide.

Definición 1.7. Si reflejamos el circuncentro O del $\square ABCD$ cíclico con respecto a P , se obtiene el punto O' y se le llama *anticentro* del $\square ABCD$ (ver figura 8).

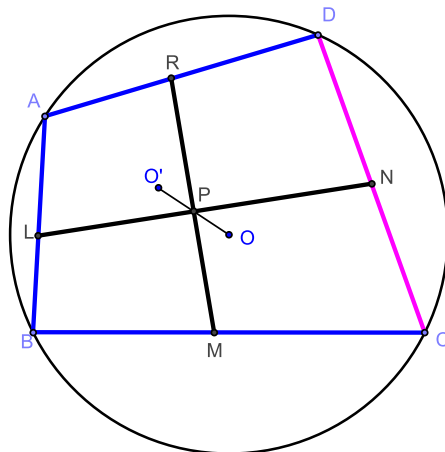


Figura 8

Cabe señalar que el anticentro aparece en varios resultados de geometría euclidiana [2], [6]; sin embargo es un concepto que no aparece antes del siglo XIX, es a partir del siglo XX cuando empieza a utilizarse como tal. El siguiente lema aporta una propiedad interesante del anticentro.

Lema 1.8. Si O' es el anticentro del $\square ABCD$ con E, F, G, H los puntos medios de los lados del $\square ABCD$ y al prolongar $\overline{EO'}$ hasta que corte a \overline{DC} en R (ver figura 9), entonces la recta que pasa por E, O' y R es perpendicular a \overline{CD} .

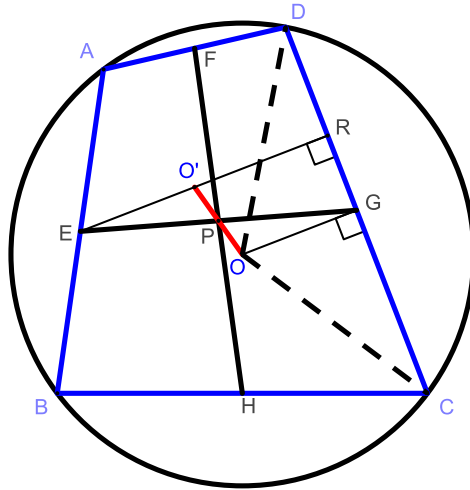


Figura 9

Demostración. Por el teorema anterior el $\square EFGH$ es un paralelogramo y como las diagonales se bisecan, entonces $\overline{EP} = \overline{PG}$, además $\angle EPO' = \angle GPO$ por ser ángulos opuestos por el vértice y $\overline{O'P} = \overline{PO}$ pues O' es el anticentro del $\square ABCD$; así $\triangle PO'E \cong \triangle POG$, luego $\overline{EO'} = \overline{OG}$.

Por otra parte $\triangle ODC$ es isósceles, porque O es el circuncentro y \overline{OG} es mediatriz del $\triangle ODC$, por lo tanto \overline{ER} es perpendicular a \overline{CD} . Además $\square EO'GO$ es un paralelogramo. ■

Observación 1.9. Por el resultado anterior, la recta que pasa por el anticentro y el punto E es perpendicular a \overline{CD} . Análogamente, la recta que pasa por el anticentro y el punto H es perpendicular a \overline{AD} . Por lo tanto, existen cuatro rectas que pasan por los puntos medios del $\square ABCD$ y que son perpendiculares a los lados opuestos y además son concurrentes en el anticentro O' (ver figura 10). Si denotamos con I, J, K, R a los puntos de intersección de estas cuatro rectas con los lados del $\square ABCD$, entonces $\overline{ER}, \overline{KF}, \overline{GJ}, \overline{IH}$ se llaman *maltitudes* del $\square ABCD$, el nombre se debe a que cada segmento es perpendicular y pasa por el punto medio de cada lado del cuadrilátero, este concepto se empezó a utilizar a mediados del siglo XIX. Por consiguiente las cuatro maltitudes del $\square ABCD$ son concurrentes en el anticentro.

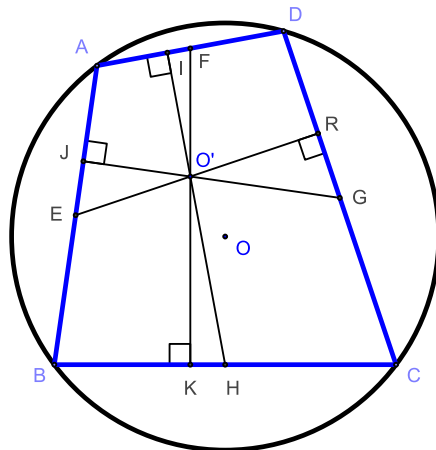


Figura 10

Teorema 1.10 (Brahmagupta). *Si el cuadrilátero ABCD es cíclico con diagonales perpendiculares entre sí, entonces la perpendicular a un lado desde el punto de intersección de las diagonales biseca el lado opuesto (ver figura 11(a)).*

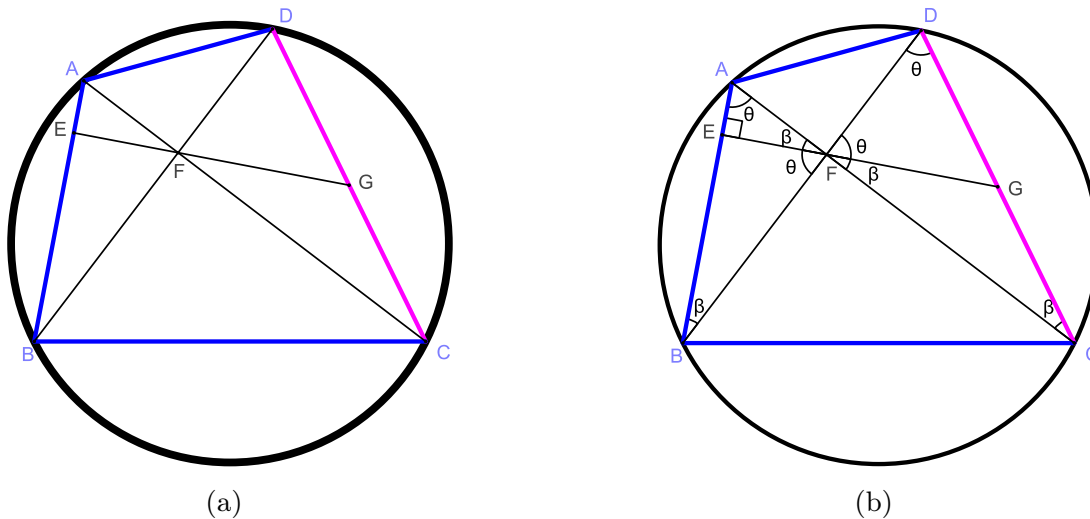


Figura 11

Demostración. Sabemos que $\angle CAB = \angle CDB$ porque abarcan el mismo arco \widehat{BC} (ver figura 11(b)), análogamente $\angle ABD = \angle ACD$ porque abarcan el mismo arco \widehat{AD} , y como además los $\triangle BFA$ y $\triangle EFB$ son rectángulos, entonces el $\triangle DFG$ resulta isósceles y de forma análoga $\triangle CFG$ también es isósceles, luego $\overline{DG} = \overline{FG}$

y $\overline{FG} = \overline{GC}$, por consiguiente $\overline{DG} = \overline{GC}$. Además F es el anticentro del $\square ABCD$ por el lema 1.8. ■

Corolario 1.11. *Si el cuadrilátero $ABCD$ es cíclico con diagonales perpendiculares entre sí, entonces la distancia del circuncentro O al lado \overline{AB} es igual a la mitad del lado opuesto, esto es, $\overline{OE} = \frac{1}{2} \overline{CD}$ (ver figura 12).*

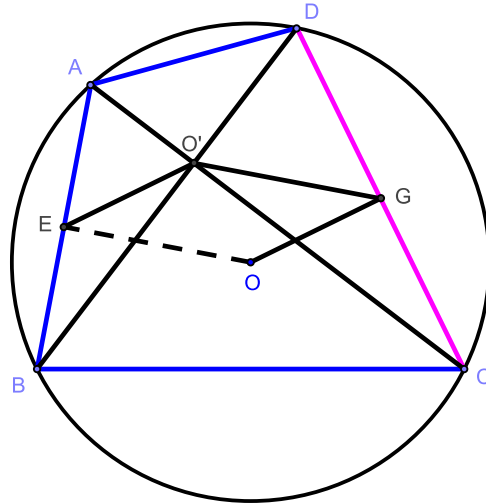


Figura 12

Demostración. Como O es el circuncentro del $\square ABCD$, entonces el $\triangle AOB$ es isósceles, pues $\overline{OA} = \overline{OB}$. Así \overline{EO} es mediatriz del lado \overline{AB} . Además, por el teorema anterior la línea que pasa por el punto medio G de \overline{CD} y el anticentro O' del $\square ABCD$ es perpendicular a \overline{AB} , por lo tanto $\overline{O'G} \parallel \overline{OE}$. Análogamente $\overline{O'E} \parallel \overline{OG}$ y por consiguiente $\square EOGO'$ es un paralelogramo, así $\overline{EO} = \overline{GO'}$; pero G es punto medio de \overline{CD} , por lo tanto G es circuncentro del $\triangle DO'C$ el cual es un triángulo rectángulo, ya que las diagonales del $\square ABCD$ son perpendiculares, luego $\overline{EO} = \overline{GO'} = \frac{1}{2} \overline{CD}$. ■

Un resultado importante y bello que caracteriza a los cuadriláteros cíclicos, es el siguiente teorema dado por el geómetra griego Claudio Ptolomeo (150 D.C.) y redescubierto por el prominente matemático de la India Brahmagupta (600 D.C.).

Teorema 1.12 (Ptolomeo). *Si el cuadrilátero $ABCD$ es cíclico (ver figura 13(a)), entonces la suma de los productos de los lados opuestos es igual al producto de las diagonales, esto es, $\overline{AB} \overline{CD} + \overline{BC} \overline{AD} = \overline{AC} \overline{BD}$.*

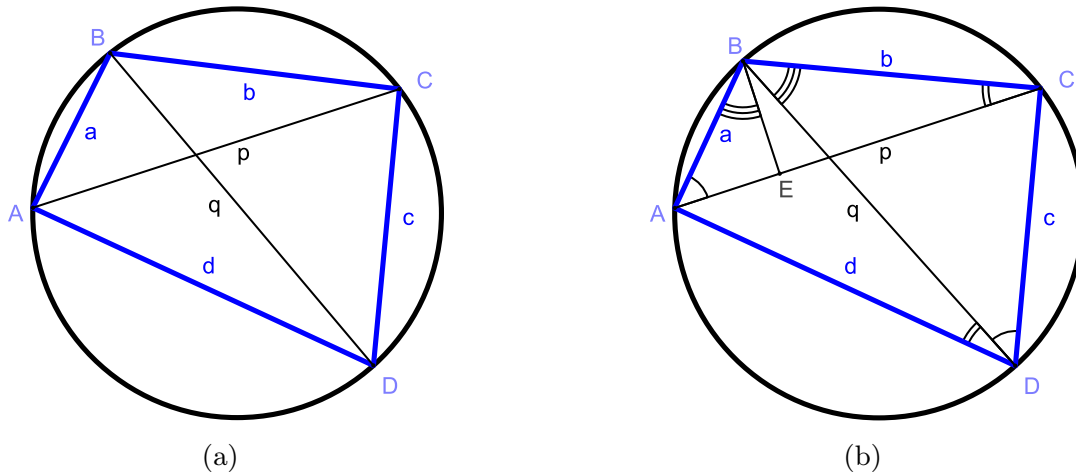


Figura 13

Demostración. En la figura 13(b) trazamos $\angle ABE = \angle DBC$, con E un punto en \overline{AC} ; $\angle BAE = \angle BDC$ por el teorema del ángulo central. Análogamente $\angle ADB = \angle ECB$, además $\angle DBA = \angle CBE$. Por lo tanto,

$$\triangle ABE \approx \triangle DBC \quad \text{y} \quad \triangle ADB \approx \triangle ECB,$$

luego

$$\frac{a}{q} = \frac{\overline{AE}}{c} \quad \text{y} \quad \frac{q}{b} = \frac{d}{\overline{EC}};$$

por consiguiente,

$$ac = q \overline{AE} \quad \text{y} \quad bd = q \overline{EC},$$

por lo tanto $ac + bd = q \overline{AE} + q \overline{EC} = q(\overline{AE} + \overline{EC}) = qp$. ■

Observación 1.13. No es muy complicado ver que el recíproco del teorema de Ptolomeo también se cumple.

Corolario 1.14. *Identidad trigonométrica*

$$\text{sen}(\alpha + \beta) = \text{sen } \alpha \cos \beta + \text{sen } \beta \cos \alpha.$$

Demostración. Consideremos el cuadrilátero cíclico ABCD, con diámetro $\overline{AC} = 1$ (ver figura 14(a)); por consiguiente $\text{sen} \alpha = \overline{BC}$, $\text{sen} \beta = \overline{CD}$, $\text{cos} \alpha = \overline{AB}$ y $\text{cos} \beta = \overline{AD}$. Ahora apliquemos el teorema de Ptolomeo al cuadrilátero cíclico ABCD, esto es, $\overline{AB} \overline{CD} + \overline{BC} \overline{AD} = \overline{AC} \overline{BD}$, por lo tanto $\text{cos } \alpha \text{ sen} \beta + \text{sen } \alpha \text{ cos} \beta = \overline{BD}$; veamos que $\text{sen}(\alpha + \beta) = \overline{BD}$, para ello trazamos el diámetro

\overline{BE} como se observa en la figura 14(b), así $\text{sen } \angle BED = \overline{BD}$, luego $\angle BED = \alpha + \beta$. Por lo tanto,

$$\text{sen}(\alpha + \beta) = \text{sen } \alpha \cos \beta + \text{sen } \beta \cos \alpha. \quad \blacksquare$$

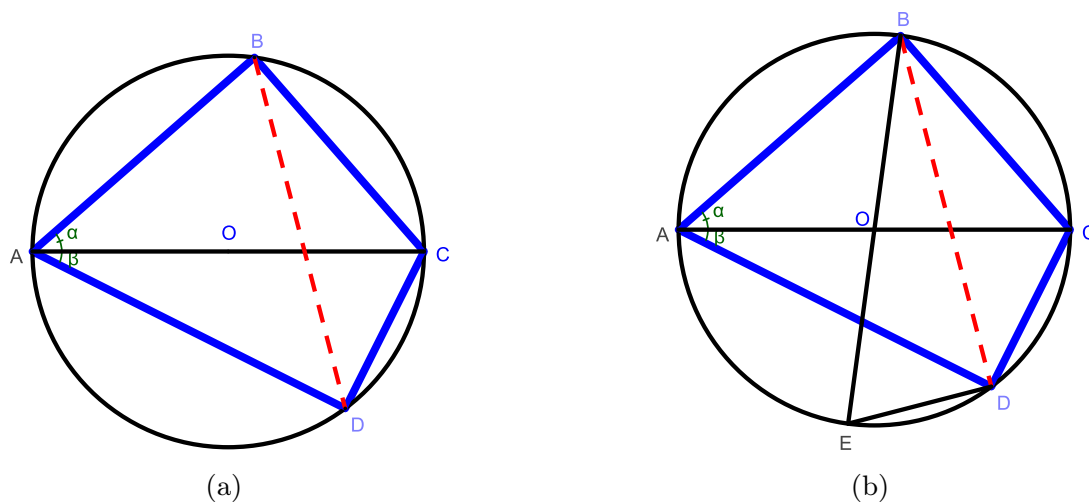


Figura 14

Teorema 1.15 (Brahmagupta). *Si el triángulo ABC es isósceles y P pertenece al arco \widehat{BC} (ver figura 15), entonces*

$$\frac{\overline{AP}}{\overline{CP} + \overline{PB}} = \frac{\overline{AC}}{\overline{BC}}.$$

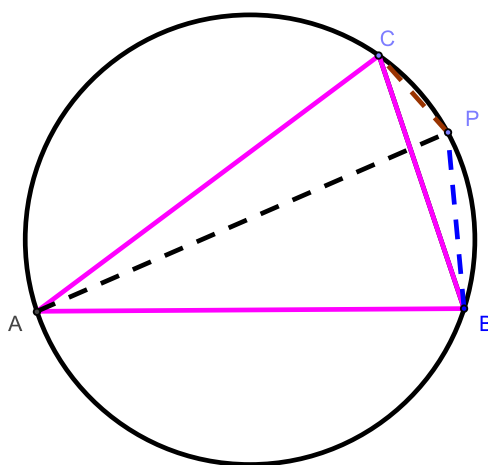


Figura 15

Demostración. Aplicando el teorema de Ptolomeo al cuadrilátero $ABPC$ obtenemos:

$$\overline{AP} \overline{BC} = \overline{CP} \overline{AB} + \overline{PB} \overline{AC} = \overline{CP} \overline{AC} + \overline{PB} \overline{AC} = \overline{AC}(\overline{CP} + \overline{PB}),$$

pues $\overline{AB} = \overline{AC}$. Por lo tanto

$$\frac{\overline{AP}}{\overline{CP} + \overline{PB}} = \frac{\overline{AC}}{\overline{BC}} \quad \blacksquare$$

Una consecuencia inmediata del teorema de Ptolomeo se debe al matemático holandés Frans van Schooten (1615-1660) y corresponde a triángulos equiláteros.

Teorema 1.16 (Van Schooten). *Si el triángulo ABC es equilátero y cada vértice es unido a un punto P en el arco \widehat{BC} (ver figura 16), entonces $\overline{AP} = \overline{CP} + \overline{PB}$.*

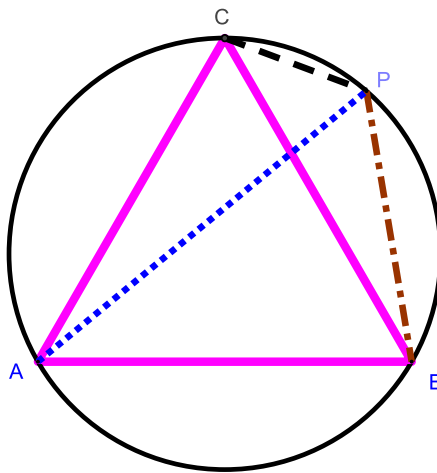


Figura 16

Demostración. Por el teorema anterior

$$\frac{\overline{AP}}{\overline{CP} + \overline{PB}} = \frac{\overline{AC}}{\overline{BC}} = 1,$$

pues $\overline{AC} = \overline{BC}$. Así $\overline{AP} = \overline{CP} + \overline{PB}$. ■

Ahora nos preguntamos ¿existe alguna fórmula para calcular el área de un triángulo o de un cuadrilátero cíclico?, uno de los geómetras más importantes del siglo XVIII, el matemático suizo Leonhard Euler (1707-1783) obtuvo algunos resultados que presentamos a continuación.

Teorema 1.17 (Euler). *Si el triángulo ABC tiene lados a, b y c , con área K y circunradio R , entonces $abc = 4KR$.*

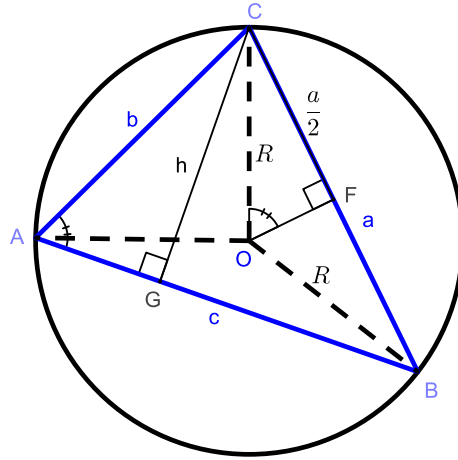


Figura 17

Demostración. Los triángulos CGA y CFO son semejantes (ver figura 17), luego

$$\frac{h}{b} = \frac{\frac{a}{2}}{R};$$

por lo tanto,

$$h = \frac{ab}{2R}.$$

Ahora calculemos el área del triángulo ABC :

$$K = \frac{ch}{2} = \frac{c}{2} \frac{ab}{2R} = \frac{abc}{4R}$$

Así $abc = 4KR$. ■

Teorema 1.18 (Euler-Ptolomeo). *Si el cuadrilátero $ABCD$ es cíclico con lados a, b, c, d y diagonales p y q (ver figura 18); de área K y circunradio R , entonces:*

$$i) \frac{p}{q} = \frac{ad + bc}{ab + cd}$$

$$ii) p = \sqrt{\frac{(ac + bd)(ad + bc)}{ab + cd}} \quad y \quad q = \sqrt{\frac{(ac + bd)(ab + cd)}{ad + bc}}$$

$$iii) 4KR = \sqrt{(ab + cd)(ac + bd)(ad + bc)}$$

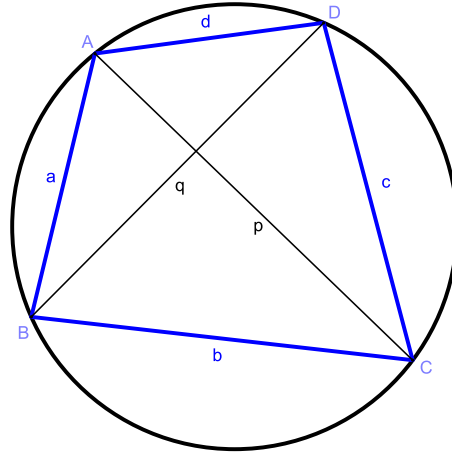


Figura 18

Demostración. Sean K_1, K_2, K_3 y K_4 las áreas de los $\triangle ABC$, $\triangle ACD$, $\triangle ABD$ y $\triangle BCD$ respectivamente. Así, aplicando el teorema anterior a los triángulos ABC y ACD (ver figura 19(a)) obtenemos:

$$K = K_1 + K_2 = \frac{pab}{4R} + \frac{pcd}{4R} = \frac{p(ab + cd)}{4R}.$$

Análogamente para los triángulos ABD y BCD (ver figura 19(b)) obtenemos:

$$K = K_3 + K_4 = \frac{qad}{4R} + \frac{qbc}{4R} = \frac{q(ad + bc)}{4R}$$

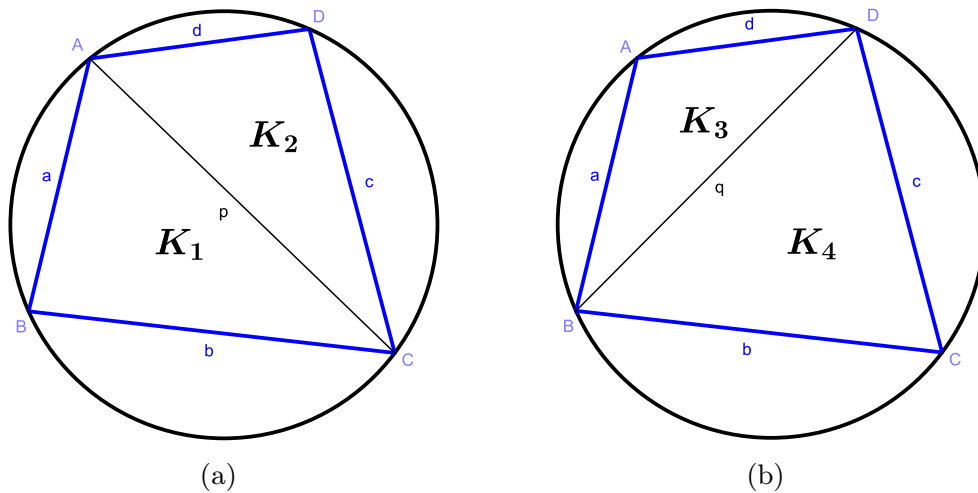


Figura 19

Por consiguiente,

$$\frac{p(ab + cd)}{4R} = \frac{q(ad + bc)}{4R};$$

luego $p(ab + cd) = q(ad + bc)$ y por lo tanto

$$\frac{p}{q} = \frac{ad + bc}{ab + cd}.$$

Para el segundo inciso $pq = ac + bd$ se cumple por Ptolomeo, por consiguiente:

$$p^2 = \frac{pq}{q} p = \frac{(ac + bd)(ad + bc)}{ab + cd}, \quad \text{luego} \quad p = \sqrt{\frac{(ac + bd)(ad + bc)}{ab + cd}}.$$

Análogamente,

$$q^2 = \frac{pq}{p} q = \frac{(ac + bd)(ab + cd)}{ad + bc}, \quad \text{luego} \quad q = \sqrt{\frac{(ac + bd)(ab + cd)}{ad + bc}}.$$

Para la tercera igualdad consideremos:

$$\begin{aligned} K^2 = K \cdot K &= \frac{p(ab + cd)}{4R} \cdot \frac{q(ad + bc)}{4R} = \frac{pq(ab + cd)(ad + bc)}{(4R)^2} \\ &= \frac{(ac + bd)(ab + cd)(ad + bc)}{(4R)^2} \end{aligned}$$

Por lo tanto $(4KR)^2 = (ac + bd)(ab + cd)(ad + bc)$, así

$$4KR = \sqrt{(ab + cd)(ac + bd)(ad + bc)}. \quad \blacksquare$$

Lema 1.19. Si el cuadrilátero $ABCD$ es cíclico con diagonales p y q (ver figura 20(a)), de área K , entonces $K \leq \frac{1}{2}pq$.

Demostración. $K = \frac{1}{2}p(h_1 + h_2)$ y como $h_1 + h_2 \leq q$, por lo tanto $K \leq \frac{1}{2}pq$. \blacksquare

Corolario 1.20. Si el cuadrilátero $ABCD$ es cíclico con diagonales p y q perpendiculares (ver figura 20(b)), de área K , entonces $K = \frac{1}{2}pq$.

Demostración. Por el lema anterior $K \leq \frac{1}{2}pq$, pero p y q son perpendiculares, así $h_1 + h_2 = q$ y por lo tanto $K = \frac{1}{2}pq$. \blacksquare

Observación 1.21. El lema 1.19 y el corolario 1.20 se cumplen para cualquier cuadrilátero convexo no necesariamente cíclico.

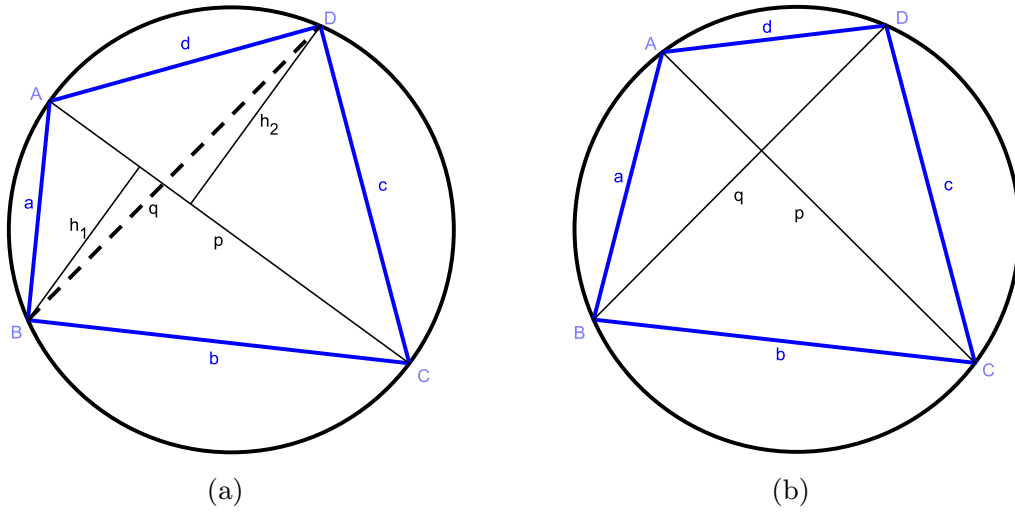


Figura 20

Para finalizar agregamos dos lemas, cuyas soluciones involucran cuadriláteros cíclicos.

Lema 1.22. *En el $\triangle ABC$, si $\overline{AB} > \overline{AC}$ y \overline{BP} es tangente al circuncírculo del $\triangle ABC$ con $\overline{AC} \cap \overline{BP} = P$, D es el simétrico de B con respecto a P y E es el simétrico de C con respecto a \overline{BP} (ver figura 21(a)), entonces $\square ABED$ es cíclico.*

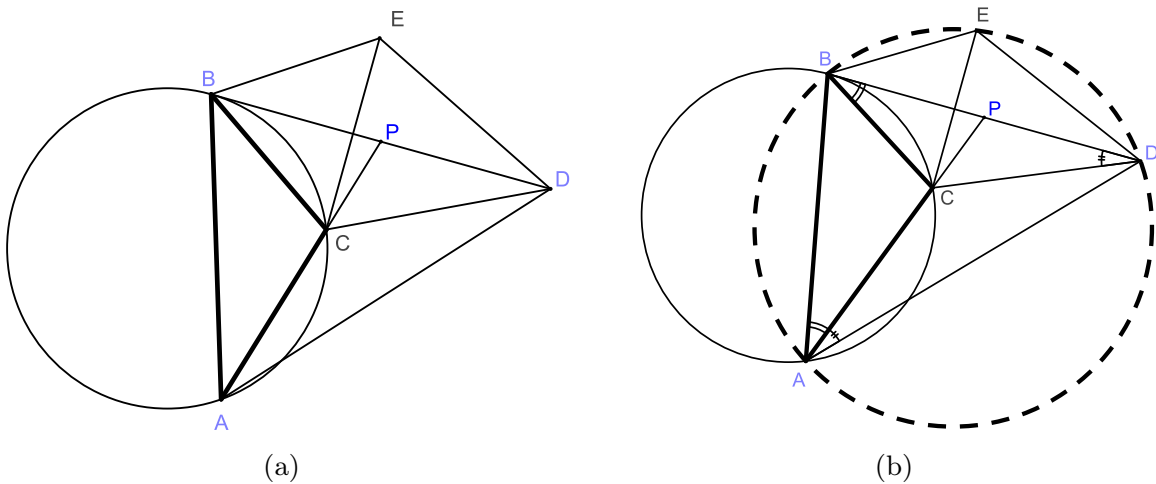


Figura 21

Demostración. $\triangle PBC \approx \triangle BAP$, luego

$$\frac{\overline{PA}}{\overline{PB}} = \frac{\overline{PB}}{\overline{PC}},$$

por lo tanto $\overline{PA} \overline{PC} = \overline{PB}^2 = \overline{PD}^2$ pues $\overline{PB} = \overline{PD}$, luego

$$\frac{\overline{PA}}{\overline{PD}} = \frac{\overline{PD}}{\overline{PC}}.$$

Además $\angle DPA = \angle DPC$, por lo tanto $\triangle DPA \approx \triangle DPC$, luego $\angle CAD = \angle PAD = \angle CDP$. Por otra parte C y E son simétricos con respecto a \overline{PB} , por consiguiente $\angle CPD = \angle PDE = \angle BDE$ y $\angle PBC = \angle EBD$. Por lo tanto:

$$\begin{aligned} \angle BAD &= \angle BAP + \angle PAD = \angle CDP + \angle PBC \\ &= \angle BDE + \angle EBD = 180^\circ - \angle BED, \end{aligned}$$

luego $\angle BAD + \angle BED = 180^\circ$, así $\square ABED$ es cíclico. ■

Lema 1.23. Dado el $\triangle ABC$ inscrito en el círculo con centro en O y el diámetro $\overline{CD} \perp \overline{AB}$, $E = \overline{CD} \cap \overline{AB}$. Si la cuerda \overline{BF} intersecta a \overline{CD} y \overline{AC} en M y N respectivamente, con $\overline{BF} = \overline{AC}$ (ver figura 22(a)), entonces

- (a) $\triangle ACM \cong \triangle BCM$
- (b) $\overline{AD} \overline{BE} = \overline{DE} \overline{BC}$
- (c) $\overline{MB}^2 = \overline{MN} \overline{MF}$

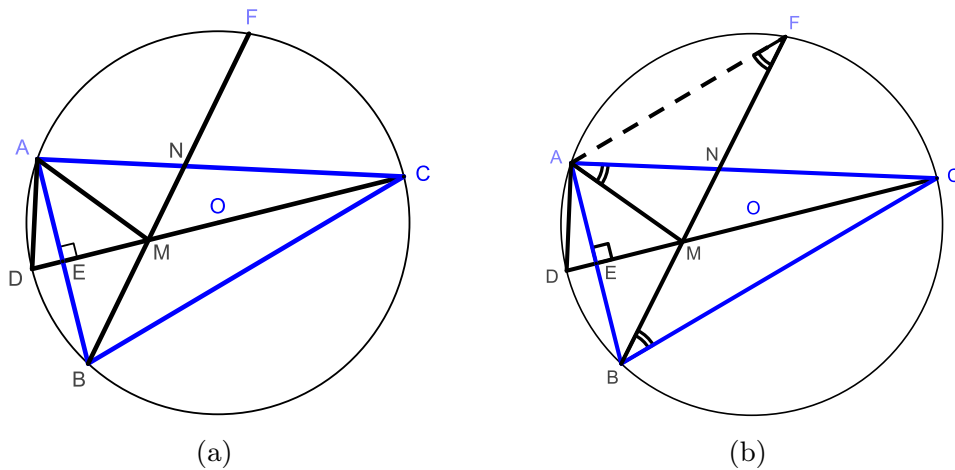


Figura 22

Demostración.

(i) Como $\overline{CD} \perp \overline{AB}$ y \overline{CD} es diámetro, entonces $\triangle AMB$ y $\triangle ACB$ son isósceles. Así $\overline{AM} = \overline{MB}$, $\overline{AC} = \overline{CB}$ y \overline{CM} es común, por lo tanto $\triangle ACM \cong \triangle BCM$.

(ii) Sabemos que $\angle EAD = \angle BCE$ por el teorema del ángulo central y $\angle DEA = \angle CEB$ porque son ángulos opuestos por el vértice, luego $\angle ADE = \angle EBC$, por lo tanto $\triangle EAD \approx \triangle BCE$ y por consiguiente

$$\frac{\overline{BE}}{\overline{DE}} = \frac{\overline{BC}}{\overline{AD}},$$

así $\overline{AD} \overline{BE} = \overline{DE} \overline{BC}$.

(iii) Trazamos \overline{AF} de tal manera que $\overline{BF} = \overline{AC}$ (ver figura 22(b)), luego $\widehat{BAF} = \widehat{AFC}$. Por lo tanto $\widehat{AB} = \widehat{FC}$, así $\angle CBF = \angle AFB$ y como $\triangle ACM \cong \triangle BCM$, por consiguiente

$$\angle MFA = \angle BFA = \angle CBF = \angle CBM = \angle MAC = \angle MAN$$

luego $\triangle NAM \approx \triangle MFA$, así

$$\frac{\overline{MN}}{\overline{AM}} = \frac{\overline{AM}}{\overline{MF}}$$

luego $\overline{AM}^2 = \overline{MN} \overline{MF}$ y como $\overline{AM} = \overline{MB}$, entonces $\overline{MB}^2 = \overline{MN} \overline{MF}$. ■

Agradecimientos

Agradezco al árbitro su paciencia en la revisión, así como sus valiosos comentarios y sugerencias para el enriquecimiento de este trabajo.

Bibliografía

- [1] C. Alsina and R. B. Nelsen, *Icons of Mathematics*, The Mathematical Association of America; (2011).
- [2] N. Altshiller-Court, *College geometry: An introduction to the modern geometry of the triangle and the circle*, Dover; (2007).
- [3] H.S.M. Coxeter, *Fundamentos de geometría*, Limusa, México, (1971).
- [4] H. Eves, *Estudio de las Geometrías Tomo 1*, UTEHA, México, (1985).

-
- [5] D. C. Kay, *College geometry a unified development*, CRC Press; (2012).
- [6] R. Honsberger, *Episodes in nineteenth and twentieth century Euclidean geometry*, The Mathematical Association of America; (1995).
- [7] A. S. Posamentier, *Advanced Euclidean geometry*, Key College Publishing; (2002).
- [8] A. S. Posamentier and Ch. T. Salkind, *Challenging problems in geometry*, Dover; (1996).

Universidad Autónoma de la Ciudad de México, UACM
Fray Servando Teresa de Mier 92 y 99, Colonia Centro,
México, D. F. C.P. 06080.

Facultad de Ciencias, UNAM
Circuito Exterior S/N, Cd. Universitaria,
México, D. F. C.P. 04510.
amersen@yahoo.com.mx

Probabilidad y Estadística

Capítulo 6

Análisis del Modelo Constante por tramos con Puntos de Cambio en la Función de Riesgo

Yoanna Arenas Martínez, Hugo Adán Cruz Suárez, Hortensia Reyes Cervantes, Francisco Tajonar Sanabria
FCFM-BUAP

Resumen

El presente trabajo se ubica en el área de estadística inferencial, en particular en el análisis de supervivencia. Las funciones de riesgo son determinantes en el análisis de supervivencia, ya que estas cuantifican el riesgo instantáneo de falla o fracaso en un momento dado. Se ha puesto atención a la prueba de hipótesis de una tasa de falla constante en contra de una que cambia en un solo punto. En el trabajo se analizará la alternativa de múltiples puntos de cambio, y se propondrá un enfoque de selección del modelo utilizando la prueba secuencial que considera al modelo de riesgo constante por tramos. Se analiza el modelo con el objetivo de estudiar el comportamiento de la función de riesgo y de observar cuando se produce un cambio.

1 Introducción

El análisis de supervivencia es una línea de investigación de la estadística inferencial, que surge como una teoría paralela a la llamada “Teoría de confiabilidad” la cual tiene como propósito el estudio del tiempo de vida o el tiempo de falla de un objeto (individuo) o un conjunto de objetos (conjunto de individuos) sometidos a un cierto tipo de tensión, ver ([3], [8]). El análisis de supervivencia tiene como propósito fundamental el estudio del tiempo de vida o falla de un evento de interés que se encuentra bajo estudio. Las funciones de riesgo son un componente muy importante en dicho análisis, ya que describen el riesgo inminente de falla en un instante de tiempo dado.

En la actualidad el uso de modelos paramétricos para representar fenómenos relacionados con tiempos de vida o tiempos de falla ha ido en aumento en diversas áreas de investigación, tales como: medicina, biología, física, industria e ingeniería, con aplicaciones que van desde el estudio de individuos con enfermedades terminales, en la comparación de la eficiencia de k tratamientos buscando prolongar la

vida de pacientes desahuciados, así como en el estudio de artículos de uso cotidiano bajo condiciones de uso normal o de vida acelerada, garantizar el funcionamiento de electrodomésticos durante t unidades de tiempo, entrega de componentes electrónicas confiables, etc., son sólo algunas situaciones donde el análisis de supervivencia o la teoría de confiabilidad está presente. En la literatura, existen modelos que se han propuesto para representar o ajustar datos de tiempo de vida, entre los más usados se encuentran, el Exponencial, el Weibull, el Gama y el Log-normal, estos se han empleado para ajustar una gran cantidad de datos. Sin embargo, existen conjuntos de datos que no pueden ser ajustados por estos modelos, por esta razón, surge la necesidad de tener una familia más general que permita modelar una clase más amplia de datos de tiempos de vida. Un modelo que han sugerido varios autores, para analizar datos de tiempos de vida es, el Gama Generalizado, este es un modelo paramétrico más general y que incluye como caso particular a cada uno de los citados anteriormente. Sin embargo, tiene una limitante significativa que lo pone en desventaja frente a los modelos clásicos, esta es, que la estimación de sus parámetros no es tan sencilla y se hace necesario usar un método numérico para la estimación.

Existen diferentes situaciones en donde la estimación explícita de la función de riesgo es de gran utilidad, uno de estos casos, es el punto de cambio de los modelos de duración, estos asumen una función con tasas de riesgo diferentes que cambian en algunos puntos en muy poco tiempo, dichos puntos son llamados puntos de cambio, los cuales se requieren determinar y estimar. Suponga que el número máximo de puntos de cambio en el modelo es k , nuestro objetivo es encontrar un modelo con k puntos de cambio, que mejor se adapte o ajuste a los datos, esto se puede pensar como una selección de modelos donde se comenzara con un modelo que no tenga puntos de cambio y se realizara la prueba de hipótesis para compararla con un modelo con al menos un punto de cambio, a continuación se puede comparar el modelo con un punto de cambio en contra del modelo con dos puntos de cambio, y así sucesivamente, solo se pasa a la siguiente prueba si se rechaza la hipótesis nula.

El presente trabajo se ubica dentro del análisis de supervivencia y tiene como objetivo principal el análisis de un modelo de regresión con múltiples puntos de cambio, dichos cambios se consideran tanto en la función de riesgo como en los parámetros de regresión, los cuales ocurren en algún instante de tiempo τ , el cual es desconocido.

Además, se proponen estimadores para los puntos de cambio en la función de riesgo y para los parámetros de regresión mediante el método de máxima verosimilitud.

2 Conceptos básicos

En la mayoría de los estudios quirúrgicos, y especialmente los relacionados con el problema de cáncer, la variable principal que se quiere valorar es el tiempo que transcurre hasta que ocurre un suceso de interés. A este tiempo, generalmente se le llama tiempo de vida o tiempo de falla, aunque en gran parte de las investigaciones el suceso en cuestión no tiene por qué ser la muerte.

Se considera que el tiempo de vida de un objeto, persona o grupo (de personas) se estudia a través de una variable aleatoria T . Más específicamente, T es una variable aleatoria continua no negativa, que representa el tiempo de vida de un individuo que se encuentra en la población de estudio. Por lo que estas variables aleatorias están definidas sobre el intervalo de tiempo $[0, \infty)$, se denota por $f(t)$ a la función de densidad de probabilidad de T .

Definición 2.1. La probabilidad de que un individuo sobreviva hasta un tiempo t está definida como

$$S(t) = P(T \geq t) = \int_t^{\infty} f(x)dx. \quad (1)$$

A la función $S(t)$ se le llama función de supervivencia de T .

Observación 2.2. En términos de la función de distribución de T , que es denotada por $F(t)$ se tiene que la función de supervivencia se expresa como

$$S(t) = 1 - P(T \leq t) = 1 - F(t). \quad (2)$$

En el caso en que se estudien sistemas que estén dados bajo ciclos o en tiempos de vida de fabricación de objetos, $S(t)$ representa a la función de confiabilidad, donde:

- * $S(t)$ es una función monótona decreciente continua
- * $S(0) = 1$
- * $S(\infty) = \lim_{t \rightarrow \infty} S(t) = 0$

Nota 2.3. En el caso en donde $S(\infty) > 0$, esto indicará que el individuo no muere o no falla, estos son tratados como casos especiales, los cuales no se trataran en el presente trabajo.

Otra función que es muy importante en el análisis de tiempos de vida es la *función de riesgo*, [5].

Definición 2.4. La función de riesgo asociada al tiempo de vida o de falla de T está dada por

$$h(t) = \frac{f(t)}{1 - F(t)}, \quad (3)$$

con $F(t) < 1$.

La función de riesgo especifica la tasa de falla instantánea en un tiempo t ; para comprender este resultado, observe que la función de riesgo está dada mediante una razón de cambio, en la cual se hace uso de la probabilidad condicional,

$$\begin{aligned} h(t) &= \frac{f(t)}{1 - F(t)} \\ &= \frac{F'(t)}{S(t)} \\ &= \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t S(t)} \\ &= \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T \leq t + \Delta t)}{\Delta t P(T \geq t)} \\ &= \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T \leq t + \Delta t \mid T \geq t)}{\Delta t}. \end{aligned} \quad (4)$$

Un primer resultado que se establece para la función de riesgo es el siguiente

Teorema 2.5. Si $f(t) = -S'(t)$ entonces

$$h(x) = -\frac{d \log S(x)}{dx} \quad y \quad \log S(x)|_0^t = -\int_0^t h(x) dx. \quad (5)$$

Demostración. Haciendo uso de (2), se obtiene la primera parte, esto es

$$-\frac{d \log S(x)}{dx} = -\left[-\frac{S'(x)}{S(x)} \right] = \frac{f(x)}{S(x)} = h(x). \quad (6)$$

Se continúa con la segunda igualdad, para lo cual partimos del lado izquierdo y de (6), de lo cual se obtiene que

$$\begin{aligned} -\int_0^t h(x) dx &= -\int_0^t -\frac{d \log S(x)}{dx} dx \\ &= \int_0^t \frac{d \log S(x)}{dx} dx = \log S(x) \Big|_0^t. \quad \square \end{aligned}$$

Un resultado que determina una relación importante entre la función de supervivencia y la función de riesgo es el siguiente

Teorema 2.6. Si $S(0) = 1$, entonces

$$S(t) = \exp\left(-\int_0^t h(x)dx\right). \quad (7)$$

Demostración. Partiendo del lado izquierdo de la igualdad y utilizando la segunda igualdad del Teorema 2.5 se tiene que

$$\begin{aligned} \exp\left(-\int_0^t h(x)dx\right) &= \exp\left(\log S(t)\Big|_0^t\right) \\ &= \exp\left(\log S(t) - \log S(0)\right) \\ &= \exp\left(\log S(t)\right) \\ &= S(t). \end{aligned} \quad \square$$

Definición 2.7. La función de riesgo acumulativa correspondiente al tiempo de vida o de falla de T está definida por

$$H(t) = \int_0^t h(x)dx. \quad (8)$$

Una relación inmediata entre la función de supervivencia y la de riesgo acumulado es la siguiente

Corolario 2.8. Si T el tiempo de vida o de falla con $S(t)$ y $H(t)$, su función de supervivencia y función de riesgo acumulada respectivamente, entonces

$$S(t) = \exp(-H(t)). \quad (9)$$

Demostración. Para probar este resultado es conveniente hacer uso de (8) y de la primera igualdad del Teorema 2.5 de la siguiente manera

$$\begin{aligned} \exp(-H(t)) &= \exp\left(-\int_0^t h(x)dx\right) \\ &= \exp\left(-\int_0^t -\frac{d \log S(x)}{dx} dx\right) \\ &= \exp(\log S(t)) \\ &= S(t), \end{aligned}$$

que es lo que se deseaba probar. □

El siguiente resultado muestra una relación única entre la función de densidad y de riesgo, la cual será de vital importancia durante el desarrollo del trabajo.

Corolario 2.9. *Si $f(t)$ y $h(t)$ son las funciones de densidad y de riesgo respectivamente de T entonces*

$$f(t) = h(t) \exp\left(-\int_0^t h(x)dx\right). \quad (10)$$

Demostración. Para realizar la prueba se utilizan (3), (7) y el Teorema 2.5, haciendo uso de estos resultados se tiene que

$$\begin{aligned} h(t) \exp\left(-\int_0^t h(x)dx\right) &= -\frac{d \log S(t)}{dt} \exp\left(-\int_0^t h(x)dx\right) \\ &= -\frac{S'(t)}{S(t)} S(t) \\ &= f(t), \end{aligned}$$

que es lo deseaba probar. □

2.1 Modelos discretos

Cuando se agrupan tiempos de vida o medidas como un número de ciclos cortos de algún tipo, usualmente T se puede tratar como una variable aleatoria discreta, para más detalles se puede revisar [5].

Suponga que T puede tomar los valores t_1, t_2, \dots , con $0 \leq t_1 \leq t_2, \dots$, y con función de probabilidad

$$f(t_j) = P(T = t_j), \quad j = 1, 2, \dots$$

Definición 2.10. Sea T una variable aleatoria discreta que determina el tiempo de vida o de falla de un individuo, se define a su función de supervivencia mediante

$$S(t) = P(T \geq t) = \sum_{j:t_j > t} f(t_j). \quad (11)$$

Algunas propiedades de la función de supervivencia son las siguientes

- * $S(t)$ es continua por la izquierda.
- * $S(t)$ es no creciente.
- * $S(0) = 1$ y $S(\infty) = 0$.

Definición 2.11. Sea T el tiempo de vida o falla de un individuo, su función de riesgo está dada por,

$$h(t_j) = P(T = t_j | T \geq t_j) = \frac{f(t_j)}{S(t_j)}. \quad (12)$$

El siguiente resultado nos indica una relación existente entre la función de supervivencia y la de riesgo.

Teorema 2.12. Si $f(t_j) = S(t_j) - S(t_{j+1})$ entonces

$$h(t_j) = 1 - \frac{S(t_{j+1})}{S(t_j)}, \quad (13)$$

y

$$S(t) = \prod_{j:t_j < t} [1 - h(t_j)]. \quad (14)$$

Demostración. Primero se muestra (13), para lo cual se hace uso de la Definición 2.11 y se procede de la siguiente manera

$$\begin{aligned} h(t_j) &= \frac{f(t_j)}{S(t_j)} \\ &= \frac{S(t_j) - S(t_{j+1})}{S(t_j)} \\ &= 1 - \frac{S(t_{j+1})}{S(t_j)}, \end{aligned}$$

una vez obtenida la igualdad (13) se continua con la ecuación (14) partiendo del lado derecho

$$\prod_{j:t_j < t} [1 - h(t_j)] = \prod_{j:t_j < t} \left[1 - \left[1 - \frac{S(t_{j+1})}{S(t_j)} \right] \right] = \prod_{j:t_j < t} \frac{S(t_{j+1})}{S(t_j)} = S(t),$$

de esta manera queda demostrado el teorema. \square

En la sección anterior se obtuvo una relación entre la función de supervivencia y la función de riesgo, procediendo de manera análoga, se obtiene el siguiente resultado, teniendo en consideración que la variable aleatoria T ahora es discreta.

Corolario 2.13. Sea T el tiempo de vida o de falla de un individuo, con función de supervivencia dada por $S(t) = \prod_{j:t_j < t} [1 - h(t_j)]$ entonces su función de riesgo acumulativa esta dada por

$$H(t) = -\log S(t).$$

De esta forma se tiene una relación entre la función de riesgo acumulativa y la de supervivencia.

3 Modelos paramétricos clásicos

En esta sección se presentan algunos modelos paramétricos que generalmente se utilizan para llevar a cabo un análisis de tiempos de vida.

3.1 Distribución exponencial

Uno de los modelos de falla más relevantes es aquel en el que los tiempo de vida se describen mediante una distribución exponencial. La cual se puede definir de varias maneras pero la más simple es suponer que la función de riesgo es constante, es decir, $h(t) = \alpha$, donde α es una constante positiva.

Definición 3.1. Sea T el tiempo de vida o de falla de un individuo (o grupo), se dice que T tiene una distribución exponencial con parámetro $\alpha > 0$, si su función de densidad de probabilidad está dada por

$$f(t) = \alpha e^{-\alpha t}, \quad t > 0. \tag{15}$$

Teorema 3.2. Si T tiene una distribución exponencial, entonces la función de supervivencia tiene la forma

$$S(t) = 1 - F(t) = e^{-\alpha t}, \tag{16}$$

y su función de riesgo es

$$h(t) = \frac{f(t)}{S(t)} = \alpha. \tag{17}$$

Demostración. La demostración se sigue de la definición se supervivencia y de la función de riesgo. □

Enseguida se enuncia el siguiente resultado importante.

Teorema 3.3. Sea T el tiempo de vida o de falla de un individuo. T tiene una distribución exponencial si y sólo si su función de riesgo es constante.

Demostración.

\Rightarrow] Si T tiene una distribución exponencial el resultado se sigue de lo anterior.

\Leftarrow] Ahora, suponga que T tiene una función de riesgo constante, es decir, $h(t) = \alpha$; el objetivo es probar que T tiene una distribución exponencial, pero por el Corolario 2.9 se tiene que

$$f(t) = \alpha \exp \left[- \int_0^t \alpha ds \right] = \alpha \exp \left(- \alpha s \Big|_0^t \right) = \alpha e^{-\alpha t},$$

por consiguiente T tiene una distribución exponencial. □

Este resultado se puede interpretar como el hecho de que no hay efecto de desgaste cuando se establece un modelo exponencial. Considere que $\Delta t > 0$, luego la probabilidad de falla o de muerte durante las siguientes Δt unidades de tiempo, dado que no falla en el instante t , esta representada por

$$P(t < T \leq t + \Delta t | T > t).$$

Aplicando la definición de probabilidad condicional se tiene que

$$P(t \leq T \leq t + \Delta t | T > t) = \frac{e^{-\alpha t} - e^{-\alpha(t+\Delta t)}}{e^{-\alpha t}} = 1 - e^{-\alpha \Delta t}. \quad (18)$$

Nota 3.4. Esta probabilidad ya no depende de T y sólo depende de Δt , lo cual indica que en una ley exponencial de fallas, la posibilidad de fallar no depende de lo que ha ocurrido antes, es decir, mientras siga funcionando se considera tan bueno como al inicio.

3.2 Distribución Weibull

La distribución Weibull representa un modelo apropiado para eventos en los cuales intervienen cierto número de componentes y cuando la falla se debe principalmente al defecto o accidente más grave, ver [10].

Definición 3.5. Sea T el tiempo de vida o de falla de un individuo, se dice que tiene una distribución Weibull con parámetros $\lambda > 0$ y $\beta > 0$, si su función de densidad de probabilidad está dada por

$$f(t) = (\alpha\beta)t^{\beta-1}e^{-\alpha t^\beta} \quad (19)$$

donde $\lambda > 0$ y $\beta > 0$ son constantes.

Teorema 3.6. Si T tiene una distribución Weibull entonces su función de riesgo es de la siguiente forma

$$h(t) = (\alpha\beta)t^{\beta-1}. \quad (20)$$

Se considera la Definición 3.5 y el Teorema 3.6 para probar que la función de supervivencia para T está dada por

$$S(t) = e^{-\alpha t^\beta}. \quad (21)$$

Algunas observaciones de esta distribución son las siguientes:

- * La distribución exponencial es un caso especial de la distribución Weibull.
- * $h(t)$ no es una constante pero es proporcional a las potencias de t , es decir, está es creciente, decreciente o constante, según sea el valor de β .

3.3 Distribución gamma

Esta es otra de las distribuciones más utilizadas en el ajuste de datos de tiempos de vida.

Definición 3.7. Sea T el tiempo de vida o de falla de un individuo (o grupo), se dice que, T tiene una distribución *Gamma* si su función de densidad de probabilidad es de la siguiente forma

$$f(t) = \frac{\lambda(\lambda t)^{k-1} \exp(-\lambda t)}{\Gamma(k)}, \quad t > 0, \quad (22)$$

donde k, λ son sus parámetros, λ^{-1} es conocido como el parámetro de escala y k es algunas veces llamado el parámetro de forma.

Esta distribución al igual que la distribución *Weibull* incluye como un caso especial a la distribución exponencial, cuando se toma $k = 1$.

La función de riesgo y la de supervivencia están dadas mediante la función *Gamma incompleta* que tienen la siguiente expresión

$$I(k, x) = \frac{1}{\Gamma(k)} \int_0^x u^{k-1} \exp(-u) du.$$

La función de supervivencia está dada por

$$S(t) = 1 - I(k, \lambda t).$$

La función de riesgo para T está denotada por

$$h(t) = \frac{\lambda(\lambda t)^{k-1} \exp(-\lambda t)}{\Gamma(k)(1 - I(k, \lambda t))}.$$

4 Censura a la derecha

La censura por la derecha se da en los límites inferiores de los tiempos de vida y sólo son permitidos para algunos individuos, esto puede ocurrir por varias razones, cuando el investigador establece criterios bajo los cuales descartara al individuo, al tomar una decisión de prueba de vida antes de que muera en algún caso no planificado, etc.

Cuando una persona es estudiada bajo cierto propósito, y estás quedan censuradas o descartadas por motivos externos al estudio, se dice que se tiene una *pérdida de seguimiento*, puesto que ya no se encuentran dentro de los criterios de

evaluación de dicho análisis. Basados en datos de censura, se necesita un modelo de probabilidad para los mecanismos de censura, [5].

Para poder llevar a cabo el análisis es necesario introducir alguna notación para los tiempos de vida y de censura. Supóngase que T_1, \dots, T_n , es una muestra aleatoria de n individuos que determinan su tiempo de vida y que C_1, \dots, C_n son los correspondientes tiempos de censura.

Se define a

$$\delta_i = \begin{cases} 1, & \text{si } T_i \leq C_i, \\ 0, & \text{si } T_i > C_i. \end{cases}$$

por lo que δ_i es la función indicadora de censura. Los datos obtenidos se analizan mediante las parejas (T_i, C_i) desde $i = 1, \dots, n$.

4.1 Censura tipo I

La censura tipo I es empleada cuando se llevan a cabo estudios sobre un periodo de tiempo específico. La censura tipo I es aplicable cuando cada individuo tiene un tiempo t_i y un potencial fijo de censura $\delta_i > 0$, el cual se denota por

$$t_i = \text{mín}(T_i, C_i), \quad \delta_i = I(T_i \leq C_i), \quad (23)$$

donde T_i y C_i representan el tiempo de vida y de censura respectivamente.

La función de verosimilitud para la muestra con censura tipo I esta basada en la función de probabilidad de (t_i, δ_i) , para $i = 1, \dots, n$, la cual está dada por

$$f(t_i)^{\delta_i} P(T_i > C_i)^{1-\delta_i}. \quad (24)$$

Observación 4.1. Si C_i es una constante fija y T_i toma valores menores o iguales a C_i se tiene que

$$\begin{aligned} P(t_i = C_i, \delta_i = 0) &= P(T_i > C_i) \\ P(t_i, \delta_i = 1) &= f(t_i), \quad t_i \leq C_i, \end{aligned}$$

donde la segunda expresión denota la función de densidad de probabilidad para T_i .

Asumimos que los tiempos de vida T_1, \dots, T_n son estadísticamente independientes, así obtenemos que la función de verosimilitud toma la forma

$$\begin{aligned} L &= \prod_{i=1}^n f(t_i)^{\delta_i} P(T_i > C_i)^{1-\delta_i} \\ &= \prod_{i=1}^n f(t_i)^{\delta_i} S(t_{i+})^{1-\delta_i} \end{aligned} \quad (25)$$

donde $S(t_{i+}) = P(T_i > t_i)$. En general, si $S(t)$ es continua entonces $S(t_{i+}) = S(t_i)$.

4.2 Censura aleatoria independiente

El proceso de censura aleatoria es muy simple ya que con frecuencia se realiza para cada individuo con un tiempo de vida T y un tiempo de censura C , que son variables aleatorias continuas e independientes, con funciones de supervivencia y de distribución $S(t)$ y $G(t)$ respectivamente, para todos los tiempos de vida y de censura se asume independencia, por lo que $G(t)$ no depende de ninguno de los parámetros de $S(t)$. Esto nos indica que es un caso de censura tipo I.

Una vez más es necesario considerar a (t_i, δ_i) , los cuales están representados mediante

$$t_i = \min(T_i, C_i) \quad \text{y} \quad \delta_i = \begin{cases} 1, & \text{si } T_i \leq C_i, \\ 0, & \text{si } T_i > C_i. \end{cases}$$

Los datos de las observaciones de n individuos consisten de las parejas (t_i, δ_i) para $i = 1, \dots, n$, el mismo resultado final es obtenido si C_i es adecuado para todo $i = 1, \dots, n$, y la función de densidad de probabilidad es obtenida para (t_i, δ_i) . Si $f(t)$ y $g(t)$ son las funciones de densidad de probabilidad de T_i y C_i entonces

$$\begin{aligned} P(t_i = t, \delta_i = 0) &= P(C_i = t, T_i > C_i) = g(t)S(t) \\ P(t_i = t, \delta_i = 1) &= P(T_i = t, T_i \leq C_i) = f(t)G(t) \end{aligned}$$

esto se puede expresar de la siguiente manera

$$P(t_i = t, \delta_i) = [f(t)G(t)]^{\delta_i} [g(t)S(t)]^{1-\delta_i},$$

y la distribución de (t_i, δ_i) , para $i = 1, \dots, n$ es la siguiente

$$\prod_{i=1}^n [f(t_i)G(t_i)]^{\delta_i} [g(t_i)S(t_i)]^{1-\delta_i}.$$

Dado que $G(t)$ y $g(t)$ no involucran parámetros de $f(t)$, la función de verosimilitud estaría dada por

$$L = \prod_{i=1}^n f(t_i)^{\delta_i} S(t_i)^{1-\delta_i}, \quad (26)$$

que es análoga a (25).

El modelo de censura independiente es de gran utilidad en los procesos de tiempos de falla o de muerte, aunque la independencia del modelo de censura aleatoria es a menudo razonable, en muchas situaciones el proceso de censura está vinculado con los procesos de tiempos de falla, ver [8].

4.3 Censura tipo II

La censura tipo II sólo considera los r primeros tiempos de vida más pequeños $t_{(1)} \leq \dots \leq t_{(r)}$ de una muestra aleatoria de tamaño n , donde r es un entero entre 1 y n .

Este tipo de censura surge cuando n individuos inician un estudio en algún tiempo específico t , y el estudio es terminado cuando se han obtenido los primeros r tiempos de fallas. El análisis de tiempos de vida que son planteados con censura tipo II, tienen una desventaja en la práctica, ya que en este tipo de censura se desea que el tiempo $t_{(r)}$ sea dado de manera aleatoria, lo cual no permite establecer un tiempo de inicio.

En el caso de la censura tipo II el valor de r es elegido antes de hacer la recolección de los datos, los cuales deben consistir en los primeros r tiempos de vida más pequeños de la muestra aleatoria.

Para las distribuciones continuas podemos denotar al tiempo de vida más pequeño como $T_{(1)} \leq \dots \leq T_{(r)}$ y dar el siguiente resultado

Teorema 4.2. Si T_i tiene una función de densidad de probabilidad $f(t)$ y función de supervivencia $S(t)$, entonces su función de densidad de probabilidad conjunta de $T_{(1)} \leq \dots \leq T_{(r)}$ es

$$\frac{n!}{(n-r)!} \left[\prod_{i=1}^r f(t_i) \right] S(t_{(r)})^{n-r}. \quad (27)$$

Demostración. Haciendo uso de la probabilidad conjunta se tiene que

$$\begin{aligned} f_{t_1, \dots, t_r}(t_1, \dots, t_r) &= n! f(t_1) \cdot f(t_2) \cdots f(t_r) \frac{1}{(n-r)!} [1 - F(t_{(r)})]^{n-r} \\ &= \frac{n!}{(n-r)!} \prod_{i=1}^r f(t_i) [1 - F(t_{(r)})]^{n-r} \\ &= \frac{n!}{(n-r)!} \prod_{i=1}^r f(t_i) [S(t_{(r)})]^{n-r}. \quad \square \end{aligned}$$

5 Formulación de supervivencia

Las distribuciones continuas, discretas y mixtas las podemos tratar en un mismo sistema, para esto, vamos a introducir dos tipos de integrales que son, la *integral de Riemann-Stieltjes* y la *integral producto*, las cuales serán de gran utilidad para definir las funciones de riesgo y de supervivencia.

Sea $G(u)$ una función no decreciente, continua por la derecha y con límite a la izquierda con un número finito de discontinuidades en un conjunto de medida

cero. Asumimos que $g(u) = G'(u)$ existe excepto en los puntos de discontinuidad de G , dichos puntos tiene la siguiente forma $G(a_j) - G(a_j-) = g_j$ donde $G(a_j-) = \lim_{\Delta a \rightarrow 0} G(a - \Delta a)$.

Definición 5.1. La integral de *Riemann-Stieltjes* de dG sobre el intervalo $(a, b]$ está definida por

$$\int_{(a,b]} dG(u) = \int_a^b g(u)du + \sum_{j:a \leq a_j \leq b} g_j. \quad (28)$$

En general, una función de distribución $F(t) = P(T \leq t)$ es continua por la derecha, no decreciente, con saltos en los puntos a_j para los cuales $P(T = a_j) = f_j$ para $f_j > 0$ y la función de densidad de probabilidad es $f(u) = F'(u)$ en los puntos donde $F(u)$ es continua, de esta forma

$$\begin{aligned} P(a < T \leq b) &= F(b) - F(a) = \int_{(a,b]} dF(u) \\ &= \int_a^b f(u)du + \sum_{j:a \leq a_j \leq b} f_j. \end{aligned} \quad (29)$$

Observación 5.2. Si $F(t)$ es continua entonces no tiene discontinuidades, y si T tiene una distribución discreta entonces $F(u)$ es una función discreta con $f(u) = 0$ en los puntos donde es continua.

Para un tratamiento general de la función de riesgo se introduce la integral producto.

Definición 5.3. Sea $a = u_0 < u_1 < \dots < u_m = b$ una partición del intervalo $(a, b]$ con $\Delta u_i = u_i - u_{i-1}$ y el $\max(\Delta u_i) \rightarrow 0$ cuando $m \rightarrow \infty$. La integral producto de $dG(u)$ está definida por

$$\prod_{(a,b]} [1 + dG(u)] = \lim_{m \rightarrow \infty} \prod_{i=1}^m [1 + G(u_i) - G(u_{i-1})]. \quad (30)$$

Teorema 5.4. Si $G(u)$ es continua para todo u en $(a, b]$ y $dG(u) = g(u)du$, entonces

$$\begin{aligned} \prod_{(a,b]} [1 + g(u)du] &= \lim_{m \rightarrow \infty} \prod_{i=1}^m [1 + g(u_i)\Delta u_i + o(\Delta u_i)] \\ &= \lim_{m \rightarrow \infty} \prod_{i=1}^m [1 + g(u_i)\Delta u_i], \end{aligned}$$

donde $o(x)$ es una función de medida de $w(x)$ tal que $\frac{w(x)}{x} \rightarrow 0$ cuando $x \rightarrow 0$, para más detalles ver [5].

Demostración. Haciendo uso de la Definición 5.3 se tiene que

$$\begin{aligned}
 \prod_{i=1}^m [1 + g(u)du] &= \prod_{i=1}^m [1 + dG(u)du] \\
 &= \lim_{m \rightarrow \infty} \prod_{i=1}^m [1 + G(u_i) - G(u_{i-1})] \\
 &= \lim_{m \rightarrow \infty} \prod_{i=1}^m [1 + P(T \leq u_i) - P(T \leq u_{i-1})] \\
 &= \lim_{m \rightarrow \infty} \prod_{i=1}^m [1 + P(u_{i-1} \leq T \leq u_i)] \\
 &= \lim_{m \rightarrow \infty} \prod_{i=1}^m [1 + g(u_i)\Delta u_i + o(\Delta u_i)] \\
 &= \lim_{m \rightarrow \infty} \prod_{i=1}^m [1 + g(u_i)\Delta u_i]. \quad \square
 \end{aligned}$$

Corolario 5.5. *Sea $\log [1 + g(u_i)\Delta u_i] = g(u_i)\Delta u_i + o(\Delta u_i)$ para un Δu_i pequeño entonces para el caso continuo se tiene que*

$$\prod_{(a,b]} [1 + g(u)du] = \exp \left[\int_a^b g(u)du \right]. \quad (31)$$

Demostración. Aplicando logaritmo en (31) y el teorema 5.4 se tiene que

$$\begin{aligned}
 \log \prod_{(a,b]} [1 + g(u)du] &= \log \lim_{m \rightarrow \infty} \prod [1 + g(u_i)\Delta U_i] \\
 &= \log \lim_{m \rightarrow \infty} \sum_{i=1}^m [1 + g(u_i)\Delta U_i] \\
 &= \lim_{m \rightarrow \infty} \sum_{i=1}^m [g(u_i)\Delta U_i] \\
 &= \int_a^b g(u)du,
 \end{aligned}$$

entonces

$$\log \prod_{(a,b]} [1 + g(u)du] = \int_a^b g(u)du,$$

de esta manera

$$\prod_{(a,b]} [1 + g(u)du] = \exp \left[\int_a^b g(u)du \right]. \quad \square$$

Corolario 5.6. Si $G(u)$ tiene saltos en los puntos a_j con $j = 1, 2, \dots$, de tamaño g_j entonces

$$\prod_{(a,b]} [1 + dG(u)] = \prod_{(a,b]} [1 + g(u)du] \prod_{j:a < a_j \leq b} (1 + g_j) \quad (32)$$

Demostración. Se puede notar que

$$(1 + g(u)du)(1 + g_j) = 1 + g(u)du + g_j = 1 + dG(u).$$

Por lo que se da la igualdad siguiente

$$\prod_{(a,b]} [1 + dG(u)] = \prod_{(a,b]} [1 + g(u)du] \prod_{j:a < a_j \leq b} (1 + g_j),$$

de esta forma se tiene lo deseado. □

Recordemos que

$$h(u) = \frac{f(u)}{S(u)},$$

representa la función de riesgo para T en los puntos donde $F(u)$ y $S(u)$ son continuas y $h_j = P(T = a_j | t \geq a_j)$ en los valores de tiempo discreto con saltos en los puntos a_j .

Definición 5.7. La función de riesgo acumulativa se define mediante la integral de *Riemman-Stieltjes* como

$$H(t) = \int_0^t dH(u) = \int_0^t h(u)du + \sum_{j:a_j \leq t} h_j. \quad (33)$$

Dada la función de riesgo acumulativa, podemos obtener la función de supervivencia, para esto es fundamental tomar $0 = u_0 < u_1 < \dots < u_m = t$ y definir

$$P(T \geq t) = \prod_{i=1}^m P(T \geq u_i | T \geq u_{i-1}). \quad (34)$$

Ahora, $\Delta u_i = u_i - u_{i-1}$ y para un intervalo $[u_{i-1}, u_i)$ suficientemente pequeño donde puede haber a lo más un salto, se tiene que

$$\begin{aligned} P(T \geq u_i | T \geq u_{i-1}) &= 1 - \frac{P(u_{i-1} \leq T \leq u_i)}{P[T \geq u_{i-1}]} \\ &= 1 - [H(u_i) - H(u_{i-1}-)] + o(\Delta u_i). \end{aligned}$$

Por lo tanto,

$$P(T \geq t) = \prod_{[0,t)} [1 - dH(u)]. \quad (35)$$

Que es la integral producto sobre el intervalo abierto $(0, t)$.

Definición 5.8. Sea $S(t) = P(T \geq t)$, continua por la izquierda y $H(u)$ es continua por la derecha, definamos

$$P(T \geq t) = \exp \left[- \int_0^t h(u) du \right] \prod (1 - h_j),$$

como la función de supervivencia para T .

Observación 5.9. Para $a < b$ tenemos que la función de supervivencia queda denotada por

$$P(T \geq b | T \geq a) = \prod_{[a,b)} [1 - dH(u)].$$

Teorema 5.10. Sea $a_0 < a_1 < \dots < a_m$ valores específicos con $a_0 = 0$ y $a_m = \infty$ y con función de riesgo de la forma

$$h(t) = \lambda_j, \quad (36)$$

donde $a_{j-1} \leq t < a_j$ y $\lambda_j > 0$ para $j = 1, 2, \dots, m$ entonces T tiene una función de riesgo constante por tramos.

Ver detalles de la demostración en [5].

Las funciones de riesgo constantes por tramos y su función de densidad de probabilidad, los hace poco atractivos en muchos entornos, en otros casos se obtiene una función de riesgo más flexible y esto se puede obtener mediante el uso de *Spline*, que consiste de polinomios constante por tramos con cortes en los puntos a_1, \dots, a_{m-1} .

Un *Spline cúbico* $g(t)$ en el intervalo (a_0, a_m) consiste de un polinomio cúbico por tramos que es asignado a $g(t)$ donde las primeras dos derivadas son continuas en todas partes de (a_0, a_m) , en particular ellas son continuas en los puntos de cambio a_1, \dots, a_{m-1} que hacen referencia a los nudos en la terminología de *Spline*, para más detalles se recomienda [1].

También se puede definir el *Spline* en otros ordenes (cuadráticos o lineales), pero el *Spline cúbico* es favorable ya que solo $m + 3$ parámetros son necesarios para

especificar un Spline cúbico con $m - 1$ nudos a_1, \dots, a_{m-1} . Una expresión de este *Spline* es la siguiente

$$g(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 + \sum_{j=1}^{m-1} \beta_j (t - a_j)_+^3, \quad (37)$$

donde $x_+ = \max(x, 0)$.

6 Generalización de censura por la derecha

La idea esencial para la aproximación general es considerar los procesos de falla y de censura sobre un tiempo determinado, de un grupo de individuos.

Supongamos que n individuos tienen un tiempo t de falla o de censura, se supone que los tiempos de vida y los de censura son discretos, por conveniencia y sin pérdida de generalidad admitimos que los valores para t están dados por $t = 0, 1, 2, \dots$, se introduce la notación para el análisis de los procesos de falla y de censura en el tiempo t .

- $Y_i(t) = I(T_i \geq t, \text{ el individuo } i \text{ no tiene censura después del tiempo } t)$
- $dN_i(t) = Y_i(t)I(T_i(t) = t)$
- $dC_i(t) = Y_i(t)I(\text{el individuo } i \text{ es censurado en un tiempo } t)$

La variable $Y_i(t)$ es llamada el indicador de riesgo.

Observación 6.1. El indicador de riesgo es 1 si y sólo si el individuo i está vivo y no es censurado justo después del tiempo t .

Las variables $dN_i(t)$ y $dC_i(t)$ registran las fallas observadas y los tiempos de censura respectivamente en el tiempo t .

Definamos los vectores

- $dN(t) = (dN_1(t), \dots, dN_n(t))$
- $dC(t) = (dC_1(t), \dots, dC_n(t))$
- $H(t) = \{(dN(s), dC(s)), s = 0, 1, \dots, t - 1\}$

$H(t)$ representa las fallas y las censuras hasta un instante antes del tiempo t .

Observación 6.2. Todos los datos observados pueden ser representados mediante

$$\text{Datos} = (dN(t), dC(t); t = 0, 1, 2, \dots).$$

Hasta este momento no se han hecho suposiciones de los mecanismos de censura, en este momento se supone que la censura será tomada en el análisis de tiempos de vida y se requiere que

$$P(dN(t)|H(t)) = \prod_{i=1}^n h_i(t)^{dN_i(t)} [1 - h_i(t)]^{Y_i(t)[1-dN_i(t)]}, \quad (38)$$

y $H(t)$, los valores de la covariable, y los mecanismos de falla para cada individuo en un tiempo t sean independientes para $t = 0, 1, 2, \dots$, es decir,

$$P(dN_i = 1|H(t)) = Y_i(t)h_i(t). \quad (39)$$

Se denota convencionalmente que $0^0 = 1$ y lo usamos en (38), asumiendo que $Y_i(t) = 0$, esto no da una información del individuo en el tiempo t , y el término se muestra en la verosimilitud, dado que el valor es igual a uno.

Observación 6.3. El valor de $Y_i(t)$ esta determinado por $H(t)$.

La ecuación (39) representa una probabilidad condicional haciendo notar la independencia de los tiempos de falla y los tiempos de censura, bajo (39) la probabilidad de que un individuo siga vivo o sea censurado justo en un tiempo t es $h_i(t)$.

* Si la falla ocurre en algún tiempo t está será denotada por

$$dN_i(t) = I(s \leq t).$$

* Si se censura en el tiempo t está dado por

$$dN_i(t) = 0, \quad Y_i(s) = I(s \leq t).$$

Haciendo uso de la Definición 2.11 y la expresión (14) se puede observar que

$$S_i(t) = \prod_{s=0}^{t-1} (1 - h_i(s)), \quad f_i(t) = h_i(t)S_i(t). \quad (40)$$

Por lo que tenemos, que la verosimilitud usando la notación (t_i, δ_i) es

$$L = \prod_{i=1}^n f_i(t_i)^{\delta_i} S_i(t_{i+1})^{1-\delta_i}. \quad (41)$$

Dado que $S_i(t_{i+1}) = S_i(t+)$, se tiene que la verosimilitud es exactamente de la forma que la ecuación (25) encontrada para la censura tipo I y los otros tipos de censura, citados anteriormente.

7 Modelo constante por tramos con múltiples puntos de cambio

El modelo paramétrico con un punto de cambio más simple y el que ocurre con más frecuencia en la literatura, es el que supone que en una muestra aleatoria de tamaño n , k sujetos tienen una función de riesgo común α_1 y los $(n - k)$ restantes, tienen otra función de riesgo común α_2 , con $\alpha_1 \neq \alpha_2$, este modelo se puede expresar de la forma siguiente

$$\lambda(t) = \begin{cases} \alpha_1, & 0 \leq t < \tau, \\ \alpha_2, & t \geq \tau. \end{cases} \quad (42)$$

El interés se centra en el modelo con múltiples puntos de cambio, y se describe de la siguiente manera:

Sean X_1, \dots, X_n que denotan los tiempos de vida o de supervivencia los cuales son independientes e idénticamente distribuidos y sean C_1, \dots, C_n los correspondientes tiempos de censura, estos son independientes de los tiempos de supervivencia. Se pondrá principal atención en las parejas (T_i, δ_i) , para $i = 1, 2, \dots, n$, ya que estas serán de vital importancia durante el análisis.

Definición 7.1. Sea T el tiempo de vida o de falla. Se dice que T tiene múltiples puntos de cambio si su función de riesgo $\lambda(t)$ está dada por

$$\lambda(t) = \begin{cases} \alpha_1, & 0 \leq t < \tau_1, \\ \alpha_2, & \tau_1 \leq t < \tau_2, \\ \vdots & \vdots \\ \alpha_{k+1}, & t \geq \tau_k, \end{cases} \quad (43)$$

donde $0 < \tau_1 < \dots < \tau_k$, k es el número de puntos de cambio en el modelo, y α_j es el valor que toma la función de riesgo entre τ_{j-1} y τ_j .

Una vez definida la función de riesgo por tramos con múltiples puntos de cambio, se obtendrá la función correspondiente de densidad de T .

Para poder obtener la función de densidad de T se hará uso de un resultado previo el cual indica que si se tiene la función de riesgo se puede determinar de manera única a la función de densidad, ver Corolario 2.2.

Utilizando dicho resultado se tiene que la función de densidad está dada por

$$f(t) = \begin{cases} \alpha_1 \exp[-\alpha_1 t], & 0 \leq t < \tau_1, \\ \alpha_2 \exp[-\alpha_1 \tau_1 - \alpha_2(t - \tau_1)], & \tau_1 \leq t < \tau_2, \\ \vdots & \vdots \\ \alpha_{k+1} \exp[-\alpha_1 \tau_1 - \alpha_2(\tau_2 - \tau_1) - \dots - \alpha_{k+1}(t - \tau_k)], & t \geq \tau_k; \end{cases} \quad (44)$$

es claro que esta es una función de densidad exponencial.

Se presenta un bosquejo de la construcción de esta función de densidad, solo se hará para los dos primeros casos y para el resto se obtienen de manera análoga

* Para $0 \leq t < \tau_1$

$$\begin{aligned} f_1(t) &= \lambda(t) \exp \left[- \int_0^t \lambda(u) du \right] = \alpha_1 \exp \left[- \int_0^t \alpha_1 du \right] \\ &= \alpha_1 \exp [-\alpha_1 t]. \end{aligned}$$

* Para $\tau_1 \leq t < \tau_2$

$$\begin{aligned} f_2(t) &= \lambda(t) \exp \left[- \int_0^t \lambda(u) du \right] = \alpha_2 \exp \left[- \int_0^{\tau_1} \alpha_1 du - \int_{\tau_1}^t \alpha_2 du \right] \\ &= \alpha_2 \exp [-\alpha_1 \tau_1 - \alpha_2 (t - \tau_1)]. \end{aligned}$$

De esta manera se va construyendo la función de densidad $f(t)$ para cada punto de cambio.

El número de fallas observadas hasta el instante de tiempo T estará dado por

$$X(t) = \sum_{i=1}^n I(T_i < t) \delta_i,$$

donde δ_i es la función indicadora de censura.

Observación 7.2. Si $X(t)$ denota el número de fallas hasta tiempo t entonces se tiene que

$$X(\tau_j) = \sum_{i=1}^n I(T_i < \tau_j) \delta_i$$

va a denotar el número de fallas observadas hasta el punto de cambio τ_j .

Ahora, la situación es similar a un problema de análisis secuencial, de esta forma se realizará una prueba de hipótesis y si se rechaza la hipótesis nula, se continúa con la siguiente prueba de hipótesis. Si no se rechaza la hipótesis nula, el proceso se detiene y se concluye que se ha encontrado el modelo final.

La prueba de hipótesis que se tiene que llevar a cabo es la siguiente

$$H_0 : \alpha_{k-1} = \alpha_k \quad vs. \quad H_1 : \alpha_{k-1} \neq \alpha_k$$

Esta es equivalente a probar la hipótesis de que $\tau_{k-1} = 0$ para $k = 2, \dots, K$.

Extendiendo la metodología existente, se desarrolla una prueba de razón de verosimilitud, la cual será de ayuda para poder realizar la prueba de hipótesis.

El primer objetivo será encontrar la función de densidad f_0 , y así poder aplicar (46).

La función de densidad para el número de fallas hasta un determinado tiempo τ_j , está dada de la siguiente manera

$$f_0(X(\tau_j)) = \begin{cases} \alpha_1 \exp [X(\tau_1)], & 0 \leq X(\tau_j) < \tau_1 \\ \alpha_2, \exp [X(\tau_2) - X(\tau_1)], & \tau_1 \leq X(\tau_j) < \tau_2, \\ \vdots & \vdots \\ \alpha_{k+1} \exp [n_u - X(\tau_{k-1})], & X(\tau_j) \geq \tau_k. \end{cases} \quad (45)$$

El tipo de censura que se va a utilizar en esta situación es censura por la derecha, por lo cual definiremos a $f(t_1) = (T_i \wedge \tau_1)$ y $s_1(t_i) = I(T_i > \tau_1)$, puesto que estos son los tiempos de vida y de censura hasta un determinado tiempo τ_1 , de esta forma tenemos que la función de densidad está dada por

$$f_1(\tau_j) = \exp \left(- \alpha_i \sum_{i=1}^n f(t_1) s_1(t_i) \right).$$

Una vez establecidas $f_1(\tau_j)$ y $f_0(X(\tau_j))$, se procede a calcular la *log*-verosimilitud de $f(t)$.

$$L_k = \sum_{j=1}^k \ln \frac{f_1(\tau_j)}{f_0(X(\tau_j))}. \quad (46)$$

Dado que la función de densidad de T está definida por tramos, se llevará acabo el proceso de máxima verosimilitud para cada tramo.

$$\begin{aligned} l_1 &= \frac{e^{(-\alpha_1 \sum_{i=1}^n (T_i \wedge \tau_1))}}{\alpha_1 e^{(-X(\tau_1))}}, \\ l_2 &= \frac{e^{(-\alpha_2 \sum_{i=1}^n (T_i \wedge \tau_2 - \tau_1) I(T_i > \tau_1))}}{\alpha_1 e^{-(X(\tau_2) - X(\tau_1))}}, \\ &\vdots \\ l_k &= \frac{e^{(-\alpha_k \sum_{i=1}^n (T_i \wedge \tau_{k-1}) I(T_i > \tau_{k-1}))}}{\alpha_k e^{-(n_u - X(\tau_{k-1}))}}. \end{aligned}$$

Donde n_u es el número de datos no censurados.

Por otro lado, se tiene que

$$L_k = \sum_{j=1}^k \ln \frac{f_1(x_j)}{f_0(x_j)} = \ln l_1 + \ln l_2 + \dots + \ln l_k,$$

luego se procede a calcular, $\ln L_1 + \ln L_2 + \dots + \ln L_k$, los cuales resultan ser

$$\begin{aligned} \ln l_1 &= (-\alpha_1 \sum_{i=1}^n (T_i \wedge \tau_1)) - (-X(\tau_1) \ln \alpha_1) \\ &= X(\tau_1) \ln \alpha_1 - \alpha_1 \sum_{i=1}^n (T_i \wedge \tau_1), \\ \ln l_2 &= (-\alpha_2 \sum_{i=1}^n (T_i \wedge \tau_2 - \tau_1) I(T_i > \tau_1)) - (-(X(\tau_2) - X(\tau_1)) \ln \alpha_2) \\ &= (X(\tau_2) - X(\tau_1)) \ln \alpha_2 - \alpha_2 \sum_{i=1}^n (T_i \wedge \tau_2 - \tau_1) I(T_i > \tau_1), \\ &\vdots \\ \ln l_k &= [-\alpha_k \sum_{i=1}^n (T_i \wedge \tau_{k-1}) I(T_i > \tau_{k-1})] - [-(n_u - X(\tau_{k-1})) \ln \alpha_{k-1}] \\ &= (n_u - X(\tau_{k-1})) \ln \alpha_{k-1} - \alpha_k \sum_{i=1}^n (T_i \wedge \tau_{k-1}) I(T_i > \tau_{k-1}). \end{aligned}$$

Por lo que

$$\begin{aligned} L_k &= \sum_{j=1}^k \ln \frac{f_1}{f_0} \\ &= X(\tau_1) \ln \alpha_1 + [X(\tau_2) - X(\tau_1)] \ln \alpha_2 + \dots + \\ &\quad [n_u - X(\tau_{k-1})] \ln \alpha_{k-1} - \alpha_1 \sum_{i=1}^n (T_i \wedge \tau_1) - \\ &\quad \alpha_2 \sum_{i=1}^n (T_i \wedge \tau_2 - \tau_1) I(T_i > \tau_1) - \dots - \alpha_k \sum_{i=1}^n (T_i \wedge \tau_{k-1}) I(T_i > \tau_{k-1}). \end{aligned} \tag{47}$$

Una vez que es obtenida la función *Log*-verosimilitud, se procede a calcular los estimadores de máxima verosimilitud, los cuales se pueden encontrar al derivar e

igualar a cero, la función antes descrita, dando como resultado

$$\begin{aligned} \hat{\alpha}_1 &= \frac{X(\tau_1)}{\sum_{i=1}^n (T_i \wedge \tau_1)}, \\ \hat{\alpha}_2 &= \frac{X(\tau_2) - X(\tau_1)}{\sum_{i=1}^n (T_i \wedge \tau_2 - \tau_1) I(T_i > \tau_1)}, \\ &\vdots \\ \hat{\alpha}_k &= \frac{n_u - X(\tau_{k-1})}{\sum_{i=1}^n (T_i \wedge \tau_{k-1}) I(T_i > \tau_{k-1})}. \end{aligned} \tag{48}$$

Ahora se procede a ver que la función de verosimilitud tiene una distribución Chi-cuadrada, para esto, vamos a sustituir los estimadores de máxima verosimilitud de los parámetros de la siguiente forma:

Cuando se tienen k puntos de cambio la función resultaría de la siguiente forma

$$\begin{aligned} L_1(\tau_1, \dots, \tau_{k-1}) &= X(\tau_1) \ln \frac{X(\tau_1)}{\sum_{i=1}^n (T_i \wedge \tau_1)} + \\ & (X(\tau_2) - X(\tau_1)) \ln \frac{X(\tau_2) - X(\tau_1)}{\sum_{i=1}^n (T_i \wedge \tau_2 - \tau_1) I(T_i > \tau_1)} + \\ & \vdots \\ & (n_u - X(\tau_{k-1})) \ln \frac{n_u - X(\tau_{k-1})}{\sum_{i=1}^n (T_i \wedge \tau_{k-1}) I(T_i > \tau_{k-1})} - n_u. \end{aligned} \tag{49}$$

Considerando a $\alpha_{k-1} = \alpha_k$ y los $k - 1$ puntos de cambio se tiene entonces que la función se puede expresar

$$\begin{aligned} L_2(\tau_1, \dots, \tau_{k-1}) &= X(\tau_1) \ln \frac{X(\tau_1)}{\sum_{i=1}^n (T_i \wedge \tau_1)} + \\ & (X(\tau_2) - X(\tau_1)) \ln \frac{X(\tau_2) - X(\tau_1)}{\sum_{i=1}^n (T_i \wedge \tau_2 - \tau_1) I(T_i > \tau_1)} + \\ & \vdots \\ & (n_u - X(\tau_{k-2})) \ln \frac{n_u - X(\tau_{k-2})}{\sum_{i=1}^n (T_i \wedge \tau_{k-2}) I(T_i > \tau_{k-2})} - n_u. \end{aligned}$$

Por lo que nuestra razón de verosimilitud queda expresada mediante

$$\begin{aligned} & L_1(\tau_1, \dots, \tau_{k-1}) - L_2(\tau_1, \dots, \tau_{k-1}) \\ &= [X(\tau_{k-1}) - X(\tau_{k-2})] \ln \hat{\alpha}_{k-1} + [n_u - X(\tau_{k-1})] \ln \hat{\alpha}_k - [n_u - X(\tau_{k-2})] \ln \hat{\alpha}_{k-1}. \end{aligned}$$

Dado que el modelo reducido tiene dos parámetros menos que el modelo completo, la prueba de razón de verosimilitud asintótica nos lleva a concluir que la función de verosimilitud tiene una distribución χ_2^2 , ver [4].

7.1 Prueba de tipo Wald para el modelo constante por tramos

El interés esta basado en las pruebas de los puntos de cambio para el siguiente modelo

$$\lambda(t) = \begin{cases} \alpha_1, & 0 \leq t < \tau_1, \\ \alpha_2, & \tau_1 \leq t < \tau_2, \\ \vdots & \vdots \\ \alpha_{k+1}, & t \geq \tau_k, \end{cases}$$

y una prueba equivalente sería analizar si la falla se encuentra antes o después de estos puntos de cambio.

Para probar la hipótesis nula de que no hay puntos de cambio contra la alternativa de que hay un punto de cambio, se tiene que probar que

$$H_0 : \alpha_1 = \alpha_2 \quad \text{o} \quad H_0 : \alpha_1 - \alpha_2 = 0,$$

para lo cual se propone el uso de una prueba de tipo Wald.

Sea $\theta' = [\alpha_1, \alpha_2, \dots, \alpha_{K+1}, \tau_1, \tau_2, \dots, \tau_K]$ la prueba de la hipótesis nula $H_0 : C'\theta = M$, donde C' es una matriz de $s \times p$, $s \leq p$ y M es un vector solución de $s \times 1$, ver [4].

La prueba de tipo Wald de la hipótesis H_0 es

$$\chi_w^2 = (C'\hat{\theta} - M)' [C' \sum_{\hat{\theta}} C]^{-1} (C'\hat{\theta} - M) \sim \chi_s^2.$$

Aunque el método propuesto es un procedimiento para múltiples puntos de cambio, este es un proceso paso a paso, por lo que se probará una hipótesis a la vez,

$$H_0 : \alpha_1 - \alpha_2 = 0 \quad \text{vs.} \quad H_1 : \alpha_1 - \alpha_2 \neq 0.$$

Para esto utilizo el estadístico de Wald de la forma

$$X_W = \frac{(\alpha_{k-1}^{\hat{}} - \alpha_k^{\hat{}})^2}{Var((\alpha_{k-1}^{\hat{}} - \alpha_k^{\hat{}}))} \sim \chi_1^2. \tag{50}$$

8 Conclusiones

El análisis de supervivencia es importante para el estudio de tiempos de vida o de falla, ya que proporciona información sobre el riesgo de falla existente en la población bajo estudio. Por lo que es conveniente obtener información del riesgo en determinados intervalos de riesgo.

El trabajo presenta una metodología para encontrar y estimar los puntos de cambio para el modelo constante por tramos.

Un estudio posterior es sobre el modelo lineal por tramos con múltiples puntos de cambio que incluye covariables, en cual se pretende realizar una simulación.

Bibliografía

- [1] De Boor Carl., *A Practical Guide to splines, Revised edition*, Springer (2001).
- [2] Hogg R., *Introduction to mathematical Statistics*, Pearson (2005).
- [3] Kleinbaum D., *Survival Analysis*, Springer (2005).
- [4] Lanchin J., *Biostatistical Methods* Wiley, (2000).
- [5] Lawless J., *Statistical Models and Methods for Lifetime Data*, Wiley, (2003).
- [6] Melody S., *Survival Analysis whith Cahnge Point Hazard Functions*, Revista Harvard University Biostatistics, Paper 40 (2006).
- [7] Mendoza M., *Importancia del análisis estadística secuencial*, Comunicación interna, (1978).
- [8] Meyer R., *Probabilidad y aplicaciones estadísticas*, Wesley, (1970).
- [9] Mood A., *Introduction to theory of statistics*, McGraw, (1974).
- [10] Palmeros O., *Algunos modelos en el análisis de supervivencia y una aplicación del modelo Weibull*, (2007).

Facultad de Ciencias Físico Matemáticas, BUAP
Avenida San Claudio y 18 Sur, Colonia San Manuel,
Puebla, Pue. C.P. 72570
212470815@fcfm.buap.mx
hcs@fcfm.buap.mx
hreyes@fcfm.buap.mx
ftajonar@fcfm.buap.mx

Capítulo 7

El modelo de Bandidos Armados, Programación Dinámica y la incorporación de nuevo personal

Hugo Adán Cruz Suárez, Víctor Hugo Vázquez Guevara
FCFM-BUAP

Resumen

Se presenta un problema de incorporación de personal en donde k nuevos colaboradores con una tasa de fallos en sus servicios desconocida son contratados en una tienda. El objetivo es maximizar el número de clientes satisfechos asignando la atención de clientes solo a uno de ellos. Este problema se abordará utilizando el modelos de bandidos armados y la técnica de programación dinámica

1 Introducción

Supongamos que una empresa que presta servicios al público se ve en la necesidad de contratar a k nuevos empleado de los cuales se desconoce sus habilidades.

Cada vez que un cliente solicita los servicios de la empresa, el administrador debe asignarle su atención a alguno de los nuevos empleados. Si el administrador conociera la proporción de clientes atendidos satisfactoriamente por parte de cada uno de ellos, solo tendría que compararlas y así asignarle la atención del cliente a aquel que sea más eficiente. Supondremos además que los clientes arriban a la empresa uno a la vez.

Si bien la eficiencia de los nuevos empleados es desconocida, supondremos que, como resultado del análisis de su currículum, las entrevistas y el periodo de prueba al que fueron sujetos antes de su contratación, se les asocia una distribución de probabilidad.

El administrador necesita entonces una estrategia de asignación de clientes basada en:

1. Las distribuciones de probabilidad asignadas a los empleados, y
2. Los desempeños observados.

Un enfoque para obtener dicha estrategia es utilizar el modelo de bandidos armados y la técnica de Programación Dinámica, mismos que en la siguiente sección son descritos.

El objetivo del administrador es diseñar una estrategia de asignación para maximizar el número de clientes atendidos satisfactoriamente de entre una cantidad fija de ellos.

2 El Modelo de Bandidos Armados

2.1 Introducción

Un Modelo de Dos Bandidos Armados en Teoría estadística de la decisión involucra selecciones secuenciales de entre k procesos estocásticos (armas, máquinas, tratamientos, etc.) en el que el tiempo puede considerarse discreto o continuo al igual que los procesos involucrados, estos procesos estocásticos son caracterizados por parámetros que normalmente son desconocidos. El proceso elegido para ser observado en cualquier instante depende de las elecciones anteriores y de los resultados observados. Un procedimiento de decisión (o estrategia) especifica qué proceso seleccionar en cualquier instante para cada historia de selecciones y observaciones previas.

En este Capítulo consideraremos el tiempo discreto, tenemos así que cada proceso genera una sucesión de variables aleatorias. Una observación en una sucesión en particular es hecha al seleccionar el correspondiente proceso. El objetivo clásico es maximizar el valor esperado de alguna ganancia $\sum_{m=1}^{\infty} \alpha_m Z_m$ en donde Z_m es la variable observada en la etapa m y α_m son números no negativos conocidos. Una estrategia es óptima si nos conduce a la ganancia esperada máxima.

2.2 Planteamiento del modelo

Denotemos por \mathcal{D} al espacio de distribuciones de probabilidad en \mathbb{R} con la topología de la convergencia en distribución. Cabe mencionar que este espacio es: separable, localmente compacto, metrizable y completo [2].

El espacio \mathcal{D}^k de vectores ordenados de dimensión k de miembros de \mathcal{D} se considerará con la correspondiente topología producto. La componente i -ésima de $(Q_1, Q_2, \dots, Q_k) \in \mathcal{D}^k$ gobierna al arma i . En el contexto del problema de incorporación de personal, las distribuciones Q_i son de Bernoulli con parámetros desconocidos, así su distribución es aleatoria. Y es por esto que el espacio $\mathcal{D}(\mathcal{D}^k)$ también será relevante en este trabajo. Un elemento $G \in \mathcal{D}(\mathcal{D}^k)$ representa la información conocida *a priori* por parte del agente encargado de la toma de decisiones concierne a las k armas. En el contexto señalado en la sección 1 el elemento $G \in \mathcal{D}(\mathcal{D}^k)$

está soportado por el subconjunto (distribuciones de Bernoulli) \times (distribuciones de Bernoulli) $\times \cdots \times$ (distribuciones de Bernoulli).

Espacio de Probabilidad

Sea Ω el espacio producto obtenido de \mathcal{D}^k e infinitas copias del intervalo unitario abierto; una para cada pareja (i, m) con $1 \leq i \leq k$ y $m = 1, 2, \dots$, esto es

$$\Omega = \mathcal{D}^k \times X_{i=1}^k X_{m=1}^\infty(0, 1). \quad (1)$$

La medida de probabilidad \mathbb{P} en Ω es el producto de G en \mathcal{D}^k y la medida de Lebesgue en cada intervalo unitario y está definida en la σ -álgebra producto \mathcal{F} de las σ -álgebras de Borel correspondientes.

Un elemento $\omega \in \Omega$ puede ser escrito en la forma

$$\omega = \{Q_i, 1 \leq i \leq k, \omega_{im}, 1 \leq i \leq k, m = 1, 2, \dots\} \quad (2)$$

en donde cada $Q_i \in \mathcal{D}$ y $\omega_{im} \in (0, 1)$. Si denotamos por Q_i^{-1} a la función inversa por la derecha de Q_i considerada como una función de distribución acumulada, definamos

$$X_{im}(\omega) = Q_i^{-1}(\omega_{im}) \quad (3)$$

para ω dada como en (2). Puede probarse que cada X_{im} es una variable aleatoria y que condicionando sobre (Q_1, \dots, Q_k) ,

$$\{X_{im} : 1 \leq i \leq k, m = 1, 2, \dots\} \quad (4)$$

es una familia independiente en donde la distribución condicional de X_{im} es Q_i [1]. Así, el valor de X_{im} es el resultado del arma i en la etapa m (se observe o no a X_{im}).

Estrategias

Una estrategia τ asigna a cada historia parcial de observaciones un entero entre 1 y k el cuál indica el arma que será seleccionada en la etapa siguiente. Así, $\tau(\phi)$ indica el arma que será seleccionada inicialmente cuando se sigue la estrategia τ , $\tau(z_1)$ indica el arma que se seleccionará en la etapa dos dado que z_1 fue observado en la etapa 1, $\tau(z_1, z_2)$ es el arma que se seleccionará en la etapa tres dado que en la etapa uno se observó z_1 y en la etapa 2 a z_2 , etc. Para que τ sea una estrategia requeriremos que el conjunto de observaciones para las cuales se elige el arma i en la etapa m ,

$$\{(z_1, z_2, \dots, z_{m-1}) : \tau(z_1, \dots, z_{m-1}) = i\} \quad (5)$$

sea un conjunto de Borel de \mathbb{R}^{m-1} .

Tenemos entonces que la sucesión de variables aleatorias observadas está definida recursivamente por

$$Z_1 = X_{\tau(\phi),1}, \quad (6)$$

$$Z_m = X_{\tau(Z_1, \dots, Z_{m-1}),m}, \quad m > 1; \quad (7)$$

el hecho de que cada Z_m es una variable aleatoria puede probarse por inducción [1].

Dado que el problema es elegir τ que maximice el valor esperado de

$$\sum_{k=1}^{\infty} \alpha_k Z_k \quad (8)$$

en donde $\mathbf{A} = (\alpha_1, \alpha_2 \dots)$ es llamada sucesión de descuento y es tal que cada $\alpha_m \geq 0$ y $\sum \alpha_m < \infty$. También se requerirá que cada componente Q_i de $(Q_1, Q_2, \dots, Q_k) \in \mathcal{D}^k$ tenga primer momento absoluto finito con probabilidad uno (con respecto a G) y que este momento a su vez tenga esperanza finita (con respecto a G). Denotaremos por $\mathcal{D}^*(\mathcal{D}^k)$ al subespacio de $\mathcal{D}(\mathcal{D}^k)$ de aquellas distribuciones que satisfacen esta condición, misma que puede expresarse como

$$\mathbb{E}[|X_{i1}|] = \mathbb{E}[\mathbb{E}(|X_{i1}| | Q_i)] = \int_{\mathbb{D}^k} \int_{\mathbb{R}} |x| Q_i(dx) G(d(Q_1, \dots, Q_k)) \quad (9)$$

$$= \int_{\mathbb{D}} \left[\int_{\mathbb{R}} |x| Q_i(dx) \right] F_i(d(Q_i)) \quad (10)$$

en donde F_i es la distribución de Q_i ; esto es, la i -ésima distribución marginal de G . Tenemos además que

$$|Z_m| \leq \sqrt[k]{\sum_{i=1}^k |X_i|} \leq \sum_{i=1}^k |X_i| \quad (11)$$

y por lo tanto Z_m es integrable y para cualquier estrategia τ

$$\left| \mathbb{E}_{\tau} \left[\sum_{m=1}^{\infty} \alpha_m Z_m \right] \right| \leq \left[\sum_{i=1}^k |X_i| \right] \sum_{m=1}^{\infty} \alpha_m < \infty \quad (12)$$

en donde el subíndice τ indica la dependencia de la esperanza sobre la estrategia τ .

A $\mathbb{E}_{\tau} [\sum_{m=1}^{\infty} \alpha_m Z_m]$ se le llama el *valor* de la estrategia τ y será denotado por $W(G, \mathbf{A}, \tau)$. Así, tenemos la siguiente definición.

Definición 2.1. El valor del (G, \mathbf{A}) -bandido está dado por

$$V(G, \mathbf{A}) = \sup_{\tau} W(G, \mathbf{A}, \tau) = \sup_{\tau} \mathbb{E}_{\tau} \left[\sum_{m=1}^{\infty} \alpha_m Z_m \right] \quad (13)$$

Una estrategia τ para la cual $V(G, \mathbf{A})$ es alcanzado se llama *óptima*

En la primera etapa de cualquier problema de bandidos armados, el administrador es enfrentado con una distribución inicial y una sucesión de descuento. En la segunda etapa (después de la selección inicial y su correspondiente observación) el administrador se enfrenta con una nueva distribución y una nueva sucesión de descuento; esto es, con un nuevo bandido armado. Así, la elección hecha en la segunda etapa puede verse como una elección inicial en este nuevo bandido armado. Un enfoque similar puede aplicarse en las demás etapas futuras. Es por esto que, si se conociera una decisión inicial óptima para cada bandido entonces se conocería una estrategia óptima para cada bandido. Este hecho le confiere a la elección inicial una relevancia especial. Sin embargo, como veremos más adelante, una elección inicial óptima no puede hacerse sin considerar las elecciones futuras.

Particionando el conjunto de estrategias de acuerdo al arma seleccionada inicialmente, tenemos que

$$V(G, A) = \bigvee_{i=1}^k \sup_{\tau(\phi)=i} W(G, A, \tau) \quad (14)$$

$$= \bigvee_{i=1}^k \left[\alpha_1 \mathbb{E}(X_{i1}) + \sup_{\tau(\phi)=i} \mathbb{E}_\tau \left(\sum_{m=2}^{\infty} \alpha_m Z_m \right) \right]. \quad (15)$$

Introduciremos ahora algunas notaciones que nos ayudarán con el estudio de (14) y que reflejan el hecho de que el administrador se enfrenta con un nuevo bandido armado en cada etapa.

Para la sucesión de descuento $A = (\alpha_1, \alpha_2, \dots)$ denotamos por $A^{(m)}$ a la sucesión $(\alpha_{m+1}, \alpha_{m+2}, \dots)$. La sucesión de descuento $A^{(1)}$ y la distribución *a posteriori* de (Q_1, \dots, Q_k) después de la etapa uno caracterizan al bandido con el que tendrá que trabajar el administrador en la etapa 2. Para indicar la dependencia que esta distribución *a posteriori* tiene sobre la observación hecha en la etapa uno, usaremos el símbolo $(x)_i G$ para denotar a la distribución condicional de (Q_1, \dots, Q_k) dada la observación x en el arma i .

En el caso en que $G = F_1 \times \dots \times F_k$ tenemos que:

$$(x)_i G = (x)_i (F_1 \times \dots \times F_k) = (F_1 \times \dots (x)F_i \times \dots \times F_k) \quad (16)$$

en donde $(x)F_i$ denota la distribución condicional de Q_i dada la observación x en el arma i .

Describiremos a continuación dos formas en que un problema de bandidos armados puede verse, para esto supongamos que el administrador tiene un asistente el cual llevará a cabo sus decisiones:

1. El administrador le proporciona a su ayudante una estrategia τ , así este evaluará $\tau(\phi)$ y elegirá el arma correspondiente y posteriormente observará un

número; digamos, z_1 , en la etapa siguiente evaluará $\tau(z_1)$, seleccionará el arma que la estrategia indique, observará z_2 , evaluará $\tau(z_1, z_2)$, etcétera.

2. En la segunda forma, el administrador debe considerar todos los pares de distribuciones en $\mathcal{D}^*(\mathcal{D}^k)$ y sucesiones de descuento; (G, A) y especificar $\tau_{G,A}(\phi)$. Cuando el administrador se enfrenta con una elección particular de (G, A) el ayudante selecciona el arma inicial de acuerdo a $\tau_{G,A}(\phi)$ y observa el número z_1 . El ayudante calcula entonces a $A^{(1)}$ y a $(z_1)_j G$ con $j = \tau_{G,A}(\phi)$. En la etapa dos el ayudante elige $\tau_{(z_1)_j G, A^{(1)}}(\phi)$ y observa z_2 , posteriormente calcula $A^{(2)}$ y a $(z_2)_i (z_1)_j G$ para a continuación seleccionar el arma $\tau_{(z_2)_i (z_1)_j G, A^{(2)}}(\phi)$, etcétera.

Nota: Cuando hablamos de un (G, A) -bandido suponemos implícitamente que $G \in \mathcal{D}^*(\mathcal{D}^k)$. Además,

$$\mathbb{E} [|X_{j2}| | G] = \mathbb{E}_\tau [\mathbb{E} [|X_{j1}| | (z_1)_i G]]. \quad (17)$$

lo cual implica que $(z_1)_i G \in \mathcal{D}^*(\mathcal{D}^k)$ con probabilidad uno y por tanto tiene sentido hablar del $((z_1)_i G, A)$ -bandido.

3 Programación Dinámica para horizontes finitos

El objetivo de esta sección es demostrar la existencia de una estrategia óptima para cada bandido con un número finito de etapas, es por esto que consideraremos sucesiones de descuento que tienen solo un número finito de elementos que no son cero.

Definición 3.1. El horizonte de una sucesión de descuento $A = (\alpha_1, \alpha_2, \dots)$ de un bandido armado se define como:

$$\inf \{n : \alpha_m = 0 \text{ para } m > n\}.$$

El horizonte de un bandido armado puede ser infinito y también puede ser igual a cero si $\alpha_m = 0$ para toda m .

Así como a los espacios \mathcal{D} y $\mathcal{D}(\mathcal{D}^k)$ se les considera con una topología, haremos lo correspondiente con el conjunto de sucesiones de descuento.

Denotemos por \mathcal{A} al conjunto de todas las sucesiones de descuento y consideremos la siguiente norma para $A = (\alpha_1, \alpha_2, \dots) \in \mathcal{A}$

$$|A|_1 = \sum_{m=1}^{\infty} \alpha_m \quad (18)$$

y la distancia entre A y $B = (\beta_1, \beta_2, \dots)$ como

$$|A - B|_1 = \sum_{m=1}^{\infty} |\alpha_m - \beta_m| \quad (19)$$

así, la topología considerada será la proveniente de esta métrica.

El siguiente teorema establece el objetivo de esta sección. Existe más de una prueba de la existencia de estrategias óptimas en Procesos de Decisión de Márkov con horizonte finito; por ejemplo [3]. Sin embargo, abordaremos una demostración propia de los problemas de bandidos armados.

Teorema 3.2. *Existe una estrategia $\tau_{G,A}$ para cada $A \in \mathcal{A}$ con horizonte finito y cada $G \in \mathcal{D}^*(\mathcal{D}^k)$ tal que*

$$\{(G, A, z_1, \dots, z_{m-1}) : \tau_{(G,A)}(z_1, \dots, z_{m-1}) = i\} \quad (20)$$

es un conjunto medible de $\mathcal{D}^*(\mathcal{D}^k) \times \mathcal{A} \times (-\infty, \infty)^{m-1}$ para cada $i = 1, 2, \dots, k$ y $m = 1, 2, \dots$ y $\tau_{G,A}$ para cada G y A que es óptima para el (G, A) -bandido. Al restringirse a las sucesiones de descuento con horizonte finito, la función

$$(G, A) \mapsto V(G, A) \quad (21)$$

es medible y satisface

$$V(G, A) = \bigvee_{i=1}^k \mathbb{E} [\alpha_i X_{i1} + V((X_{i1})_i G, A^{(1)} | G)] \quad (22)$$

Demostración. El conjunto de sucesiones de descuento que tienen un horizonte particular (digamos n) es un conjunto medible de \mathcal{A} . Para ver esto podemos considerar la función

$$f : \mathcal{A} \rightarrow \mathbb{R} \quad (23)$$

tal que $\forall A \in \mathcal{A}$, $f(A) = |A|_1$ y notar que es una función continua ya que

$$\left| \sum_{m=1}^{\infty} \alpha_m - \sum_{m=1}^{\infty} \beta_m \right| \leq \sum_{m=1}^{\infty} |\alpha_m - \beta_m| \quad (24)$$

Así, podemos hacer la prueba de este teorema por inducción sobre el horizonte de la sucesión de descuento.

La única sucesión con horizonte cero es $\mathbf{0} = (0, 0, \dots)$. Claramente, tenemos que $V(G, \mathbf{0}) = 0$ para cada G y además todas las estrategias son óptimas. Por ejemplo,

$$\tau_{(G,\mathbf{0})}(z_1, \dots, z_m) = 1 \quad (25)$$

para cada $m = 1, 2, \dots$ y z_1, z_2, \dots, z_m . Así, la ecuación (22) se satisface con $A = \mathbf{0}$.

Supongamos ahora que $\tau_{G,A}$ ha sido definida y que la conclusión del teorema es verdadera para todas las sucesiones de descuento $A \in \mathcal{A}$ con horizonte menor que n . Procedamos a definir una estrategia $\tau_{G,A}$ para $A \in \mathcal{A}$ con horizonte igual a n .

Definamos una estrategia $\tau_{G,A}^i$ que elige el arma i inicialmente y que procede de manera óptima:

$$\tau_{G,A}^i(\phi) = i \quad (26)$$

$$\tau_{G,A}^i(z_1, \dots, z_{m-1}) = \tau_{(z_1)_i G, A^{(1)}(z_2, \dots, z_{m-1})} \text{ , si } m > 1. \quad (27)$$

Tenemos que la función $A \mapsto A^{(1)}$ es continua (y por tanto medible) ya que

$$\sum_{m=2}^{\infty} |\alpha_m - \beta_m| \leq \sum_{m=1}^{\infty} |\alpha_m - \beta_m| \quad (28)$$

además, puede probarse que la función $(z_1, G) \mapsto (z_1)_i G$ es medible [1] así, la hipótesis de inducción implica la medibilidad del conjunto

$$\{(G, A, z_1, \dots, z_{m-1}) : \tau_{(G,A)}^i(z_1, \dots, z_{m-1}) = j\} \quad (29)$$

para cada i y m .

Para probar (22) consideremos el lado derecho de (14):

$$\sup_{\tau(\phi)=i} \mathbb{E}_{\tau} \left[\sum_{m=2}^{\infty} \alpha_m Z_m | G \right] = \sup_{\tau(\phi)=i} \mathbb{E}_{\tau} \left(\mathbb{E}_{\tau} \left[\sum_{m=2}^{\infty} \alpha_m Z_m | Xim \right] | G \right) \quad (30)$$

$$\leq \mathbb{E}_{\tau} \left(\sup_{\tau(\phi)=i} \left(\mathbb{E}_{\tau} \left[\sum_{m=2}^{\infty} \alpha_m Z_m | Xim \right] | G \right) \right). \quad (31)$$

Puede observarse que cada supremo del lado derecho es alcanzado por $\tau_{(G,A)}^i$ y que el supremo correspondiente es igual a $V((X_{i1})_i G, A^{(1)})$ el cuál por la hipótesis de inducción y la medibilidad de $(x, G) \mapsto (x)_i G$ es una función medible de (ω, G, A) . Así (22) se sigue de (14) y a su vez, de (22) y la hipótesis de inducción tenemos la medibilidad de $(G, A) \mapsto V(G, A)$.

Finalmente, completamos la prueba al definir a $\tau_{G,A}$ como a aquella $\tau_{G,A}^i$ con i el entero mas pequeño para el que el máximo en (22) es alcanzado. \square

Para propósitos computacionales, reescribiremos a (22) como:

$$V(G, A) = \bigvee_{i=1}^k \int_{\mathcal{D}^k} \int_{\mathbb{R}} [\alpha_1 x + V((x)_i G, A)] Q_i(dx) G(d(Q_1, \dots, Q_k)) \quad (32)$$

$$= \bigvee_{i=1}^k \int_{\mathcal{D}} \int_{\mathbb{R}} [\alpha_1 x + V((x)_i G, A)] Q_i(dx) F_i(dQ_i) \quad (33)$$

Veamos cómo utilizar esta última expresión para hallar estrategias óptimas. Pero primero introduzcamos una notación que hará más simple la escritura de la metodología correspondiente.

Denotemos por $G^{(m)}$ a la distribución condicional en \mathcal{D}^k después de la etapa m . Supongamos ahora que el (G, A) -bandido tiene horizonte 1. Como el horizonte de $A^{(1)}$ es cero, tenemos que para cualquier distribución $G^{(1)}$

$$V(G^{(1)}, A^{(1)}) = 0 \tag{34}$$

y así (32) se simplifica:

$$V(G, (\alpha_1, 0, \dots)) = \alpha_1 \prod_{i=1}^k \int_{\mathcal{D}^k} \int_{\mathbb{R}} x Q_i(dx) G(d(Q_1, \dots, Q_k)) \tag{35}$$

$$= \alpha_1 \prod_{i=1}^k \int_{\mathcal{D}} \int_{\mathbb{R}} x Q_i(dx) F_i(dQ_i) \tag{36}$$

Una selección inicial óptima es cualquier arma para la cual el supremo en (35) es alcanzado.

Supongamos ahora que el (G, A) -bandido tiene horizonte 2. Para poder utilizar la expresión (32) requerimos los valores de los bandidos que tienen como sucesión de descuento a $A^{(1)}$ (cuyo horizonte es 1). Observamos que las únicas distribuciones que necesitamos considerar son aquellas de la forma $(x)_i G$ para las observaciones x en el arma i , los valores $V((x)_i G, A^{(1)})$ son obtenidos con (35). Cuando utilizamos (32) para encontrar a $V(G, A)$, las selecciones iniciales óptimas son aquellas indexadas por los i 's tales que el máximo en (32) es alcanzado.

En general, cuando se tenga un (G, A) -bandido con horizonte $n < \infty$ lo primero que es necesario hacer para cada $m = 1, 2, \dots, n - 1$ las posibles distribuciones condicionales $G^{(m)}$ de (Q_1, \dots, Q_k) dadas las observaciones correspondientes. Dado que $A^{(n-1)}$ tiene horizonte uno, cada $V(G^{(n-1)}, A^{(n-1)})$ es obtenido mediante el uso de (35). A continuación, cada $V(G^{(n-2)}, A^{(n-2)})$ es calculado a través de (32). Este proceso se repite de manera iterada hasta que $V(G, A)$ es obtenido.

4 Aplicación de la Técnica de Programación Dinámica al problema de incorporación de personal

Como ya se mencionó, en el problema de asignación de personal el elemento $G \in \mathcal{D}(\mathcal{D}^k)$ correspondiente está soportado por un subconjunto de (distribuciones de Bernoulli) \times (distribuciones de Bernoulli) $\times \dots \times$ (distribuciones de Bernoulli). Es decir, en este caso cada Q_i es una distribución de Bernoulli y por tanto tiene asociado un parámetro θ_i , así que podemos considerar que G es también una distribución sobre $[0, 1]^k$. Tenemos además que, dado el objetivo del administrador

solo consideraremos una sucesión de descuento para cada horizonte n ; a saber, $A = (1, 1, 1, \dots, 1, 0, 0, \dots)$. Dada la naturaleza de las distribuciones de las armas, denotaremos a sus dos únicos posibles resultados por $(1)_i$, para un éxito en el arma i (el empleado i atendió satisfactoriamente a un cliente) y por $(0)_i$, para un fracaso en el arma i (el empleado i no atendió satisfactoriamente a un cliente).

4.1 Ejemplos

En esta parte del trabajo presentaremos algunos ejemplos de distribuciones G que podrían aparecer en un problema concreto.

1. Armas Dependientes

Supongamos que, como consecuencia de una análisis de los currículums y del periodo de prueba al que fueron sometidos dos nuevos empleados, se concluye que ambos son muy buenos o ambos son muy malos, inclinándose más por la primer posibilidad; explícitamente:

$$G = \frac{4}{5}\delta_{(1/10, 15/100)} + \frac{1}{5}\delta_{(9/10, 8/10)}, \quad (37)$$

en donde $\delta_{(x,y)}$ es la medida delta en (x, y) . Supongamos que en este caso el número de clientes que deben ser repartidos es $n = 2$, así que $A = (1, 1, 0, \dots)$.

Comenzamos hallando las distribuciones *a posteriori* $G^{(n-1)} = G^{(2)}$ (que en este caso existen cuatro de ellas) utilizando tanto a la función de verosimilitud y la distribución *a priori* G [4]. Haremos los detalles para hallar la distribución $(1)_1G$, las otras tres distribuciones pueden encontrarse de manera similar. Tenemos que hallar $a, b \in \mathbb{R}$, en donde a y b son tales que:

$$(1)_1G = a\delta_{(1/10, 15/100)} + b\delta_{(9/10, 8/10)}. \quad (38)$$

Así,

$$\begin{aligned} a = \mathbb{P}[B_1|(1)_1] &= \frac{\mathbb{P}[(1)_1|B_1]\mathbb{P}[B_1]}{\mathbb{P}[(1)_1|B_1]\mathbb{P}[B_1] + \mathbb{P}[(1)_1|B_2]\mathbb{P}[B_2]} \\ &= \frac{\frac{1}{10}\frac{4}{5}}{\frac{1}{10}\frac{4}{5} + \frac{9}{10}\frac{1}{5}} = \frac{4}{13}, \end{aligned}$$

donde $B_1 = \{\theta_1 = \frac{1}{10}, \theta_2 = \frac{15}{100}\}$ y $B_2 = \{\theta_1 = \frac{9}{10}, \theta_2 = \frac{8}{10}\}$. De manera similar,

tenemos que

$$\begin{aligned} b = \mathbb{P}[B_2|(1)_1] &= \frac{\mathbb{P}[(1)_1|B_2] \mathbb{P}[B_2]}{\mathbb{P}[(1)_1|B_1] \mathbb{P}[B_1] + \mathbb{P}[(1)_1|B_2] \mathbb{P}[B_2]} \\ &= \frac{\frac{9}{10} \frac{1}{5}}{\frac{1}{10} \frac{4}{5} + \frac{9}{10} \frac{1}{5}} = \frac{9}{13}, \end{aligned}$$

Tenemos que

$$(1)_1G = \frac{4}{13} \delta_{(1/10,15/100)} + \frac{9}{13} \delta_{(9/10,8/10)}. \quad (39)$$

Análogamente

$$(0)_1G = \frac{36}{37} \delta_{(1/10,15/100)} + \frac{1}{37} \delta_{(9/10,8/10)} \quad (40)$$

$$(1)_2G = \frac{3}{7} \delta_{(1/10,15/100)} + \frac{4}{7} \delta_{(9/10,8/10)} \quad (41)$$

$$(0)_2G = \frac{17}{18} \delta_{(1/10,15/100)} + \frac{1}{18} \delta_{(9/10,8/10)}. \quad (42)$$

De (35), tenemos que

$$V((1)_1G, A^{(1)}) = \left(\frac{4}{13} \frac{1}{10} + \frac{9}{13} \frac{9}{10} \right) \vee \left(\frac{4}{13} \frac{15}{100} + \frac{9}{13} \frac{8}{10} \right) \quad (43)$$

$$= \frac{85}{130} \vee \frac{132}{130} = \frac{132}{130} \quad (44)$$

$$V((0)_1G, A^{(1)}) = \left(\frac{36}{37} \frac{1}{10} + \frac{1}{37} \frac{9}{10} \right) \vee \left(\frac{36}{37} \frac{15}{100} + \frac{1}{37} \frac{8}{10} \right) \quad (45)$$

$$= \frac{45}{370} \vee \frac{62}{370} = \frac{62}{370} \quad (46)$$

$$V((1)_2G, A^{(1)}) = \left(\frac{3}{7} \frac{1}{10} + \frac{4}{7} \frac{9}{10} \right) \vee \left(\frac{3}{7} \frac{15}{100} + \frac{4}{7} \frac{8}{10} \right) \quad (47)$$

$$= \frac{39}{70} \vee \frac{365}{700} = \frac{39}{70} \quad (48)$$

$$V((0)_2G, A^{(1)}) = \left(\frac{17}{18} \frac{1}{10} + \frac{1}{18} \frac{9}{10} \right) \vee \left(\frac{17}{18} \frac{15}{100} + \frac{1}{18} \frac{8}{10} \right) \quad (49)$$

$$= \frac{26}{180} \vee \frac{335}{1800} = \frac{335}{1800}. \quad (50)$$

Tenemos ahora, a través de (32) que

$$\begin{aligned} V(G, A) &= \left\{ \frac{4}{5} \left[\frac{62}{370} \frac{9}{10} + \left(1 + \frac{132}{130} \right) \frac{1}{10} \right] + \frac{1}{5} \left[\frac{62}{370} \frac{1}{10} + \left(1 + \frac{132}{130} \right) \frac{9}{10} \right] \right\} \vee \\ &\quad \left\{ \frac{4}{5} \left[\frac{335}{1800} \frac{85}{100} + \left(1 + \frac{39}{70} \right) \frac{15}{100} \right] + \frac{1}{5} \left[\frac{335}{1800} \frac{2}{10} + \left(1 + \frac{39}{70} \right) \frac{8}{10} \right] \right\} \\ &= \frac{81}{125} \vee \frac{1709}{1000} = \frac{1709}{1000}. \end{aligned} \quad (51)$$

Tenemos entonces, observando a (51), que solo existe una selección inicial óptima; a saber, el empleado 2. La segunda selección depende del resultado de la primer observación: si el resultado es un éxito $((1)_2)$ el administrador debe cambiar al empleado 1, ya que $\left(\frac{39}{70} > \frac{365}{700}\right)$. Por otro lado, si el resultado es un fracaso $((0)_2)$ el administrador debe asignarle el segundo cliente al empleado 2 nuevamente. Dicho de otra forma, para la selección inicial tenemos que:

$$\tau(\phi) = 2 \quad (52)$$

mientras que para la segunda selección

$$\tau(z_1) = \begin{cases} 1 & z_1 = (1)_2 \\ 2 & z_1 = (0)_2 \end{cases}$$

algo que vale la pena resaltar es que si el empleado 2 (que es seleccionado inicialmente) atiende satisfactoriamente a un cliente, el administrador no le asignará la atención del segundo cliente.

2. Armas con distribuciones continuas

Supongamos ahora que tenemos a tres empleados. Supongamos además que uno de ellos (digamos el tercero) es conocido en el sentido de que se conoce cabalmente su eficacia, supongamos que $\theta_3 = 8/15$. Y que la distribución de los otros dos empleados es uniforme en $[0, 1]^2$. Así que tenemos que $G = F_1 \times F_2 \times F_3$ en donde F_1 y F_2 son distribuciones uniformes en $[0, 1]$ y $F_3 = \delta_{8/15}$. Supongamos además que el número de clientes a ser repartidos es $n = 3$ ($A = (1, 1, 1, 0, \dots)$).

Para calcular las distribuciones $G^{(2)}$ tenemos que notar que tanto θ_1 como θ_2 tienen distribuciones *a posteriori* de tipo Beta [4]. Hallemos ahora algunos de los 15 valores diferentes de $V(G^{(2)}, A^{(2)})$:

- (a) De acuerdo con (35) para encontrar a $V((1)_1(1)_1G, A^{(2)})$ tenemos que calcular y comprar las siguientes 3 integrales

$$\frac{\Gamma(4)\Gamma(2)}{\Gamma(3)\Gamma(1)^3} \int_0^1 \int_0^1 \theta_1^3 d\theta_1 d\theta_2 = \frac{3}{4} \quad (53)$$

	$(0)_1$	$(1)_1$	$(0)_1 (0)_1$	$(0)_1 (1)_1$	$(1)_1 (1)_1$
$(0)_2$	$(1/2, 1/2, 8/15)$	$(1/3, 1/2, 8/15)$	$(2/3, 1/2, 8/15)$	$(1/4, 1/2, 8/15)$	$(1/2, 1/2, 8/15)$
$(1)_2$	$(1/2, 1/3, 8/15)$	$(1/3, 1/3, 8/15)$	$(2/3, 1/3, 8/15)$		
$(0)_2 (0)_2$	$(1/2, 2/3, 8/15)$	$(1/3, 2/3, 8/15)$	$(2/3, 2/3, 8/15)$		
$(0)_2 (1)_2$	$(1/2, 1/4, 8/15)$				
$(1)_2 (1)_2$	$(1/2, 1/2, 8/15)$				
	$(1/2, 3/4, 8/15)$				

Figura 1: $m=2$

$$\frac{\Gamma(4)\Gamma(2)}{\Gamma(3)\Gamma(1)^3} \int_0^1 \int_0^1 \theta_1^2 \theta_2 d\theta_1 d\theta_2 = \frac{1}{2} \quad (54)$$

$$\int_{\mathbb{R}} x d\delta_{8/15} = \frac{8}{15} \quad (55)$$

y así $V((1)_1(1)_1G, A^{(2)}) = \frac{3}{4} \vee \frac{1}{2} \vee \frac{8}{15} = \frac{3}{4}$.

(b) Encontremos ahora a $V((1)_1(0)_2G, A^{(2)})$, para esto, tenemos que calcular:

$$\frac{\Gamma(3)\Gamma(3)}{\Gamma(2)^2\Gamma(1)^2} \int_0^1 \int_0^1 \theta_1^2 (1 - \theta_2) d\theta_1 d\theta_2 = \frac{2}{3} \quad (56)$$

$$\frac{\Gamma(3)\Gamma(3)}{\Gamma(2)^2\Gamma(1)^2} \int_0^1 \int_0^1 \theta_1 \theta_2 (1 - \theta_2) d\theta_1 d\theta_2 = \frac{1}{12} \quad (57)$$

$$\int_{\mathbb{R}} x d\delta_{8/15} = \frac{8}{15} \quad (58)$$

y así $V((1)_1(0)_2G, A^{(2)}) = \frac{2}{3} \vee \frac{1}{12} \vee \frac{8}{15} = \frac{2}{3}$.

Podemos observar también que $V((a)_i(b)_jG, A^{(2)}) = V((b)_j(a)_iG, A^{(2)})$, para $a, b = 0, 1$ e $i, j = 1, 2, 3$. Además, $V((a)_3(b)_3G, A^{(2)}) = \frac{1}{2} \vee \frac{1}{2} \vee \frac{8}{15}$ (celda azul de la Figura 1). Por otro lado, las celdas verdes corresponden a $V((a)_3(b)_iG, A^{(2)})$, para $a, b = 0, 1$ e $i = 1, 2$.

Posteriormente, bajo las mismas consideraciones, hallaremos los cinco valores

	(0) ₁	(1) ₁
(0) ₂	(198/180, 198/180, 192/180)	(156/180, 198/180, 192/180) (242/180, 210/180, 216/180)
(1) ₂	(198/180, 156/180, 192/180)	
	(210/180, 242/180, 216/180)	

Figura 2: m=1

diferentes de $V(G^{(1)}, A^{(1)})$. Por ejemplo, para calcular $V((1)_1G, A^{(1)})$ tenemos que

$$\begin{aligned}
& V((1)_1G, A^{(1)}) \\
&= \left(2 \int_0^1 [0 + V((0)_1(1)_1G, A^{(2)})] (1 - \theta_1) + [1 + V((1)_1(1)_1G, A^{(2)})] \theta_1 d\theta_1 \right) \vee \\
&\quad \left(\int_0^1 [0 + V((0)_2(1)_1G, A^{(2)})] (1 - \theta_1) + [1 + V((1)_2(1)_1G, A^{(2)})] \theta_1 d\theta_1 \right) \vee \\
&\quad \left(\frac{7}{15} [0 + V((0)_3(1)_1G, A^{(2)})] + \frac{8}{15} [1 + V((1)_3(1)_1G, A^{(2)})] \right) \\
&= \left(2 \int_0^1 \left[0 + \frac{8}{15} \right] (1 - \theta_1) + \left[1 + \frac{3}{4} \right] \theta_1 d\theta_1 \right) \vee \\
&\quad \left(\int_0^1 \left[0 + \frac{2}{3} \right] (1 - \theta_1) + \left[1 + \frac{2}{3} \right] \theta_1 d\theta_1 \right) \vee \\
&\quad \left(\frac{7}{15} \left[0 + \frac{2}{3} \right] + \frac{8}{15} \left[1 + \frac{2}{3} \right] \right) \\
&= \frac{242}{180} \vee \frac{210}{180} \vee \frac{216}{180} = \frac{242}{180}
\end{aligned}$$

En la Figura 2 observamos los valores de $V(G^{(1)}, A^{(1)})$ en donde la celda azul corresponde tanto a $V((1)_3G, A^{(1)})$ como a $V((0)_3G, A^{(1)})$.

Finalmente, calculamos $V(G, A)$:

$$\begin{aligned}
V(G, A) &= \int_0^1 [(0 + V((0)_1G, A^{(1)})) (1 - \theta_1) + (1 + V((1)_1G, A^{(1)})) \theta_1] d\theta_1 \vee \\
&\quad \int_0^1 [(0 + V((0)_2G, A^{(1)})) (1 - \theta_2) + (1 + V((1)_2G, A^{(1)})) \theta_2] d\theta_2 \vee \\
&\quad \frac{7}{15} (0 + V((0)_3G, A^{(1)})) + \frac{8}{15} (1 + V((1)_3G, A^{(1)})) \\
&= \frac{310}{180} \vee \frac{310}{180} \vee \frac{294}{180}
\end{aligned}$$

Tenemos entonces que existen dos elecciones iniciales óptimas ($\tau(\phi) = 1$ y $\tau(\phi) = 2$). Observamos finalmente en las Figuras 3 y 4 las estrategias óptimas resultantes.

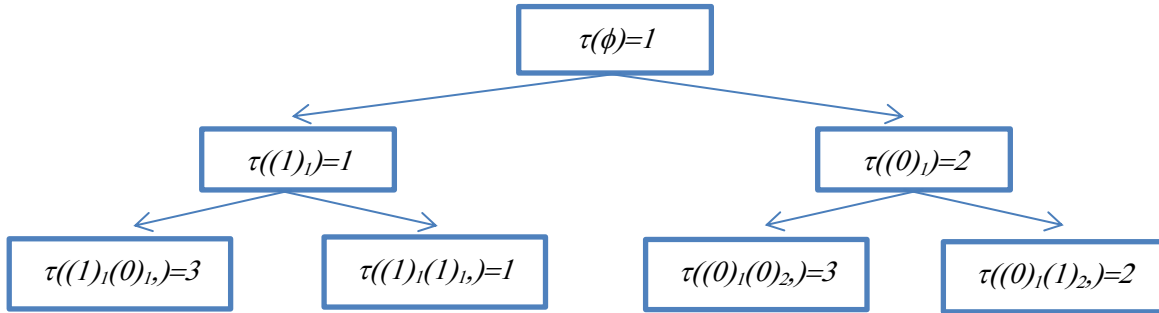


Figura 3: Estrategias óptimas con $\tau(\phi) = 1$

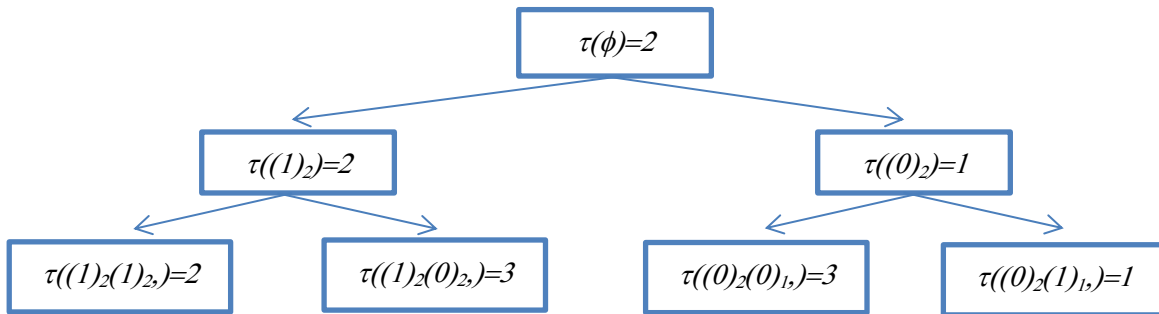


Figura 4: Estrategias óptimas con $\tau(\phi) = 2$

5 Conclusiones

La técnica de Programación Dinámica y el Modelo de Bandidos Armados resultaron ser eficaces para abordar el problema en que nuevo personal con tasas de fallos desconocidas deben ser asignados uno a la vez para atender a los clientes de una empresa y maximizar el número en que estos son atendidos satisfactoriamente.

Agradecimientos

Esta trabajo fue apoyado por el proyecto VIEP “El modelo de dos bandidos armados”.

Bibliografía

- [1] Berry D.A. and Fristedt B., *Bandit Problems*, Chpaman and Hall, 1985.
- [2] Parthasarathy K. R. *Probability measures on metric spaces*, Academic Press, 1967.

- [3] Hernández-Lerma O. y Lasserre J.B. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. Springer-Verlag, 1996.
- [4] Wakefield J. *Bayesian and Frequentist Regression Methods*. Springer Verlag, 2013.

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

Capítulo 8

Diseño y análisis de un experimento de uso eficiente de nitrógeno en maíz

Bulmaro Juárez Hernández¹, Jenaro Reyes Matamoros², Thomas Lübberstedt³

FCFM-BUAP¹, ICUAP-BUAP², ISU³

Resumen

El objetivo del estudio fue evaluar la respuesta morfológica de plántulas de maíz a la aplicación de niveles contrastantes de nitrógeno. En calidad de material vegetal fueron estudiados 4 cultivares de maíz sembrado en el estado de Puebla: variedad Sb 302 Berentsen, criollo azul (*Hueytamalco*), criollo blanco (*Tepexi de Rodríguez*) y criollo azul (*Tepexi de Rodríguez*). El experimento se llevó a cabo considerando los factores, variedad de cultivares y dosis de nitrógeno aplicada, utilizando un diseño completamente aleatorizado. Las plántulas fueron cultivadas en solución nutritiva durante 21 días en condiciones controladas. Se utilizó una solución nutritiva modificada con las dosis de 20 % y 100 % de nitrógeno. Para el análisis de las variables medidas se utilizó primero el método de componentes principales con el propósito de tener una sola variable respuesta, y en seguida se realizó un análisis de varianza para determinar la respuesta de los cultivares de maíz a la fertilización nitrogenada. Cabe indicar que se presenta una justificación estadística para utilizar sólo la variable de longitud total de raíces secundarias como variable respuesta, sin que esto represente una pérdida considerable de información para realizar el análisis estadístico mediante un modelo lineal bifactorial. Por último, se muestra que con un nivel de significancia $\alpha = 0.05$ no existe efecto de interacción entre los factores cultivar y dosis de nitrógeno. Asimismo, con un modelo sin el término de interacción y con el mismo nivel de significancia se obtiene que no existe efecto de tratamiento en el factor cultivar, en tanto que existe efecto de tratamiento en el factor dosis de nitrógeno.

1 Introducción

México es el centro de origen y de diversidad genética del maíz (*Zea mays L.*), el cual se ha ubicado en el primer término a nivel mundial superando al trigo y al arroz. La importancia del maíz es mucho mayor que la de cualquier otro cultivo,

ya que se usa como forraje, alimento humano y en diversos procesos industriales. La superficie destinada a la producción de maíz a nivel mundial ascendió a 176.9 millones de hectáreas para el año 2012 [13] y el cultivo de este cereal, en los países en desarrollo, representa aproximadamente el 50 % del área total sembrada. En estas regiones el cultivo se ha hecho usualmente bajo condiciones de baja disponibilidad de nitrógeno, como resultado de una natural baja fertilidad del suelo.

En México el maíz es el cultivo de mayor importancia en la agricultura. Y tiene desde el punto de vista económico y social un lugar relevante en nuestro país, ya que en este cereal se basa un alto porcentaje de la alimentación de sus habitantes. Por lo que poco más de 7.3 millones de hectáreas (40 % de la superficie agrícola) son sembradas año con año [14]. Este se cultiva en todas las regiones geográficas del país y con diferentes niveles de tecnología, donde el uso de fertilizantes juega un papel fundamental para mantener la producción ya que estos insumos proporcionan los nutrientes que las plantas necesitan para su desarrollo y son los que representan mayor costo de producción [1].

El nitrógeno (N) como fertilizante es un elemento importante para lograr rendimientos satisfactorios, sin embargo, el uso indiscriminado de N ha traído cambios negativos en el ambiente, principalmente en la atmósfera, mantos acuíferos y sistemas lagunares, entre otros componentes de los ecosistemas, adicionalmente el alto costo energético hace necesario incrementar la eficiencia en el uso de N, una de las alternativas es el uso de variedades que absorban y metabolicen el N de manera más eficiente. Entre los elementos minerales esenciales, el N es el que con más frecuencia limita el crecimiento y el rendimiento del maíz. Esta condición ocurre porque las plantas requieren cantidades relativamente grandes de N (de 1.5 a 3.5 % de peso seco de la planta) y porque la mayoría de los suelos no tienen suficiente N en forma disponible para mantener los niveles deseados de producción. Debido a que la deficiencia de N puede disminuir el rendimiento y la calidad del grano, es necesario tomar medidas para asegurar que niveles adecuados de N estén disponibles para las plantas. Estimativos globales sugieren que los fertilizantes nitrogenados cubren el 80 % del costo total de fertilizantes y el 30 % de toda la energía asociada con la producción agrícola moderna de maíz [2].

Los productores de maíz reconocen que son necesarias concentraciones adecuadas de N en la planta para obtener altos rendimientos, sin embargo, el dilema está en conocer qué cantidades aplicar para lograr estas concentraciones. El N del fertilizante que no es aprovechado produce, además del perjuicio económico, daño ambiental por pérdida del nutriente a capas inferiores del perfil del suelo. Las pérdidas excesivas de fertilizante nitrogenado de los cultivos pueden contaminar las aguas profundas con NO_3^- [2]. Con el creciente interés de la opinión pública en la calidad ambiental aumentan también las presiones sobre los agricultores en

Cuadro 1: Cultivares de maíz utilizado en el estudio

Número	Cultivar	Localidad
1	SB 302 Berentsen*	Teziutlán
2	Criollo azul	Hueytamalco
3	Criolla blanco	Tepexi de Rodríguez
4	Criollo azul	Tepexi de Rodríguez

*La variedad mejorada se adquirió en esta localidad.

mejorar el manejo del N. Con la adopción creciente de prácticas de conservación de suelos también aumentarán las prácticas relacionadas con el mejor manejo de los fertilizantes nitrogenados.

Considerando la importancia que tiene incrementar el uso eficiente de nitrógeno en el cultivo de maíz, el objetivo del estudio fue evaluar la respuesta morfológica de plántulas de maíz a la aplicación de niveles contrastantes de nitrógeno. Para realizar el análisis de datos se llevó a cabo un estudio de las variables medidas mediante el método de componentes principales (CP) con el propósito de tener una sola variable o índice como respuesta para llevar a cabo una evaluación biométrica a través de un análisis de varianza (ANVA) aplicado a un modelo bifactorial para determinar la respuesta de los cultivares de maíz a las dosis de nitrógeno. Se está interesado en evaluar la respuesta de algunas características de la morfología de la raíz de plantulas de maíz a la aplicación de diferentes dosis de nitrógeno.

2 Material y Métodos

Localidad

El experimento se llevó a cabo en condiciones de invernadero en las instalaciones del Departamento de Investigación en Ciencias Agrícolas de la Benemérita Universidad Autónoma de Puebla, localizada a $19^{\circ} 14'$ latitud norte, y $98^{\circ} 18'$ longitud oeste, a una altitud de 2150 msnm [3].

Material vegetal

En calidad de materia vegetal se utilizaron 4 cultivares de maíz, los cuales provienen de distintas regiones ecológicas del estado de Puebla (Cuadro 1).

Cuadro 2: Parámetros evaluados en las plántulas de maíz.

Variable	Abreviatura
Longitud de mesocótilo	LM
Longitud de la raíz primaria	LRPr
Número de raíces secundarias	NRS
Longitud total de raíces secundarias	LTRS
Número de raíces adventicias	NRA
Longitud total de raíces adventicias	LTRA
Longitud de la plántula	LP1
Peso fresco de la raíz	PFR
Peso fresco de la plántula	PRP1
Peso seco de la raíz	PSR
Peso seco de la plántula	PSP1

Métodos

El experimento se llevó a efecto considerando los factores variedad de cultivar y dosis de nitrógeno aplicada, siguiendo un diseño completamente aleatorizado (DCA). Las plántulas fueron cultivadas en solución nutritiva por 21 días usando el método de rollo de papel descrito por Woll et al. [12]. Se utilizó una solución nutritiva modificada de Hoagland y Arnon [6]. En la solución nutritiva se evaluaron las concentraciones de 20 % y 100 % de nitrógeno. De cada cultivar, en cada tratamiento, se eligieron para la medición de manera aleatoria seis plántulas sanas.

VARIABLES EVALUADAS

Las plántulas de maíz fueron evaluadas a los 21 días de germinación. Las variables estudiadas se presentan en el Cuadro 2.

Las variables LM, LRPr, LTRS, LTRA y LP1 fueron medidas de manera manual usando una regla. Particularmente, LP1 se midió tomando el nodo marcado por las raíces adventicias, hasta la hoja más larga de la plántula. Las variables PFR y PFPl fueron medidas con una báscula de precisión. El PSR y PSP1 fueron medidos después de que las muestras permanecieron 24 horas a 70 °C en la estufa de secado, cabe aclarar que las mediciones se efectuaron por una sola persona.

Cuadro 3: Matriz de correlaciones para los datos obtenidos con dosis de 20 % de N.

	LM	NRS	LTRS	NRA	LTRA	LP1	PFP1	PFR	PSP1	PSR
LRPr	0.388	0.477	0.476	0.302	0.504	0.132	0.323	0.398	0.272	0.486
LM		0.465	0.400	0.135	0.103	0.133	0.284	0.338	0.227	0.455
NRS			0.799	0.488	0.551	0.576	0.676	0.460	0.656	0.675
LTRS				0.328	0.515	0.624	0.622	0.555	0.563	0.675
NRA					0.641	0.375	0.456	0.160	0.451	0.437
LTRA						0.519	0.669	0.560	0.654	0.678
LP1							0.839	0.567	0.729	0.649
PFP1								0.790	0.941	0.837
PFR									0.781	0.877
PSP1										0.829

3 Análisis estadístico

Una vez obtenidos los datos, se realizó un análisis de las variables con ayuda del software estadístico R ([4], [10]) mediante el método de componentes principales (CP)[7], con el propósito de establecer una sola variable respuesta, la cual se estudió por medio del análisis de varianza (ANVA) utilizando un modelo bifactorial para determinar la respuesta de los cultivares de maíz a las dosis de nitrógeno ([5], [8], [9], [11]).

Análisis de los datos de las plántulas de maíz con dos tratamientos de nitrógeno

Para determinar la respuesta de los cultivares de maíz a las dosis de nitrógeno se analiza si existe correlación entre las diferentes variables medidas, de lo cual se podría derivar la necesidad de reducir el número de variables que se deben utilizar. Asimismo, se podría considerar la posibilidad de usar un índice para representar de forma conveniente el resultado en el ANVA para detectar diferencias entre los distintos tratamientos propuestos. El análisis de datos de las variables de las plántulas de maíz con los dos tratamientos de nitrógeno se muestra a continuación.

En el Cuadro 3 se presenta la matriz de correlaciones entre las variables en los datos con dosis de 20 % de N. La matriz de correlaciones muestra coeficientes con valores positivos muy altos entre las variables PFP1 y LPI, PSP1 y PFP1, y entre PSR y los parámetros PFP1, PFR y PSP1 para la dosis de 20 % de N, a los 21 días de crecimiento de las plántulas. Cabe señalar que las variables LTRS y NRS, PFR y PFP1, y entre PSP1 y los parámetros LPI, PFR presentaron una correlación

Cuadro 4: Primeras cinco componentes principales para el 20 % de N.

	CP1	CP2	CP3	CP4	CP5
LRPr	0.034	0.052	-0.747	-0.642	0.141
LM	0.002	-0.002	-0.021	-0.023	-0.038
NRS	0.041	0.014	-0.005	-0.222	-0.768
LTRS	0.968	-0.246	-0.008	0.047	0.019
NRA	0.013	0.052	0.031	-0.077	-0.594
LTRA	0.242	0.966	0.011	0.081	0.033
LP1	0.045	0.036	0.662	-0.720	0.179
PFPr	0.005	0.007	0.038	-0.059	-0.009
PFR	0.004	0.006	0.009	-0.033	0.059
PSP1	0.000	0.001	0.004	-0.004	-0.003
PSR	0.001	0.001	0.001	-0.004	0.000

positiva alta. En particular puede notarse que la variable LTRS está altamente correlacionada con todas las variables excepto con NRA. Es conveniente mencionar que las correlaciones observadas sí tienen una explicación natural, ya que si la LTRS es grande generalmente se tendrá un mayor NRS; por otro lado sí el NRS es grande, en general la LRPr será grande, lo cual lleva a tener una alta correlación con la LTRS. De tal manera que las correlaciones mostradas tienen una base natural, siendo de importancia considerar la LTRS. Las observaciones anteriores permiten considerar la conveniencia de realizar un análisis de CP para investigar si es posible proponer un índice como variable respuesta e intentar interpretar tal índice en función de las ponderaciones asignadas a cada variable de esta CP ([7], [10]).

En el Cuadro 4 se registran las primeras cinco componentes principales para la dosis de 20 % de N.

Con los resultados obtenidos para las variables medidas de las plántulas cultivadas con la dosis de 20 % de N, se realiza el análisis de CP para tratar de proponer un índice que esté en términos de los resultados obtenidos para las diferentes variables en estudio y que este sea utilizado como variable respuesta en el ANVA. En el Cuadro 4 se presentan las primeras cinco CP y en el Cuadro 5 se puede observar la importancia de cada uno de estos CP a través de sus proporciones de varianzas, asimismo se observa la proporción de la varianzas acumulada de los primeros cinco CP. Con base en estos resultados se decide considerar sólo a la primera CP, sin tener una pérdida grande respecto a la variación fenotípica de las plántulas cultivadas con

Cuadro 5: Importancia de los componentes principales para el 20 % de N.

	CP1	CP2	CP3	CP4	CP5	CP6	CP7
Desviación estándar	37.335	12.873	2.378	1.608	1.169	0.849	0.219
Proporción de varianza	0.888	0.105	0.004	0.002	0.001	0.000	0.000
Proporción acumulada	0.888	0.993	0.997	0.999	1.000	1.000	1.000

Cuadro 6: Matriz de correlaciones para los datos obtenidos con dosis de 100 % de N.

	LM	NRS	LTRS	NRA	LTRA	LP1	PFP1	PFR	PSP1	PSR
LRPr	0.266	-0.016	0.242	-0.153	0.105	-0.204	0.135	0.030	0.166	-0.031
LM		0.042	0.098	-0.177	0.055	-0.058	-0.076	0.210	0.070	0.101
NRS			0.686	0.324	0.409	0.427	0.509	0.381	0.462	0.418
LTRS				0.467	0.644	0.409	0.594	0.653	0.596	0.629
NRA					0.529	0.455	0.618	0.348	0.455	0.408
LTRA						0.615	0.580	0.601	0.399	0.544
LP1							0.647	0.602	0.520	0.669
PFP1								0.675	0.668	0.745
PFR									0.501	0.921
PSP1										0.631

la dosis del 20 % de N, pues ella sola retiene el 88.87 % de la variabilidad total. Por otro lado, también se puede observar en el Cuadro 4, que en relación a los pesos asociados a la primera componente principal, el peso más alto está asociado a la variable LTRS con un 72 % respecto al resto de las variables, por lo que se decide sólo trabajar con esta variable más que con el índice representado por la primera CP para llevar a cabo el ANVA para determinar el efecto del tratamiento en el experimento.

En el Cuadro 6 se observa la matriz de correlaciones entre las variables en estudio para los datos con dosis de 100 % de N. La matriz de correlaciones muestra coeficientes con valores positivos muy altos entre las variables PSR y PFR para la dosis de 100 % de N, a los 21 días de germinación de las plántulas. Las variables PSR y PFP1 presentaron una correlación positiva alta. Se puede advertir que la variable LTRS nuevamente está altamente correlacionada con todas las variables excepto con NRA y LP1, aun cuando la correlación entre estas últimas variables no es muy baja.

En el Cuadro 7 se presentan los resultados del análisis de CP para la dosis de 100 % de N. Estos resultados y los presentados en el Cuadro 8 confirman que la

Cuadro 7: Primeras cinco componentes principales para el 100 % de N.

	CP1	CP2	CP3	CP4	CP5
LRPr	0.021	-0.042	0.689	0.695	-0.110
LM	0.001	-0.001	0.022	0.018	-0.015
NRS	0.054	-0.022	-0.253	0.067	-0.954
LTRS	0.969	-0.238	-0.007	-0.023	0.058
NRA	0.023	0.053	-0.117	-0.085	0.152
LTRA	0.234	0.945	0.174	-0.125	-0.068
LP1	0.049	0.211	-0.644	0.695	0.209
PFP1	0.010	0.015	-0.032	-0.070	0.000
PFR	0.008	0.011	-0.024	0.026	0.048
PSP1	0.001	0.000	-0.003	0.007	0.002
PSR	0.001	0.001	-0.003	0.003	0.004

Cuadro 8: Importancia de los componentes principales para el 100 % de N.

	CP1	CP2	CP3	CP4	CP5	CP6	CP7
Desviación estándar	5.085	6.474	2.591	1.748	1.322	0.927	0.260
Proporción de varianza	0.920	0.061	0.010	0.004	0.003	0.001	0.000
Proporción acumulada	0.920	0.981	0.991	0.995	0.998	0.999	0.999

primera CP (CP1) que resulta de la transformación de las variables bajo estudio de las plántulas cultivadas con la dosis de 100 % de N es más que suficiente para considerarla como un índice que represente los resultados obtenidos para las variables estudiadas, sin tener una pérdida grande respecto a la variación fenotípica de las plántulas, de forma que la CP1 puede ser utilizada como variable respuesta. Por otro lado, en el Cuadro 7, se confirma que el peso más alto correspondiente a la primera componente principal, está asociado a la variable LTRS con un 72 % respecto a todas las variables bajo estudio, por lo que se decide sólo trabajar con esta variable más que con el índice representado por CP1.

Modelo lineal bifactorial asociado al diseño del experimento

El experimento se realizó aplicando de forma completamente aleatoria dos diferentes dosis de nitrógeno a cuatro variedades de maíz, resultando un diseño bifactorial

[9] al cual se le puede asociar el siguiente modelo lineal para describir el tipo de observaciones obtenidas:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \quad \text{para} \quad \begin{cases} i = 1, 2, 3, 4 \\ j = 1, 2 \\ k = 1, 2, \dots, 6. \end{cases}$$

Donde, y_{ijk} representa la respuesta observada cuando el factor cultivar (tipo de maíz) se encuentra en el i -ésimo nivel con $i = 1, 2, 3, 4$; el factor dosis de N, se encuentra en el j -ésimo nivel con $j = 1, 2$, y se tienen 6 repeticiones o replicas para cada combinación de los niveles de los factores.

Por otro lado, en este modelo, los términos en la parte derecha de la igualdad representan:

μ , el efecto medio general;

τ_i , el efecto del i -ésimo nivel del factor cultivar;

β_j , el efecto del j -ésimo nivel del factor dosis de N;

$(\tau\beta)_{ij}$, el efecto de la interacción entre τ_i y β_j , y

ε_{ijk} , el componente del error aleatorio.

En este diseño se supone que ambos factores son fijos y que los efectos de tratamiento se definen como desviaciones de la media general, por lo que:

$$\sum_{i=1}^4 \tau_i = 0 \quad \text{y} \quad \sum_{j=1}^2 \beta_j = 0,$$

además, se supone que los efectos de interacción son fijos y que se definen de manera que:

$$\sum_{i=1}^4 \sum_{j=1}^2 (\tau\beta)_{ij} = 0.$$

Finalmente, se supone que los ε_{ijk} conforman una muestra aleatoria de una distribución normal con media cero.

Nótese que hay un total de $4 \cdot 2 \cdot 6 = 48$ observaciones ya que se realizan 6 replicas.

En un diseño bifactorial, tanto los factores (o tratamientos) de renglón como de columna tienen la misma importancia. Esto es, se está interesado en probar hipótesis acerca de la igualdad de los efectos de tratamiento de renglón, es decir,

$$H_o : \tau_1 = \tau_2 = \tau_3 = \tau_4 = 0 \quad \text{vs} \quad H_a : \text{al menos una } \tau_i \neq 0,$$

Cuadro 9: ANVA de dos factores: cultivar vs. dosis de N (20 % y 100 %), con variable respuesta LTRS y con término de efecto de interacción.

Fuente	GL	SC	CM	F	p-valor
Cultivar	3	4520.7	1506.9	2.00	0.129
Dosis de N	1	14698.3	14698.3	19.55	0.000
Interacción	3	887.4	295.8	0.39	0.758
Error	40	30076.5	751.9		
Total	47	50182.8			

de la igualdad de los efectos de tratamiento de columna, esto es,

$$H_o : \beta_1 = \beta_2 = 0 \text{ vs } H_a : \text{ al menos una } \beta_j \neq 0, i = 1, \dots, 4; j = 1, 2.$$

También se está interesado en determinar si los tratamientos de renglón y columna interaccionan. Es decir, es conveniente probar

$$H_o : (\tau\beta)_{ij} = 0, \text{ para todas las } (i,j). \text{ vs } H_a : \text{ al menos una } (\tau\beta)_{ij} \neq 0.$$

Las hipótesis anteriores se contrastan a través de un análisis de varianza (ANVA) en dos sentidos.

Análisis de varianza para el modelo con el término de efecto de interacción

En el Cuadro 9 se presentan los resultados del ANVA obtenidos para el modelo lineal con factores cultivar y dosis de N ([5], [8]).

De acuerdo con los resultados del análisis de varianza se puede concluir que con un nivel de significancia del 5 % hay evidencia estadística de que existe efecto de las dosis de N, en tanto que el tipo de cultivar de maíz, no afecta significativamente la LTRS. Además, se puede observar que estadísticamente no existe efecto de interacción para la LTRS entre Tipo de Cultivar y Dosis de Nitrógeno.

El diagnóstico del modelo propuesto se evalúa con los valores obtenidos de los estadísticos S (desviación estándar) y R^2 (R-cuadrada), en este caso los valores de los estadísticos son: $S = 27.42$ y $R^2 = 40.07\%$. Lo cual indica que el modelo propuesto es moderadamente adecuado pues sólo explica un 40 % de la variabilidad en la respuesta.

Después del ANVA, en el análisis de medias para probar la hipótesis de igualdad de medias se utiliza el procedimiento ANOM (Análisis de Medias) mediante el paquete estadístico Minitab. El resultado se presenta en la Figura 1.

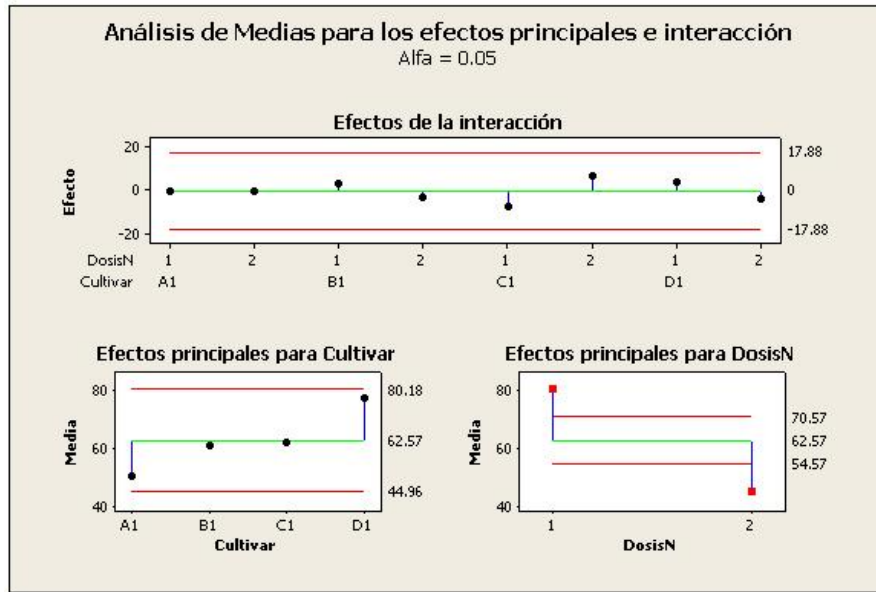


Figura 1: Análisis de medias para probar igualdad de medias para los efectos de interacción, cultivar y dosis.

El análisis de medias bajo el supuesto de normalidad muestra los efectos de interacción, los efectos principales para el primer factor y los efectos principales para el segundo factor. Las gráficas de ANOM tienen una línea central y límites de decisión. Si un punto se ubica fuera de los límites de decisión, entonces existe una evidencia significativa de que la media representada por ese punto es diferente de la media principal. Con un ANOM de dos factores, primero se busca los efectos de interacción. Si existe evidencia significativa para la interacción, usualmente no tiene sentido considerar los efectos principales porque el efecto de un factor depende del nivel del otro factor.

En este caso, los efectos de interacción se encuentran plenamente dentro de los límites de decisión, lo que significa que no hay evidencia de interacción. Ahora, se pueden observar los efectos principales. Las dos gráficas de la parte inferior de la Figura 1 muestran las medias de los niveles de los dos factores, donde se observa que el efecto principal es la diferencia entre la media y la línea central. Como los puntos que representan las medias de los cuatro niveles del factor cultivar, se encuentran dentro de los límites de decisión, esto indica que no existe evidencia significativa de que la media de cuatro niveles sea diferente de la media principal con un nivel de significancia $\alpha = 0.05$. Por otra parte, los efectos principales para los niveles 1 y 2 del factor dosis de N se encuentran fuera de los límites de decisión en la gráfica inferior derecha, lo que significa que existe evidencia de que estas medias son diferentes de la media principal con un nivel de significancia $\alpha = 0.05$.

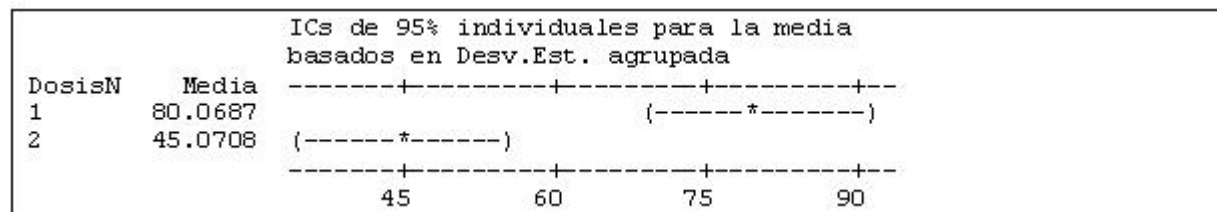


Figura 2: Intervalos de confianza individuales de 95 % para la media de dosis de N, basados en una prueba simple de medias con la desviación estándar agrupada.

En seguida se presenta un análisis gráfico para probar la diferencia de tratamientos para el factor dosis de N, asimismo, se realiza una comparación de medias, presentando los intervalos de confianza correspondientes a cada media.

En la Figura 2 se presenta el análisis para “dosis de N” con un nivel de significancia de 5 % que indica el ANVA.

En el caso de los intervalos de confianza para la media de dosis de N se puede observar que hay evidencia estadística de que para un nivel de significancia del 5 % la respuesta promedio no es la misma si se aplican diferentes dosis de nitrógeno, ya que los intervalos de confianza del 95 % no se intersectan.

Idoneidad del modelo

A continuación se realiza un análisis estadístico para determinar si los supuestos del modelo del ANVA se cumplen. En la Figura 3 se observa que el supuesto de normalidad es satisfecho dado que los puntos residuales vs porcentaje, se distribuyen de manera cercana alrededor de la recta, además de que el estadístico de Anderson-Darling para probar normalidad proporciona un p-valor de 0.476. La Figura 4 muestra que no hay evidencias contra el supuesto de varianzas iguales de los errores en el modelo, puesto que los puntos del valor ajustado contra el residuo simple se distribuyen sin seguir un patrón determinado; en la Figura 5 se presentan los intervalos de confianza de Bonferroni y pruebas de Bartlett y Levene para verificar igualdad de varianzas, observándose que no existe evidencia estadística en contra de la igualdad de varianzas al considerar un tamaño de prueba del 5 %.

Análisis de varianza para el modelo sin el término de efecto de interacción

Respecto al análisis de varianza de los factores principales (sin efecto de interacción): cultivar vs dosis de N se puede señalar, considerando el ANVA cuyo resultado se encuentra en el Cuadro 10, que con un nivel de significancia del 5 % no hay evidencia estadística de que exista efecto de tratamiento entre los cultivares, no obstante, si

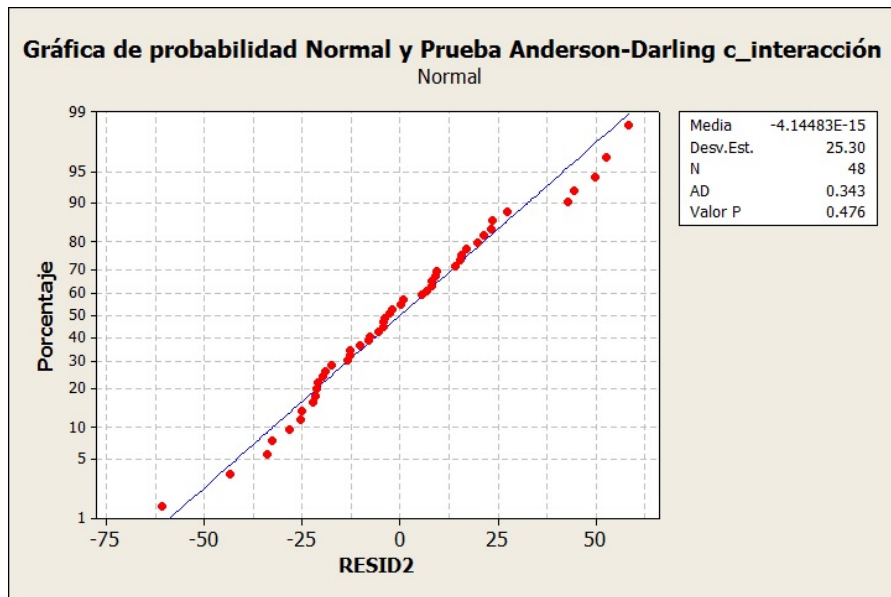


Figura 3: Prueba de normalidad para los residuos en el modelo lineal propuesto con interacción.

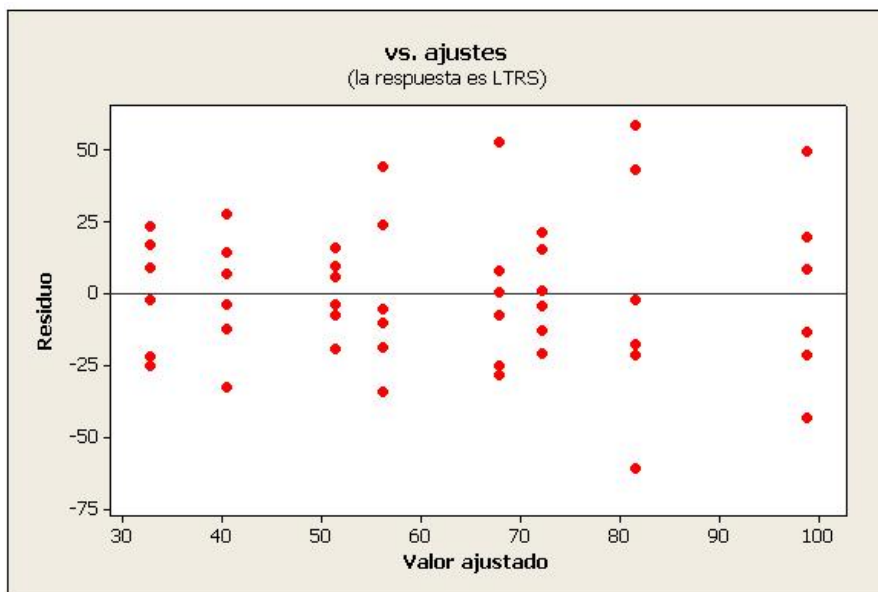


Figura 4: Valores ajustados de los residuales vs los valores residuales simples para el modelo lineal propuesto.

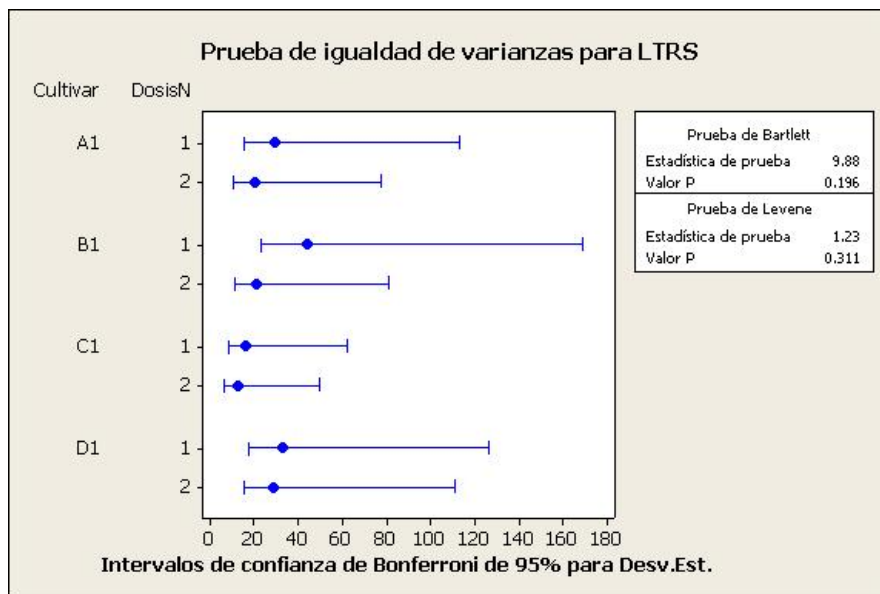


Figura 5: Intervalos de confianza de Bonferroni y pruebas de Bartlett y Levene para verificar igualdad de varianzas.

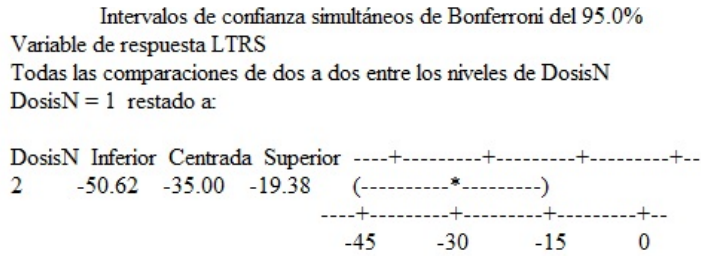
Cuadro 10: ANVA de dos factores: cultivar vs. dosis de N (20 % y 100 %), con variable respuesta LTRS sin término de efecto de interacción.

Fuente	GL	SC	MC	F	p-valor
Cultivar	3	4520.7	1506.9	2.09	0.115
Dosis de N	1	14698.3	14698.3	20.41	0.000
Error	43	30963.8	720.1		
Total	47	50182.8			

hay efecto de tratamiento para las dosis de N sobre la variable respuesta LTRS. Se trabaja con el modelo sin interacción debido a que anteriormente se ha probado que no existe efecto de interacción y a que las pruebas de hipótesis para los efectos principales se pueden utilizar para determinar si los factores afectan el resultado de la variable LTRS.

En la Figura 6 se realiza una comparación múltiple de medias con un nivel de significancia del 5 % para el factor dosis de N, presentando los intervalos de confianza simultáneos de Bonferroni y las pruebas simultaneas de Bonferroni.

Relativo a los intervalos de confianza para la media de dosis de N, para el caso del modelo sin interacción, se puede indicar que hay evidencia estadística de que existe efecto de tratamiento al considerar un nivel de significancia del 5 %, para el



Pruebas simultáneas de Bonferroni

Variable de respuesta LTRS

Todas las comparaciones de dos a dos entre los niveles de DosisN

DosisN = 1 restado a:

DosisN	Diferencia de medias	EE de diferencia	Valor T	Valor P ajustado
2	-35.00	7.746	-4.518	0.0000

Figura 6: Intervalos de confianza simultáneos de 95 % y pruebas simultáneas del 5 % de Bonferroni para la media de Dosis de N, para el modelo sin interacción.

factor dosis de N, debido a que el intervalo de confianza del 95 % de la diferencia de los niveles al 100 % y 20 % del factor DosisN no contiene al cero. Observándose un mayor efecto sobre la variable de respuesta LTRS de la dosis de nitrógeno al 20 % que el de la dosis de nitrógeno al 100 %.

Idoneidad del modelo sin interacción

En la Figura 7 se observa que el supuesto de normalidad para el término del error para el modelo sin interacción es satisfecho dado que los puntos residuales vs porcentaje, se distribuyen de manera cercana alrededor de la recta, además de que el estadístico de Anderson-Darling para probar normalidad proporciona un p-valor de 0.522.

En la Figura 8 se muestra que el supuesto de homogeneidad de varianzas de los errores en el modelo sin interacción, efectivamente se cumple, puesto que los puntos del valor ajustado contra el residuo simple se distribuyen sin seguir un patrón determinado. Además la prueba de Levene para igualdad de varianzas, considerando como variable respuesta a LTRS y como factores a Cultivar y DosisN, no es significativa al 5 % dado que el p-valor obtenido en dicha prueba es de 0.311.

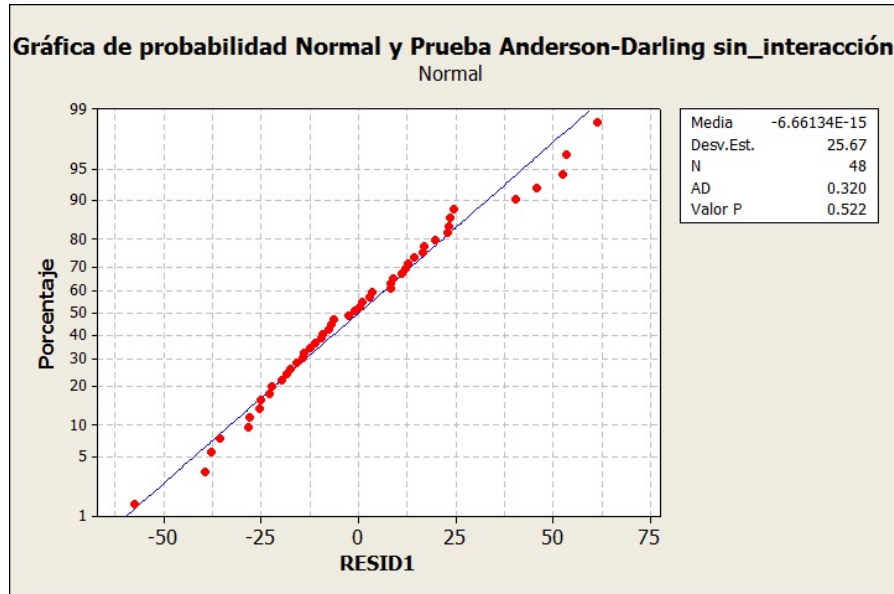


Figura 7: Prueba de normalidad para los residuos en el modelo lineal sin interacción.

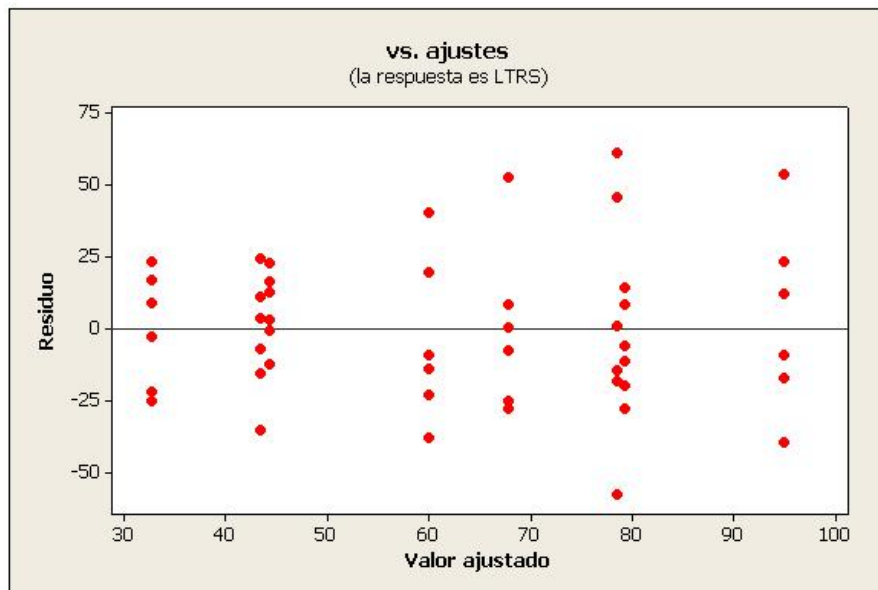


Figura 8: Valores ajustados de los residuales vs los valores residuales simples para el modelo lineal sin interacción.

Conclusiones

El análisis de componentes principales con 92 % de variabilidad acumulada muestra que para evaluar la respuesta morfológica de plántulas de maíz a los 21 días de germinación bajo condiciones de 20 % y 100 % de nitrógeno sólo se necesita considerar la primera de las componentes principales.

La variable que proporciona mayor información en esta primera CP (72 %) es la longitud total de raíces secundarias, lo cual permite sin mucha pérdida de información considerar sólo a esta variable para realizar el ANVA para probar las hipótesis de igualdad de tratamientos principales e interacción.

El ANVA muestra que no existe efecto de interacción entre los factores Dosis de Nitrógeno y Tipo de Cultivar, y sólo el factor Dosis de Nitrógeno resulta significativo.

Con un p-valor de aproximadamente cero, los datos proporcionan evidencia de que el promedio de la variable respuesta no es el mismo para las dosis 20 % y 100 % de nitrógeno, y según el análisis de comparación de medias, para el modelo sin interacción, se obtiene un resultado a favor de la dosis del 20 % de nitrógeno al considerar un nivel de significancia del 5 %.

Agradecimientos

A la Vicerrectoría de Investigación y Estudios de Posgrado por el financiamiento del proyecto “Uso eficiente de nitrógeno en maíz”.

Bibliografía

- [1] Adesemoye A.O., Torbert H.A. and Klopper J.W., *Increased plant uptake of nitrogen from N-depleted fertilizer using plant growth-promoting rhizobacteria*, Applied Soil Ecology, **46** (2010), 54–58.
- [2] Below F., *Fisiología, nutrición y fertilización nitrogenada del maíz*, Informaciones Agronómicas, **54** (2002), 3–9.
- [3] INEGI, *Cartas de división política y carta de municipios. Instituto Nacional de Estadística y Geografía*, (2010), México.
- [4] Dalgaard P., *Introductory statistics with R*. Springer, 2008.
- [5] Grima Cintas P., Marco Almagro L., Tort-Martorell Llabrés X., *Estadística Práctica con MINITAB*. Pearson Educación, S. A., 2004.

- [6] Hoagland D.R. and Arnon D. I., *The water-culture method for growing plants without soil. California, Agricultural Experiment Station, Circular*, **347** (1950), 1–32.
- [7] Johnson D.E., *Métodos multivariados aplicados al análisis de datos*. International Thomson Editores, 2000.
- [8] Minitab 15, *Statistical Software*. Minitab Inc., 2008.
- [9] Montgomery D.C., *Diseño y análisis de experimentos*. Grupo Editorial Iberoamérica, 1999.
- [10] R-i386 3.0.2, *Statistical Software*. The R foundations for statistical computing, 2013.
- [11] William N., *Statistics for engineers and scientists*. International Thomson Editores, 2006.
- [12] Woll K., Borsuk L., Stransky H., Nettleton D., Schnable P.S. and Hochholdinger F. , *Isolation characterization and pericycle specific transcriptome analyses of the novel maize (Zea mays L.) lateral and seminal root initiation mutant rum1*, *Plant Physiol.*, **139** (2005), 1255–1267.
- [13] FAO, *Statistical data bases*, (2013), <http://faostat3.fao.org>
- [14] SIAP, *Cierre de la producción agrícola por estado*, (2014), <http://www.siap.gob.mx>

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

DICA-Instituto de Ciencias, BUAP
14 Sur No. 6301, Colonia San Manuel,
Puebla, Pue. C.P. 72570

Department of Agronomy, Iowa State University
Ames, IA 50011, USA

bjuares@fcfm.buap.mx
jenaro.reyes@correo.buap.mx
thomasl@iastate.edu

Topología

Capítulo 9

Continuos localmente conexos sin hiperespacio único

$$C_n(X)$$

José Gerardo Ahuatzí Reyes, David Herrera Carrasco, Fernando Macías Romero
FCFM-BUAP

Resumen

Un continuo es un espacio métrico compacto, conexo y no vacío. Un continuo X es casi enrejado si el conjunto de los puntos de X que poseen una vecindad que es una gráfica finita es denso en X . Dado un continuo X y un número natural n , el n -ésimo hiperespacio de X , denotado por $C_n(X)$, es el conjunto formado por los subconjuntos cerrados de X dotado de la métrica de Hausdorff. En este capítulo se detallan las demostraciones de que los continuos localmente conexos que no son continuos casi enrejados, así como una subclase de los continuos localmente conexos casi enrejados, no tienen hiperespacio único $C_n(X)$.

1 Introducción

Un **continuo** es un espacio métrico compacto, conexo y no vacío. Dado un continuo X , se denota por 2^X al espacio $\{A \subset X : A \text{ es un cerrado de } X\}$ dotado de la topología de Vietoris o, equivalentemente, dotado con la métrica de Hausdorff. Un **hiperespacio** de X es cualquier subespacio de 2^X . Para cualquier número natural n , el n -ésimo **hiperespacio** de X es el subespacio de 2^X

$$C_n(X) = \{A \in 2^X : A \text{ tiene a lo más } n \text{ componentes}\};$$

en el caso $n = 1$ se acostumbra omitir el subíndice, es decir, $C(X)$ es el **hiperespacio de subcontinuos** de X . Se denomina **cubo de Hilbert** a cualquier espacio topológico homeomorfo al producto topológico de una cantidad numerable de intervalos cerrados. Denotaremos por \mathbb{N} al conjunto de los números naturales y por \mathbb{R} al conjunto de los números reales. Dado un espacio topológico Y y $B \subset Y$, la cerradura de B en Y se denota por $\text{cl}_Y(B)$ y el interior de B en Y , por $\text{int}_Y(B)$. Los conceptos no definidos en este trabajo se consideran como en [14], [13] y [4].

Dada una clase de hiperespacios \mathcal{H} para la cual a cada continuo X le corresponde el hiperespacio $\mathcal{H}(X)$, es natural preguntarse si esta correspondencia es biunívoca

(topológicamente). Es sencillo probar que si X y Y son continuos homeomorfos, entonces $C_n(X)$ y $C_n(Y)$ (así como 2^X y 2^Y) son homeomorfos. El recíproco de la afirmación anterior es mucho más interesante y se sabe que no siempre se cumple. Un resultado clásico de Curtis y Schori [5, 6] muestra que $C(X)$ es homeomorfo al cubo de Hilbert siempre que X es un continuo localmente conexo sin arcos libres (véase la Definición 2.7). Asimismo, G. Acosta y D. Herrera [1] muestra que para cada dendrita (véase la Definición 2.6), X , cuyo conjunto de puntos extremos no es cerrado (véase la Definición 9.3 y el Ejercicio 6.25 de [14]) existe una cantidad no numerable de dendritas distintas (no homeomorfas) entre sí y a X tales que su hiperespacio de subcontinuos es homeomorfo a $C(X)$. Considérese el siguiente concepto.

Definición 1.1. Sea X un continuo y n un número natural. Decimos que X tiene **hiperespacio único** $C_n(X)$ si para cada continuo Y tal que $C_n(Y)$ es homeomorfo a $C_n(X)$, se cumple que X es homeomorfo a Y .

En estos términos, [5, 6] y [1] muestran que tanto los continuos localmente conexos que no contienen arcos libres como las dendritas cuyo conjunto de puntos extremos no es cerrado no tienen hiperespacio único $C(X)$. En este trabajo se presentan pruebas detalladas de dos resultados análogos a los dos anteriores, presentados por R. Hernández Gutiérrez, A. Illanes y V. Martínez de la Vega [8], para el caso general de los n -ésimos hiperespacios y para dos clases de continuos: los continuos que no son casi enrejados (véase la Definición 2.8) y una subclase de los continuos casi enrejados. Estos resultados se enuncian en los Teoremas 6.6 y 7.1.

2 Gráficas finitas, dendritas y continuos casi enrejados

En este primer apartado vamos a introducir las clases de continuos que trataremos a lo largo de este capítulo, así como algunos resultados básicos relacionados con dichas clases.

Definición 2.1. Sea X un espacio topológico. X es un **arco** si es homeomorfo al intervalo cerrado $[0, 1]$. X es una **curva cerrada simple** si es homeomorfo a la circunferencia unitaria del plano euclideo. Si X es un arco, el conjunto de puntos extremos de X , es decir, el conjunto que bajo un homeomorfismo de X en $[0, 1]$ corresponde al conjunto $\{0, 1\}$, se denota por $E(X)$.

Definición 2.2. Un continuo X es una **gráfica finita** si es la unión de una familia finita de arcos tales que cada par de ellos o son ajenos o se intersectan en uno o dos de sus puntos extremos.

Lema 2.3 ([14], Proposición 9.2). *Si X y Y son gráficas finitas, tales que $X \cap Y$ es un conjunto finito y no vacío, entonces $X \cup Y$ es una gráfica finita.*

Definición 2.4. Un continuo X es un **árbol** si es una gráfica finita y no contiene curvas cerradas simples.

Lema 2.5. *Sea X un espacio topológico arco conexo. Si F es un subconjunto finito de X , entonces existe un árbol T contenido en X , tal que $F \subset T$.*

Demostración. Demostraremos este lema por inducción sobre el número de elementos de F . El caso $n = 2$ se sigue directamente de que X es arco conexo. Supongamos que el lema se cumple para todo conjunto finito con a lo más n elementos. Sean F un subconjunto de X con n elementos, x un punto en X y T un árbol contenido en X con $F \subset T$. Si $x \in T$, entonces $F \cup \{x\} \subset T$. Supongamos entonces que $x \notin T$. Obsérvese que existe un arco α contenido en X , tal que x es punto extremo de α y $\alpha \cap T = \{q\}$, donde q es el punto extremo de α distinto de x . Por el Lema 2.3, se tiene que $T \cup \alpha$ es una gráfica finita. Ahora, mostremos que $T \cup \alpha$ es un árbol.

En efecto, supongamos que $T \cup \alpha$ contiene una curva cerrada simple S . Como T es un árbol y α es un arco, S no está contenido ni en T ni en α . Además, como T y α son compactos, los conjuntos $T - \alpha = (T \cup \alpha) - \alpha$ y $\alpha - T = (T \cup \alpha) - T$ son subconjuntos abiertos y ajenos de $T \cup \alpha$ que intesectan a S . En esta forma, el conjunto $S - (T \cap \alpha) = S - \{q\}$ es desconexo. Esto es una contradicción pues una curva cerrada simple no se hace desconexa al quitarle solo uno de sus puntos. Por lo tanto, $T \cup \alpha$ es un árbol. Así, se concluye la demostración. \square

Definición 2.6. Una **dendrita** es un continuo localmente conexo que no contiene curvas cerradas simples.

Definición 2.7. Sea X un continuo. Denotamos

$$\mathcal{G}(X) = \{p \in X : p \text{ tiene una vecindad } M \text{ en } X, \text{ tal que}$$

$$M \text{ es una gráfica finita}\},$$

$$\mathcal{P}(X) = X - \mathcal{G}(X).$$

Definición 2.8. Un continuo X es **casi enrejado** si $\mathcal{G}(X)$ es denso en X .

Lema 2.9. *Sean X un continuo y U un subconjunto de X . Si X es casi enrejado, entonces $\text{int}_X(U) \subset \text{cl}_X(U - \mathcal{P}(X))$.*

Demostración. Supongamos que $x \in \text{int}_X(U)$. Sea V un subconjunto abierto de X , tal que $x \in V$. Luego, $\text{int}_X(U) \cap V$ es un subconjunto abierto y no vacío de X . Además, como X es casi enrejado, $\text{int}_X(\mathcal{P}(X)) = \emptyset$. De esta manera $\text{int}_X(U) \cap V \not\subset \mathcal{P}(X)$; es decir, $V \cap (\text{int}_X(U) - \mathcal{P}(X)) = (V \cap \text{int}_X(U)) - \mathcal{P}(X) \neq \emptyset$. Así, $V \cap (U - \mathcal{P}(X)) \supset V \cap (\text{int}_X(U) - \mathcal{P}(X)) \neq \emptyset$. Por tanto, $x \in \text{cl}_X(U - \mathcal{P}(X))$. \square

Una formulación equivalente de continuo casi enrejado puede obtenerse del siguiente resultado.

Definición 2.10. Sea X un continuo. Un arco A contenido en X es un **arco libre** de X si $A - E(A)$ es un abierto de X .

Definición 2.11. Sea X un continuo. Denotamos

$$\mathcal{FA}(X) = \bigcup \{ \text{int}_X(J) : J \text{ es un arco libre de } X \}.$$

Lema 2.12. Si X es un continuo, entonces $\text{cl}_X(\mathcal{G}(X)) = \text{cl}_X(\mathcal{FA}(X))$. Por lo tanto, un continuo X es casi enrejado si, y solo si, $\mathcal{FA}(X)$ es denso en X .

Demostración. Sea X un continuo con métrica d . Como $\mathcal{FA}(X) \subset \mathcal{G}(X)$, basta con mostrar que $\text{cl}_X(\mathcal{G}(X)) \subset \text{cl}_X(\mathcal{FA}(X))$.

Supongamos que $x \in \text{cl}_X(\mathcal{G}(X))$. Luego, existe una sucesión $\{x_n\}_{n \in \mathbb{N}}$ en $\mathcal{G}(X)$, tal que $\lim x_n = x$. Fijemos $n \in \mathbb{N}$. Como $x_n \in \mathcal{G}(X)$, existe una gráfica finita G contenida en X , tal que $x_n \in \text{int}_X(G)$. Sean $k \in \mathbb{N}$ y arcos A_1, \dots, A_k en G , tales que $\bigcup_{i=1}^k A_i = G$ y $A_i \cap A_j \subset E(A_i) \cap E(A_j)$ si $i \neq j$. Sea $j_0 \in \{1, \dots, k\}$, tal que $x_n \in A_{j_0}$.

Obsérvese que $V = \text{int}_X(G) \cap B_d(x_n, \frac{1}{n})$ es una vecindad de x_n en X . Además, es posible hallar un subarco B de A_{j_0} , tal que $x \in B \subset V$.

Obsérvese que $B - E(B) \subset A_{j_0} - E(A_{j_0}) \subset G - \bigcup_{i \neq j_0} A_i$. Luego, $G - (B - E(B)) = \bigcup_{i \neq j_0} A_i \cup (A_{j_0} - (B - E(B)))$. Además, $A_{j_0} - (B - E(B))$ es un subconjunto cerrado de A_{j_0} y, por tanto, de G . Así, $G - (B - E(B))$ es cerrado en G ; es decir, $B - E(B)$ es abierto en G . Como además $B - E(B) \subset \text{int}_X(G) \subset G$, se tiene que $B - E(B)$ es un subconjunto abierto de X . De este modo, B es un arco libre de X .

Sea $y_n \in B - E(B) \subset \text{int}_X(B) \subset \mathcal{FA}(X)$. Obsérvese que, para cada $n \in \mathbb{N}$, se cumple que $d(x_n, y_n) < \frac{1}{n}$. Así, $\lim y_n = x$, con cada y_n elemento de $\mathcal{FA}(X)$. Por lo tanto, $x \in \text{cl}_X(\mathcal{FA}(X))$. Esto concluye la prueba del lema. \square

3 Z-conjuntos y el Teorema de Toruńczyk

El Teorema 5.6, uno de los resultados que cimentan la demostración de los Teoremas 6.6 y 7.1, se basa en una caracterización del cubo de Hilbert debida a H. Toruńczyk (Teorema 3.8). En este apartado se presenta este último resultado así como algunos conceptos y resultados relacionados. Los primeros conceptos importantes que definimos son el de extensor absoluto y el de retracts absoluto. Resulta interesante el hecho que estas dos propiedades son equivalentes.

Definición 3.1. Sean X y Y espacios topológicos, $A \subset Y$ y $f : A \rightarrow X$ una función continua. Decimos que la función $F : X \rightarrow Y$ es una **extensión continua** de f a Y si F es continua y $F|_A = f$. Un espacio normal X es un **extensor absoluto** si, para cada espacio normal Y , cada subconjunto cerrado A de Y y cada función continua $f : A \rightarrow X$, se cumple que f tiene una extensión continua a Y .

Definición 3.2. Un subconjunto cerrado A de un espacio topológico Y es un **retracto** de Y si Id_A tiene una extensión continua a Y . Un espacio normal X es un **retracto absoluto** si, para cada espacio normal Y y cada subconjunto cerrado B de Y homeomorfo a X , se satisface que B es un retracto de Y .

Teorema 3.3 ([4], Teorema (4.B.19)). *Sea X un espacio topológico. Entonces, X es un retracto absoluto si, y solo si, X es un extensor absoluto.*

Para tratar con el Teorema de Toruńczyk, resulta conveniente la siguiente notación (en realidad la función d_∞ que se define a continuación resulta ser una métrica sobre el espacio de funciones continuas que van de X a Y , pero esta propiedad no es requerida aquí).

Definición 3.4. Dados X y Y espacios métricos, X con métrica acotada, y funciones continuas $f, g : Y \rightarrow X$, denotaremos

$$d_\infty(f, g) = \sup\{d(f(y), g(y)) : y \in Y\}.$$

Definición 3.5. Sean X y Y espacios métricos y $\{f_n\}_{n \in \mathbb{N}}$ una sucesión de funciones continuas de X en Y . Decimos que $f : X \rightarrow Y$ es el **límite uniforme** de $\{f_n\}_{n \in \mathbb{N}}$ si $\lim d_\infty(f_n, f) = 0$.

Definición 3.6. Sean X un espacio métrico compacto y A un subconjunto cerrado de X . Decimos que A es un **Z-conjunto** en X si Id_X es el límite uniforme de funciones continuas cuyas imágenes no intersectan a A . Decimos que una función continua f entre espacios métricos compactos X_1 y X_2 es una **Z-función** si $f(X_1)$ es un Z-conjunto en X_2 .

Al tratar con Z-conjuntos vamos a utilizar el siguiente lema, el cual garantiza que, al realizar ciertas operaciones sobre estos, se siguen obteniendo Z-conjuntos.

Lema 3.7. *Sea X un espacio métrico compacto. Entonces, se cumplen las siguientes afirmaciones:*

- (1) *Un subconjunto cerrado de un Z-conjunto en X es, a su vez, un Z-conjunto en X .*

(2) La unión finita de Z -conjuntos en X es un Z -conjunto en X .

Demostración. Probaremos (1). Supongamos que A es un Z -conjunto en X y B es un subconjunto cerrado de A . Sea $\varepsilon > 0$. Luego, existe una función continua $f_\varepsilon : X \rightarrow X - A$, tal que $d_\infty(f_\varepsilon, Id_X) < \varepsilon$. Obsérvese que $X - A \subset X - B$. De esta manera, podemos considerar que $f_\varepsilon : X \rightarrow X - B$. Además, B es un subconjunto cerrado de X . Por lo tanto, B es un Z -conjunto en X .

Ahora mostremos (2). Bastará probar la afirmación para dos elementos. Sean A_1 y A_2 dos Z -conjuntos en X . Tomemos una función continua $f_1 : X \rightarrow X - A_1$, tal que $d_\infty(f_1, Id_X) < \frac{\varepsilon}{2}$. Obsérvese que, dado $a \in A$, se cumple que $d(f_1(X), A_1) \leq d(f_1(a), a) \leq d_\infty(f_1, Id_X)$. Así, $d(f_1(X), A_1) < \frac{\varepsilon}{2}$. Como X es compacto, $f_1(X)$ es compacto. Además, $f_1(X) \cap A_1 = \emptyset$. Luego, $d(f_1(X), A_1) > 0$. Así, existe una función continua $f_2 : X \rightarrow X - A_2$, tal que $d_\infty(f_2, Id_X) < \frac{1}{2}d(f_1(X), A_1)$. Obsérvese que para cualquier $x \in X$ se cumple que $d(f_1(x), f_2(f_1(x))) < \frac{1}{2}d(f_1(X), A_1) < d(f_1(X), A_1)$ y, por tanto, $f_2(f_1(x)) \notin A_1$. Se sigue de esto último que $f_2(f_1(X)) \subset X - A_1$. Como también $f_2(X) \subset X - A_2$, podemos considerar la función continua $g = f_2 \circ f_1 : X \rightarrow X - (A_1 \cup A_2)$. Obsérvese también que $d(g(x), x) \leq d(g(x), f_1(x)) + d(f_1(x), x) < \frac{1}{2}d(f_1(X), A_1) + \frac{\varepsilon}{2} < \frac{\varepsilon}{4} + \frac{\varepsilon}{2} = \frac{3\varepsilon}{4}$, para cualquier $x \in X$. Luego, $d_\infty(g, Id_X) < \varepsilon$. Por tanto, $A_1 \cup A_2$ es un Z -conjunto en X . \square

Ahora, estamos en mejores condiciones para enunciar el Teorema de Toruńczyk.

Teorema 3.8 ([9], Teorema 9.3). *Sea X un retracto absoluto. Si la función identidad sobre X es un límite uniforme de Z -funciones, entonces X es un cubo de Hilbert.*

4 Hiperespacios de crecimiento

En este apartado se presenta otra clase de hiperespacios: los hiperespacios de crecimiento. Asimismo, se revisan algunas de sus propiedades que, junto con la propiedad de convexidad (dentro de los espacios localmente conexos), proveen de algunas características de la función Φ_ε (Definición 4.16) que se utilizan en las pruebas de los Teoremas 5.6, 6.6 y 7.1.

Definición 4.1. Sean X un continuo y \mathcal{B} un subconjunto cerrado de 2^X . Decimos que \mathcal{B} es un **hiperespacio de crecimiento** si para cualesquiera $A \in \mathcal{B}$ y $B \in 2^X$, tales que $A \subset B$ y cada componente de B intersecta a A , se tiene que $B \in \mathcal{B}$.

El primer resultado acerca de los hiperespacios de crecimiento es que contiene a los hiperespacios que interesan a este estudio, es decir, los n -ésimos hiperespacios y los hiperespacios que describe la siguiente definición. Esto se expresa en la Observación 4.3.

Definición 4.2. Sean X un continuo, n un número natural y K un subconjunto cerrado de X . Denotamos

$$C_n^K(X) = \{A \in C_n(X) : K \subset A\}$$

$$C_n(X, K) = \{A \in C_n(X) : K \cap A \neq \emptyset\}$$

Observación 4.3. Dado un continuo localmente conexo X , un subconjunto cerrado K de X y $n \in \mathbb{N}$, los hiperespacios $C_n(X)$, $C_n^K(X)$ y $C_n(X, K)$ son hiperespacios de crecimiento.

Demostración. Sean $A \in C_n(X)$ y $B \in 2^X$, tales que $A \subset B$ y cada componente de B intersecta a A . Sean A_1, \dots, A_k las componentes de A , con $k \leq n$. Como $A \subset B$, para cada $i \in \{1, \dots, k\}$ existe una componente B_i de B , tal que $A_i \subset B_i$. Luego, $A \subset \bigcup_{i=1}^k B_i$. Como las componentes de un conjunto cualquiera son ajenas por pares y cada componente de B intersecta a A , tenemos que B_1, \dots, B_k son todas las componentes de B . Así, $B \in C_n(X)$.

Si, además, $A \in C_n^K(X)$; es decir, $K \subset A$, entonces $K \subset A \subset B$ y $B \in C_n^K(X)$. Si, en lugar de tal condición, A es elemento de $C_n(X, K)$; en otras palabras, si $A \cap K \neq \emptyset$, entonces $B \cap K \supset A \cap K \neq \emptyset$ y $B \in C_n^K(X)$. Esto prueba que $C_n(X)$, $C_n^K(X)$ y $C_n(X, K)$ son hiperespacios de crecimiento. \square

La siguiente característica importante de los hiperespacios de crecimiento que vamos a demostrar es que estos espacios son retractos absolutos (Teorema 4.14). Para ello vamos a requerir algunas definiciones más, así como varios resultados previos.

Definición 4.4. Sean X un espacio con métrica d y $\varepsilon > 0$. Dado cualquier $A \in 2^X$, la d -bola cerrada generalizada en X de radio ε alrededor de A , es el conjunto $C_d(A, \varepsilon) = \{x \in X : d(x, A) \leq \varepsilon\}$.

Lema 4.5. Sea X un espacio métrico. Si A es un subespacio compacto de X y x es un elemento cualquiera de X , existe $a \in A$, tal que $d(x, A) = d(x, a)$.

Demostración. Dado cualquier $n \in \mathbb{N}$, sea $a_n \in A$, tal que $d(x, a_n) < d(x, A) + \frac{1}{n}$. Como A es compacto, podemos suponer que $\lim a_n = a$, para alguna $a \in A$. Obsérvese que, para cada $n \in \mathbb{N}$, se cumple que $d(x, A) \leq d(x, a_n) \leq d(x, A) + \frac{1}{n}$. Así, tomando límites, $d(x, A) = \lim d(x, a_n)$. Además, por la continuidad de d , obtenemos que $\lim d(x, a_n) = d(x, a)$. Por lo tanto, $d(x, A) = d(x, a)$. \square

Definición 4.6. Sean X y Y espacios con métricas d y e , respectivamente. Una **isometría** es un isomorfismo $h : X \rightarrow Y$, tal que, para cada $a, b \in X$, $e(h(a), h(b)) = d(a, b)$.

Definición 4.7. Sea X un espacio compacto con métrica d . Decimos que d es una **métrica convexa** para X si, para cada $p, q \in X$, existe una isometría $\gamma : [0, d(p, q)] \rightarrow \gamma([0, d(p, q)])$, tal que $\gamma(0) = p$ y $\gamma(d(p, q)) = q$.

Teorema 4.8 ([12], Teorema 4). *Sea X un continuo. Si X es localmente conexo, entonces X admite una métrica convexa.*

Aunque la definición de métrica convexa aquí mostrada es distinta a la dada en [12], la Proposición 10.4 de [9] garantiza que estas dos definiciones son equivalentes en el caso de los continuos (más aún, para espacios métricos compactos).

Lema 4.9. *Sean X, Y, D continuos y $p \in Y$, tales que $Y = X \cup D$ y $X \cap D = \{p\}$. Si d es una métrica convexa de Y , la métrica inducida en X por d es una métrica convexa.*

Demostración. Supongamos que d es una métrica convexa de Y . Sean $a, b \in X$. Luego, existe una isometría $\gamma : [0, d(a, b)] \rightarrow \gamma([0, d(a, b)]) \subset Y$, tal que $\gamma(0) = a$ y $\gamma(d(a, b)) = b$.

Si p es un punto extremo del arco $A = \gamma([0, d(a, b)])$, digamos $p = b$, entonces $A - \{p\}$ es conexo. Como $X \cap D = \{p\}$ y tanto X como Y son cerrados de Y , los conjuntos $X - \{p\} = X - D$ y $D - \{p\} = X - D$ son abiertos de Y ajenos entre sí. Además, $(X - \{p\}) \cup (D - \{p\}) = Y - \{p\}$. En consecuencia, $A - \{p\} \subset X - \{p\}$ o $A - \{p\} \subset D - \{p\}$. Como $a \in (X - \{p\}) \cap (A - \{p\})$, se tiene que $A - \{p\} \subset X - \{p\}$. Así, $A \subset X$.

Por otro lado, si p no es un punto extremo de A , entonces $A - \{p\}$ tiene dos componentes, cada una de ellas conteniendo un punto extremo de A . Sean K_a y K_b tales componentes, con $a \in K_a$ y $b \in K_b$. Como K_a y K_b son conexos y $a \in K_a \cap (X - \{p\})$ y $b \in K_b \cap (X - \{p\})$, podemos aplicar un argumento similar al aplicado a $A - \{p\}$ en el caso anterior, para obtener que $K_a, K_b \subset X - \{p\}$. Por tanto, $A = K_a \cup K_b \cup \{p\} \subset X$.

De este modo, en cualquier caso, se cumple que $A \subset X$; es decir, existe un arco de a a b contenido en X y que es isométrico a un intervalo cerrado. Esto prueba que la métrica inducida por d sobre X es convexa. \square

El siguiente lema no solo resulta útil en la prueba del Teorema 4.14, sino también para demostrar el Lema 4.17 y el Teorema 5.6.

Lema 4.10. *Sea X continuo. Si d es una métrica convexa para X y $\varepsilon > 0$, entonces:*

- (1) *Para cada $A \in 2^X$, se tiene $C_d(A, \varepsilon) = \text{cl}_X(\text{int}_X(C_d(A, \varepsilon)))$.*
- (2) *Si \mathcal{B} es un hiperespacio de crecimiento, entonces, para cada $A \in \mathcal{B}$, se cumple que $C_d(A, \varepsilon) \in \mathcal{B}$.*

(3) Para cada $n \in \mathbb{N}$ y $A \in C_n(X)$, se tiene $C_d(A, \varepsilon) \in C_n(X)$.

Demostración. (1) Primero mostremos que $\text{cl}_X(\text{int}_X(C_d(A, \varepsilon))) \subset C_d(A, \varepsilon)$. Obsérvese que $C_d(A, \varepsilon)$ contiene a $\text{int}_X(C_d(A, \varepsilon))$, por lo cual basta probar que $C_d(A, \varepsilon)$ es cerrado en X . Para ello, tomemos $x \in X - C_d(A, \varepsilon)$. Luego, $d(x, A) > \varepsilon$. Sea $\delta > 0$, tal que $d(x, A) - \varepsilon > 2\delta$. Dado cualquier $y \in B_d(x, \delta)$ se tiene que $d(x, A) \leq d(x, y) + d(y, A) < \delta + d(y, A)$. Así, $d(y, A) > d(x, A) - \delta > \delta + \varepsilon$, siempre que $a \in A$. Por tanto, $d(y, A) \geq \delta + \varepsilon > \varepsilon$; es decir, $y \in X - C_d(A, \varepsilon)$. Esto muestra que $B_d(y, \delta) \subset X - C_d(A, \varepsilon)$ y que $C_d(A, \varepsilon)$ es cerrado en X . Obsérvese que esta parte de la prueba no requiere ninguna condición adicional al continuo X .

Ahora probaremos que $C_d(A, \varepsilon) \subset \text{cl}_X(\text{int}_X(C_d(A, \varepsilon)))$. Para tal efecto, tomemos $x \in C_d(A, \varepsilon)$. Si $x \in A$, entonces $x \in N_d(A, \varepsilon) \subset \text{int}_X(C_d(A, \varepsilon))$. Podemos suponer entonces que $x \notin A$. Por el Lema 4.5, existe $a \in A$, tal que $d(x, a) = d(x, A)$. Como d es una métrica convexa para X , existe una isometría $\gamma : [0, d(a, x)] \rightarrow X$, tal que $\gamma(0) = a$ y $\gamma(d(a, x)) = x$. Obsérvese que para cualquier $t \in [0, d(a, x))$ se cumple que $d(a, \gamma(t)) = d(\gamma(0), \gamma(t)) = |t - 0| = t < d(a, x) = d(x, A) \leq \varepsilon$. Esto implica que $\gamma(t) \in B_d(a, \varepsilon)$. En otras palabras, $\gamma([0, d(a, x))) \subset B_d(a, \varepsilon) \subset C_d(A, \varepsilon)$. Así, $\gamma([0, d(a, x))) \subset \text{int}_X(C_d(A, \varepsilon))$ y $x \in \gamma([0, d(a, x)]) = \text{cl}_X(\gamma([0, d(a, x)))) \subset \text{cl}_X(\text{int}_X(C_d(A, \varepsilon)))$. Esto muestra que $C_d(A, \varepsilon) \subset \text{cl}_X(\text{int}_X(C_d(A, \varepsilon)))$. Más aún, prueba que cada elemento de $C_d(A, \varepsilon) - A$ puede ser unido a un elemento de A por un arco contenido en $C_d(A, \varepsilon)$. Por lo tanto, se da la igualdad de (1).

(2) Mostraremos primero que si K es conexo, entonces $C_d(K, \varepsilon)$ es conexo. Para ello, supongamos que K es conexo. Dado $x \in C_d(K, \varepsilon) - A$, por el párrafo anterior, existen $a_x \in K$ y un arco α_x contenido en $C_d(K, \varepsilon)$ que une a a_x con x . De este modo, $C_d(K, \varepsilon) = K \cup (\bigcup \{\alpha_x : x \in C_d(K, \varepsilon) - K\})$. Como K es conexo y, para cada $x \in C_d(K, \varepsilon)$, se cumple que α_x es conexo y $a_x \in \alpha_x \cap K$, tenemos, por [7, Teorema 6.1.9], que $C_d(K, \varepsilon)$ es conexo.

Ahora, supongamos que $A \in \mathcal{B}$. Probaremos que $C_d(A, \varepsilon) \in \mathcal{B}$. Con tal motivo, sea C una componente de $C_d(A, \varepsilon)$. Sea $x \in C$, tal que $d(x, A) \leq \varepsilon$. Por el Lema 4.5, existe $a \in A$, tal que $d(x, a) = d(x, A) \leq \varepsilon$. Como $C_d(\{a\}, \varepsilon)$ está contenido en $C_d(A, \varepsilon)$ y, por el párrafo anterior, es conexo, tenemos que $C_d(\{a\}, \varepsilon) \subset C$. Así, $a \in C$ y $A \cap C \neq \emptyset$. Por tanto, cada componente de $C_d(A, \varepsilon)$ intersecta a A ; es decir, $C_d(A, \varepsilon) \in \mathcal{B}$. Esto prueba (3).

(3) Por la Observación 4.3, $C_n(X)$ es un hiperespacio de crecimiento. Aplicando (2), obtenemos el resultado.

Esto concluye la demostración del lema. □

Lema 4.11. Dado un continuo con métrica convexa d , la función $\Phi : 2^X \times [0, \infty) \rightarrow 2^X$ dada por

$$(A, r) \xrightarrow{\Phi} C_d(A, r)$$

es continua.

Demostración. Usaremos la métrica máx para $2^X \times \mathbb{R}$. Tomemos $\delta > 0$. Sean $A, B \in 2^X$ y $\varepsilon, \eta \leq 0$, tales que $H_d(A, B) < \frac{\delta}{4}$ y $|\varepsilon - \eta| < \frac{\delta}{4}$. Queremos demostrar que $H_d(C_d(A, \varepsilon), C_d(B, \eta)) < \delta$. Sea $x \in C_d(A, \varepsilon)$. Por el Lema 4.5, existe $a \in A$, tal que $d(x, a) = d(x, A) \leq \varepsilon$. Como $A \subset N_d(B, \frac{\delta}{4})$, existe $b \in B$, tal que $d(a, b) < \frac{\delta}{4}$. Luego, $d(x, b) \leq d(x, a) + d(a, b) < d(x, A) + \frac{\delta}{4} \leq \varepsilon + \frac{\delta}{4} < \frac{\delta}{4} + \eta + \frac{\delta}{4} = \eta + \frac{\delta}{2}$. Sea $\gamma : [0, d(b, x)] \rightarrow X$ una isometría. Si $d(b, x) \leq \eta$, entonces $x \in C_d(B, \eta)$. Podemos suponer entonces que $d(b, x) > \eta$. Entonces, $d(b, \gamma(\eta)) = d(\gamma(0), \gamma(\eta)) = \eta$. Así, $\gamma(\eta) \in C_d(B, \varepsilon)$ y $d(x, \gamma(\eta)) = d(\gamma(d(b, x)), \gamma(\eta)) = d(b, x) - \eta \leq \eta + \frac{\delta}{2} - \eta = \frac{\delta}{2}$. Por tanto, $x \in N_d(C_d(B, \eta), \frac{\delta}{2})$. De esta manera, $C_d(A, \varepsilon) \subset N_d(C_d(B, \eta), \frac{\delta}{2})$. Del mismo modo, se puede probar que $C_d(B, \eta) \subset N_d(C_d(A, \varepsilon), \frac{\delta}{2})$. Esto muestra que $H_d(C_d(A, \varepsilon), C_d(B, \eta)) < \delta$. Así, se concluye que Φ es continua. \square

Los siguientes dos teoremas son resultados muy notables relacionados con el cubo de Hilbert. No obstante su gran importancia, en este escrito el uso de estos en este capítulo se limita a la prueba del Teorema 4.14.

Teorema 4.12 ([5], Teorema 1). *Sea X un continuo. Entonces, X es un continuo localmente conexo si, y solo si, 2^X es homeomorfo al cubo de Hilbert.*

Teorema 4.13 ([9], Teorema 9.2). *Sea X un espacio métrico compacto. Si X es un retracto del cubo de Hilbert, entonces X es un retracto absoluto.*

Teorema 4.14. *Sea X un continuo localmente conexo. Si \mathcal{B} es un hiperespacio de crecimiento, entonces \mathcal{B} es un retracto absoluto.*

Demostración. Sea Φ la función continua definida en el Lema 4.11. Dado $A \in 2^X$, se define el valor t_A como sigue. Como X es compacto, X es acotado según la métrica d . Luego, $\Phi(A, 2 \text{ diám}(X)) = X$. Obsérvese que el conjunto $\Phi^{-1}(\mathcal{B}) \cap (\{A\} \times [0, \infty)) = \{A\} \times \{t \in [0, \infty) : \Phi(A, t) \in \mathcal{B}\}$ es cerrado en $2^X \times [0, \infty)$ y contiene al menos al conjunto $\{(A, 2 \text{ diám}(X))\}$. Así, $\{t \in [0, \infty) : \Phi(A, t) \in \mathcal{B}\}$ es un subconjunto cerrado y no vacío de $[0, \infty)$. Sea $t_A = \min\{t \in [0, \infty) : \Phi(A, t) \in \mathcal{B}\}$.

Mostraremos que la función $g : 2^X \rightarrow [0, \infty)$ dada por $A \mapsto t_A$ es continua. Para ello, sean $\delta > 0$ y $A, B \in 2^X$, tales que $H(A, B) < \frac{\delta}{2}$.

Afirmamos que $C_d(A, t_A) \subset C_d(B, t_A + \frac{\varepsilon}{2})$. En efecto, sea $x \in C_d(A, t_A)$, es decir, $x \in X$, tal que $d(x, A) \leq t_A$. Por el Lema 4.5, existe $a \in A$, tal que $d(x, a) = d(x, A) \leq t_A$. Como $A \subset N_d(B, \frac{\varepsilon}{2})$, existe $b \in B$, tal que $d(a, b) < \frac{\varepsilon}{2}$. Luego, $d(x, b) \leq d(x, a) + d(a, b) < t_A + \frac{\varepsilon}{2}$; es decir, $x \in C_d(B, t_A + \frac{\varepsilon}{2})$. Esto prueba nuestra afirmación.

Ahora, mostremos que cada componente de $C_d(B, t_A + \frac{\delta}{2})$ intersecta a $C_d(A, t_A)$. Sea C una componente de $C_d(B, t_A + \frac{\delta}{2})$. Tomemos $y \in C$. Luego, por el Lema 4.5,

existe $b \in B$, tal que $d(b, y) \leq t_A + \frac{\varepsilon}{2}$. Como $B \subset B(A, \frac{\varepsilon}{2})$, existe $a \in A$, tal que $d(a, b) < \varepsilon$. Obsérvese que $y, a \in C_d(\{b\}, t_A + \frac{\varepsilon}{2})$. Además, por el Lema 4.10 (3), este último conjunto es conexo. Por tanto, $C_d(\{b\}, t_A + \frac{\varepsilon}{2}) \subset C$. Así, $a \in C$ y $C_d(A, t_A) \cap C \supset A \cap C \neq \emptyset$.

Como \mathcal{B} es un hiperespacio de crecimiento y $C_d(A, t_A) \in \mathcal{B}$, los dos párrafos anteriores muestran que $C_d(B, t_A + \frac{\varepsilon}{2}) \in \mathcal{B}$. Por tanto, $t_B \leq t_A + \frac{\varepsilon}{2}$. De la misma forma, se puede probar que $t_A \leq t_B + \frac{\varepsilon}{2}$. Así, $-\frac{\varepsilon}{2} \leq t_A - t_B \leq \frac{\varepsilon}{2}$; es decir, $|t_A - t_B| \leq \frac{\varepsilon}{2} < \varepsilon$. Esto muestra que g es continua.

De esta forma, la función $\Theta : 2^X \rightarrow \mathcal{B}$ dada por $A \mapsto \Phi(A, t_A)$ es continua. Además, $\Theta(A) = \Phi(A, 0) = A$ siempre que $A \in \mathcal{B}$. De este modo, Θ es una retracción de 2^X en \mathcal{B} . Además, por el Teorema 4.12, el hiperespacio 2^X es homeomorfo al cubo de Hilbert. Por tanto, \mathcal{B} es un retracto del cubo de Hilbert. Aplicando el Teorema 4.13, concluimos que \mathcal{B} es un retracto absoluto. \square

Una consecuencia fundamental del teorema anterior es que los hiperespacios que más nos interesan son retractos absolutos, como lo afirma el siguiente corolario.

Corolario 4.15. *Sean X un continuo localmente conexo, K un subconjunto cerrado de X y $n \in \mathbb{N}$. Si X es localmente conexo, entonces los hiperespacios $C_n(X)$, $C_n^K(X)$ y $C_n(X, K)$ son retractos absolutos.*

Demostración. Es una consecuencia inmediata de la Observación 4.3 y del Teorema 4.14. \square

Como último punto de este apartado, vamos a mostrar algunas propiedades de la función Φ_ε , la cual se define a continuación.

Definición 4.16. Dado un espacio métrico X y $\varepsilon > 0$, la función $\Phi_\varepsilon : 2^X \rightarrow 2^X$ está dada por

$$A \mapsto \Phi_\varepsilon(A) = C_d(A, \varepsilon).$$

Lema 4.17. *Sea X un espacio métrico. Si la métrica de X es convexa, entonces para cualquier $\varepsilon > 0$ se cumplen las siguientes tres afirmaciones:*

- (1) *Si \mathcal{B} es un hiperespacio de crecimiento, entonces podemos restringir $\Phi_\varepsilon|_{\mathcal{B}} : \mathcal{B} \rightarrow \mathcal{B}$.*
- (2) *Φ_ε es continua.*
- (3) *$(H_X)_\infty(Id_{2^X}, \Phi_\varepsilon) \leq \varepsilon$.*

Demostración. (1) es consecuencia inmediata del Lema 4.10 (3).

(2) es una implicación directa del Lema 4.11.

(3) Dado $A \in 2^X$, probemos que $H_d(A, \Phi_\varepsilon(A)) \leq \varepsilon$. Para tal fin, sea $\eta > 0$. Obsérvese que $A \subset C_d(A, \varepsilon) = \Phi_\varepsilon(A) \subset N_d(\Phi_\varepsilon(A), \varepsilon + \eta)$. Por otro lado, dada $x \in \Phi_\varepsilon(A)$, por el Lema 4.5, existe $a \in A$, tal que $d(x, a) = d(x, A)$. Luego, $d(x, a) \leq \varepsilon < \varepsilon + \eta$ y $x \in N_d(A, \varepsilon + \eta)$. Esto implica que $\Phi_\varepsilon(A) \subset N_d(A, \varepsilon + \eta)$. Por lo tanto, $H_d(A, \Phi_\varepsilon(A)) \leq \varepsilon + \eta$. Como η es un número positivo arbitrario, deducimos que $H_d(A, \Phi_\varepsilon(A)) \leq \varepsilon$. Al ser A un elemento cualquiera de 2^X , concluimos la desigualdad en (3). \square

5 Cubos de Hilbert en $C_n(X)$

En este apartado vamos mostrar una forma muy general de hallar copias del cubo de Hilbert en el n -ésimo hiperespacio de cualquier continuo localmente conexo X para el cual $\mathcal{P}(X) \neq \emptyset$ (Teorema 5.6, presentado originalmente en [8]).

Teorema 5.1 ([4], Teorema 1.F.6). *Sean X y Y espacios topológicos y $f : X \rightarrow Y$ una función. Si A y B son subconjuntos cerrados de X , tales que $X = A \cup B$ y las restricciones $f|_A : A \rightarrow Y$ y $f|_B : B \rightarrow Y$ son continuas, entonces f es continua.*

Teorema 5.2 ([13], Lema 1.48). *Sea X un continuo. Entonces, se satisface lo siguiente:*

(1) *Si \mathcal{A} es un subconjunto cerrado y no vacío de 2^X , entonces $\bigcup \mathcal{A} \in 2^X$.*

(2) *Para cualesquiera $A, B \in 2^X$ se cumple que $H_d(\bigcup \mathcal{A}, \bigcup \mathcal{B}) \leq H_{H_d}(A, B)$.*

(3) *La función $\Psi : 2^{2^X} \rightarrow 2^X$, dada por $\mathcal{A} \mapsto \bigcup \mathcal{A}$, es continua.*

Teorema 5.3 ([11], Definición 1.8.21 y Teorema 1.8.22). *Si $f : X \rightarrow Y$ es una función continua entre los espacios métricos compactos X y Y , entonces la función $f^* : 2^X \rightarrow 2^Y$ dada por*

$$A \xrightarrow{f^*} f(A)$$

es continua y $f^(C_n(X)) \subset C_n(Y)$, para cualquier $n \in \mathbb{N}$.*

El siguiente resultado no es difícil de probar.

Lema 5.4. *Sea X un espacio métrico. Entonces, la función $f : X \rightarrow F_1(X)$, dada por $x \xrightarrow{f} \{x\}$, es un homeomorfismo.*

En la demostración del siguiente lema, se hará uso de la notación B_d^X para especificar que los conjuntos B_d se toman sobre el espacio X . En otras palabras, dados un subespacio X de algún espacio con métrica d , un punto $p \in X$ y $\varepsilon > 0$, denotaremos $B_d^X(p, \varepsilon) = B_d(p, \varepsilon) \cap X$.

Lema 5.5. *Sean X un continuo localmente conexo, R un subconjunto cerrado de $\mathcal{P}(X)$ y $K \in C(X)$, tal que $\text{int}_X(K) \cap R \neq \emptyset$. Entonces, $C_n^K(X)$ es un Z -conjunto de $C_n(X, R)$.*

Demostración. Sea d una métrica para d . Obsérvese primero que, si $A \in C_n^K(X)$, entonces $R \cap A \supset R \cap K \supset R \cap \text{int}_X(K) \neq \emptyset$; es decir, $A \in C_n(X, R)$. Como $C_n^K(X)$ es compacto y $C_n(X, R)$ es un espacio métrico, lo anterior implica que $C_n^K(X)$ es un subconjunto cerrado de $C_n(X, R)$.

Fijemos $p \in \text{int}_X(K) \cap R$ y sea $\varepsilon > 0$, tal que $B_d^X(p, \varepsilon) \subsetneq \text{int}_X(K)$. Por [14, Teorema 8.10], existen $m \in \mathbb{N}$ y subcontinuos localmente conexos X_1, \dots, X_m de X , tales que $X = \bigcup_{i=1}^m X_i$ y $\text{diám}(X_i) < \frac{\varepsilon}{4}$, para cada $i \in \{1, \dots, m\}$. Sean $G = \{i \in \{1, \dots, m\} : p \in X_i\}$ y $S = \bigcup_{i \in G} X_i$. Como $p \in \bigcap_{i \in G} X_i$, se tiene que, para cada $x \in S$, $d(p, x) < \frac{\varepsilon}{4}$; es decir, $S \subset B_d^X(p, \frac{\varepsilon}{4}) \subsetneq X$.

Sean $F = \{i \in \{1, \dots, m\} : p \notin X_i \text{ y } X_i \cap S \neq \emptyset\}$ y $H = \{i \in \{1, \dots, m\} : X_i \cap S = \emptyset\}$. Obsérvese que S y $\bigcup_{i \in H} X_i$ son subconjuntos cerrados y ajenos de X y $F \cup G \cup H = \{1, \dots, m\}$. Si $F = \emptyset$, como $S \subsetneq X$, se tiene que $\bigcup_{i \in H} X_i \neq \emptyset$ y $X = S \cup (\bigcup_{i \in H} X_i)$, lo cual contradice la conexidad de X . Así, $F \neq \emptyset$.

Para cada $i \in F$, fijemos $p_i \in X_i \cap S$. Aplicando [14, Ejercicio 8.5], tenemos que S es un continuo localmente conexo y, por [14, Teorema 8.26], S es arco conexo. Luego, por el Lema 2.5, existe un árbol T contenido en S , tal que $\{p\} \cup \{p_i : i \in F\} \subset T$.

Sea $Y = T \cup (\bigcup_{i \in F} X_i)$. Como para cada $i \in F$, $p_i \in T \cap X_i$ se tiene, por [14, Ejercicio 8.51], que Y es un continuo localmente conexo. Obsérvese que, para cada $i \in F$ y $x \in X_i$ se cumple $d(x, p) \leq d(x, p_i) + d(p, p_i) \leq \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{\varepsilon}{2}$; es decir, $\bigcup_{i \in F} X_i \subset B_d^X(p, \frac{\varepsilon}{2})$. Luego, $Y \subset B_d^X(p, \frac{\varepsilon}{2}) \subset \text{int}_X(K)$.

Sea $Z = Y \cap R$. Obsérvese que $p \in Z$ y, por el Corolario 4.15 y el Teorema 3.3, $C(Y)$ y $C(Y, Z)$ son extensores absolutos.

Consideremos las funciones $\varphi : X \rightarrow C(X)$, dada por $x \mapsto \{x\}$; $\alpha = \varphi|_Y : Y \rightarrow C(Y)$ y $\beta = \varphi|_Z : Z \rightarrow C(Y, Z)$. Por el Teorema 5.4, φ es continua y, por tanto, α y β son continuas. Como Z es un subconjunto cerrado de $(S \cup Y) \cap R$, podemos extender β a una función continua $\bar{\beta} : (S \cup Y) \cap R \rightarrow C(Y, Z)$. Obsérvese que $\bar{\beta}|_Z = \beta|_Z = \alpha|_Z$, $(S \cup Y) \cap R$ y Y son subconjuntos cerrados de $((S \cup Y) \cap R) \cup Y = (S \cap R) \cup Y$ y $((S \cup Y) \cap R) \cap Y = Y \cap R = Z$. De esta manera, por el Lema 5.1, la función $\alpha \cup \bar{\beta} : (S \cap R) \cup Y \rightarrow C(Y, Z)$ dada por

$$x \xrightarrow{\alpha \cup \bar{\beta}} \begin{cases} \alpha(x) & \text{si } x \in Y, \\ \bar{\beta}(x) & \text{si } x \in (S \cup Y) \cap R \end{cases}$$

es continua. Como $(S \cap R) \cup Y$ es un subconjunto cerrado de $S \cup Y$, podemos extender $\alpha \cup \bar{\beta}$ a una función continua $\bar{\alpha} : S \cup Y \rightarrow C(Y)$.

Consideremos ahora la función $\gamma : X \rightarrow C(X)$ dada por

$$x \xrightarrow{\gamma} \begin{cases} \bar{\alpha}(x) & \text{si } x \in S \cup Y, \\ \varphi(x) & \text{si } x \in X - (S \cup Y). \end{cases}$$

Como $X - (S \cup Y) = X - \bigcup_{i \in G \cup F} X_i \subset \bigcup_{i \in H} X_i$, se tiene que $\text{cl}_X(X - (S \cup Y)) \subset \bigcup_{i \in H} X_i$. Luego, si $A = \text{cl}_X(X - (S \cup Y)) \cap (S \cup Y)$, entonces $A \subset (S \cup Y) \cap (\bigcup_{i \in H} X_i) = (Y - S) \cap (\bigcup_{i \in H} X_i) \subset Y$. Así, $\gamma|_A = \bar{\alpha}|_A = \alpha|_A = \varphi|_A$. Como $\text{cl}_X(X - (S \cup Y)) = (X - (S \cup Y)) \cup A$, esto último implica que $\gamma|_{\text{cl}_X(X - (S \cup Y))} = \varphi|_{\text{cl}_X(X - (S \cup Y))}$. Por tanto, por el Lema 5.1, γ es continua.

Ahora, definiremos la función g_ε . Obsérvese que, para cada $A \in C_n(X)$, si A_0 es una componente de A , entonces $\gamma(A_0)$ es un subcontinuo de $C(X)$. Aplicando [3, Lema 2.2], se tiene que $\bigcup \gamma(A_0) \in C(X)$. Como $\bigcup \gamma(A) = \bigcup \{\bigcup \gamma(A_0) : A_0 \text{ es una componente de } A\}$, y A tiene a lo más n componentes, deducimos que $\bigcup \gamma(A) \in C_n(X)$. De esta manera, podemos considerar la función

$$g_\varepsilon : C_n(X) \rightarrow C_n(X) \text{ dada por } A \xrightarrow{g_\varepsilon} \bigcup g(A).$$

Obsérvese que, por los Teoremas 5.2 y 5.3, las funciones $\Psi : 2^{2^X} \rightarrow 2^X$ y $\gamma^* : 2^X \rightarrow 2^{C(X)}$ dadas por $\mathcal{A} \xrightarrow{\Psi} \bigcup \mathcal{A}$ y $A \xrightarrow{\gamma^*} \gamma(A)$ son continuas. Como $g_\varepsilon(B) = \Psi \circ \gamma^*(B)$, para cada $B \in C_n(X)$, se tiene que g_ε es una función continua.

Tomemos $A \in C_n(X)$. Dado $x \in A$, si $x \in S \cup Y$, entonces $\gamma(x) = \bar{\alpha}(x) \in C(Y) \subset C(S \cup Y)$ y $\{x\} \in C(S \cup Y)$. Como $S \cup Y \subset B_d^X(p, \frac{\varepsilon}{2})$, se tiene que $H_X(\{x\}, \gamma(x)) \leq \text{diám}(S \cup Y)$. Si, por el contrario, $x \notin S \cup Y$, entonces $\gamma(x) = \{x\}$, y, consecuentemente, $H_X(\{x\}, \gamma(x)) = 0 \leq \text{diám}(S \cup Y)$. De esta manera $H_{H_d}(\{\{x\} : x \in A\}, \{\gamma(x) : x \in A\}) \leq \text{diám}(S \cup Y) < \varepsilon$. Aplicando el Teorema 5.2 (2) obtenemos que $H_X(A, g_\varepsilon(A)) = H_X(\bigcup_{x \in A} \{x\}, \bigcup_{x \in A} \gamma(x)) < \varepsilon$.

Por otro lado, $X - S \subset \bigcup_{i \in F \cup H} X_i$ así que $\text{cl}_X(X - S) \subset \bigcup_{i \in F \cup H} X_i$. Luego, $\text{Fr}_X(S) = S \cap \text{cl}_X(X - S) \subset S \cap (\bigcup_{i \in F \cup H} X_i) = S \cap (\bigcup_{i \in F} X_i) \subset \bigcup_{i \in F} X_i$ y, por ende, $S - \bigcup_{i \in F} X_i = \text{int}_X(S) \cap (X - \bigcup_{i \in F} X_i)$. Así, $S - \bigcup_{i \in F} X_i$ es un subconjunto abierto de X . Como $p \in S - \bigcup_{i \in F} X_i$ y $p \in \mathcal{P}(X)$, tenemos que $S - \bigcup_{i \in F} X_i \not\subseteq T$. Tomemos $s \in (S - \bigcup_{i \in F} X_i) - T = S - Y \subset K$. Luego, dada $x \in A$, si $x \in S \cup Y$, entonces $s \notin Y \supset \bar{\alpha}(x) = \gamma(x)$; es decir, $s \notin \gamma(x)$, y si $x \notin S \cup Y$, entonces $s \notin \{x\} = \gamma(x)$. Así, $s \notin \bigcup_{x \in A} \gamma(x) = g_\varepsilon(x)$ y, por tanto, $K \not\subseteq g_\varepsilon(A)$.

De esta manera $(H_X)_\infty(\text{Id}_{C_n(X)}, g_\varepsilon) < \varepsilon$ y $g_\varepsilon : C_n(X) \rightarrow C_n(X) - C_n^K(X)$.

Más aún, supongamos que $A \in C_n(X, R)$ y fijemos $a \in A \cap R$. Si $a \in X - (S \cup Y)$, entonces $\gamma(a) = \{a\}$ y, por consiguiente, $g_\varepsilon(A) \cap R \supset \gamma(a) \cap R = \{a\} \neq \emptyset$. Si $a \in S \cup Y$, entonces $a \in (S \cup Y) \cap R$. En consecuencia, $\gamma(a) = \bar{\beta}(a) \in$

$C(Y, Z)$, por lo cual $g_\varepsilon(A) \cap R \supset \gamma(a) \cap Z \neq \emptyset$. De cualquier manera, $g_\varepsilon(A) \in C_n(X, R)$. Por lo tanto, $g_\varepsilon|_{C_n(X, R)} : C_n(X, R) \rightarrow C_n(X, R) - C_n^K(X)$. Como, además, $(H_X)_\infty(Id_{C_n(X, R)}, g_\varepsilon|_{C_n(X, R)}) < \varepsilon$ y ε puede ser tomado arbitrariamente pequeño, concluimos que $C_n^K(X)$ es un Z-conjunto de $C_n(X, R)$. \square

En la siguiente demostración se hace uso de la notación B_d^X , utilizada en el lema previo.

Teorema 5.6. *Sean X un continuo localmente conexo y R un subconjunto cerrado de $\mathcal{P}(X)$. Entonces, $C_n(X, R)$ es un cubo de Hilbert.*

Demostración. Obsérvese en primer lugar que, por el Corolario 4.15, el hiperespacio $C_n(X, R)$ es un retracto absoluto. De esta manera, para aplicar el Teorema 3.8, basta mostrar que la función $Id_{C_n(X, R)}$ es un límite uniforme de Z-funciones.

Sea $\varepsilon > 0$. Por el Teorema 4.8, podemos suponer que la métrica para X es convexa. Además, por la Observación 4.3, $C_n(X, R)$ es un hiperespacio de crecimiento. Aplicando el Lema 4.17 obtenemos que la restricción $\Phi_\varepsilon|_{C_n(X, R)} : C_n(X, R) \rightarrow C_n(X, R)$ es continua y que

$$(H_X)_\infty(Id_{C_n(X, R)}, \Phi_\varepsilon|_{C_n(X, R)}) < \varepsilon.$$

Como R es compacto, existen $s \in \mathbb{N}$ y $p_1, \dots, p_s \in R$, tales que

$$R \subset \bigcup_{i=1}^s B_d^X(p_i, \frac{\varepsilon}{2}).$$

Para cada $i \in \{1, \dots, s\}$, sea $K_i = C_d(\{p_i\}, \frac{\varepsilon}{2})$. Luego, $R \subset \bigcup_{i=1}^s K_i$. Por otro lado, para cada $i \in \{1, \dots, s\}$, se cumple que $p_i \in \text{int}_X(K_i) \cap R$ y, por el Lema 4.10, $K_i \in C(X)$. Aplicando el Lema 5.5, obtenemos que $C_n^{K_i}(X)$ es un Z-conjunto en $C_n(X, R)$. De esta manera, por el Lema 3.7, el conjunto $\mathcal{C} = \bigcup_{i=1}^s C_n^{K_i}(X)$ es un Z-conjunto en $C_n(X, R)$.

Sea $A \in C_n(X, R)$. Entonces, $A \cap (\bigcup_{i=1}^s K_i) \supset A \cap R \neq \emptyset$; es decir, existe $j \in \{1, \dots, s\}$, tal que $A \cap K_j \neq \emptyset$. Sea $t \in A \cap K_j$. Luego, $K_j = C_d(\{p_j\}, \frac{\varepsilon}{2}) \subset C_d(\{t\}, \varepsilon) \subset C_d(A, \varepsilon) = \Phi_\varepsilon(A)$ y, en consecuencia, $\Phi_\varepsilon(A) \in C_n^{K_j}(X)$. Por tanto, $\Phi_\varepsilon(C_n(X, R)) \subset \mathcal{C}$.

Como $C_n(X, R)$ es compacto, $\Phi_\varepsilon(C_n(X, R))$ es compacto y, por ende, un subconjunto cerrado de \mathcal{C} . Luego, aplicando el Lema 4.10, se obtiene que $\Phi_\varepsilon(C_n(X, R))$ es un Z-conjunto en $C_n(X, R)$; es decir, $\Phi_\varepsilon|_{C_n(X, R)}$ es una Z-función. Por lo tanto, $Id_{C_n(X, R)}$ es un límite uniforme de Z-funciones. Así, se concluye, por el Teorema 3.8, que $C_n(X, R)$ es un cubo de Hilbert. \square

6 Los continuos que no son casi enrejados no tienen hiperespacio único

En esta sección se presenta la demostración de que cualquier continuo que no es casi enrejado no tiene n -ésimo hiperespacio único.

Observación 6.1. Para cualesquiera continuos X y Z , tales que $Z \subset X$, se satisface que $\mathcal{P}(Z) \subset \mathcal{P}(X)$ o, equivalentemente, $\mathcal{G}(X) \cap Z \subset \mathcal{G}(Z)$.

Demostración. Sea $x \in \mathcal{G}(X) \cap Z$. Luego, es posible hallar una colección finita de arcos libres de X , \mathcal{A} , tal que $x \in \text{int}_X(\bigcup \mathcal{A})$ y, si \mathcal{A} tiene más de un elemento, entonces cada par de elementos de \mathcal{A} se intersecta únicamente en x . Sea $A \in \mathcal{A}$. Si $x \in A - E(A)$, es posible obtener dos subarcos A_1 y A_2 de A tales que $x \in E(A_1) \cap E(A_2)$, $A_1 \cap A_2 = \{x\}$ y $A_1 \cup A_2 = A$. Así, podemos suponer que $x \in E(A)$. Supongamos que $A - E(A)$ no está contenido en Z ni en $X - Z$. Como $A - E(A)$ es conexo, este conjunto debe intersectar $\text{Fr}_X(Z)$. Si $\text{Fr}_X(Z) \cap (A - E(A))$ tiene tres o más elementos, entonces existe un subarco B de A tal que sus puntos extremos no intersectan a Z y $B \cap Z \neq \emptyset$. Esto contradice la conexidad de Z , pues $B - E(B)$ y $X - B$ son abiertos ajenos que intersectan a Z . Luego, podemos hacer $\text{Fr}_X(Z) \cap (A - E(A)) = \{p, q\}$. Si $p = q$, entonces $(C - E(C)) \cap \text{Fr}_X(Z) = \emptyset$, donde C es el subarco de A con puntos extremos p y x . Como $(C - E(C))$ es conexo, esto último implica que $C - E(C) \subset Z$ o $C - E(C) \subset X - Z$. Si $p \neq q$, entonces podemos suponer que el subarco de A con puntos extremos p y x no contiene a q . Luego, si nombramos D a dicho subarco, se sigue, de manera similar al caso anterior, que $D - E(D) \subset Z$ o $D - E(D) \subset X - Z$. En cualquier caso, es posible hallar un subarco F_A de A tal que $F_A \cap Z = \{p\}$ o bien $F_A \subset Z$. Si $\mathcal{F} = \{F_A : A \in \mathcal{A}\}$, entonces $\bigcup \mathcal{F}$ es una vecindad de x en X tal que $(\bigcup \mathcal{F}) \cap Z$ es una gráfica finita. Así, $x \in \mathcal{G}(Z)$. \square

El siguiente resultado es una interesante propiedad de extensión que posee el cubo de Hilbert dada por R. Anderson. Se utiliza en la prueba del Lema 6.5.

Teorema 6.2 ([2], Teorema 10.1). *Si $h : A \rightarrow B$ es un homeomorfismo entre Z -conjuntos en un cubo de Hilbert \mathcal{Q} , entonces h se puede extender a un homeomorfismo de \mathcal{Q} sobre \mathcal{Q} .*

Teorema 6.3 ([8], Teorema 18). *Sean X un continuo localmente conexo y $p \in X$. Entonces, existe una familia no numerable \mathcal{D} de dendritas ajenas y no homeomorfas por pares, tal que se satisfacen las siguientes condiciones:*

(a) *Para cada $D \in \mathcal{D}$, se tiene que D no contiene arcos libres.*

(b) El continuo localmente conexo $X \cup_p D$ no es homeomorfo a X .

(c) Si B y D son elementos distintos de \mathcal{D} , entonces $X \cup_p B$ y $X \cup_p D$ no son homeomorfos.

Lema 6.4. Sean X, Y y D continuos y $p \in Y$, tales que $Y = X \cup D$ y $X \cap D = \{p\}$. Si E es un subconjunto cerrado de X que contiene a p , entonces se satisfacen las siguientes igualdades:

$$(a) C_n(X) \cap C_n(Y, E \cup D) = C_n(X, E),$$

$$(b) C_n(Y) - C_n(Y, E \cup D) = C_n(X) - C_n(X, E) \text{ y}$$

$$(c) \text{Fr}_{C_n(X)}(C_n(X, E)) = \text{Fr}_{C_n(Y)}(C_n(Y, E \cup D)).$$

Demostración. Demostraremos primero que $C_n(X) \cap C_n(Y, E \cup D) = C_n(X, E)$. Como $C_n(X, E) \subset C_n(Y, E) \subset C_n(Y, E \cup D)$, resta probar una contención. Supongamos para ello que $A \in C_n(X) \cap C_n(Y, E \cup D)$. Luego, $A \subset X$, $\emptyset \neq A \cap (E \cup D)$ y

$$\begin{aligned} A \cap (E \cup D) &= (A \cap E) \cup (A \cap D) \\ &= (A \cap E) \cup (A \cap X \cap D) \\ &= (A \cap E) \cup (A \cap \{p\}) \\ &= A \cap (E \cup \{p\}) \\ &= A \cap E; \end{aligned}$$

es decir, $A \in C_n(X, E)$. Esto muestra la primera igualdad.

Ahora probemos que $C_n(Y) - C_n(Y, E \cup D) = C_n(X) - C_n(X, E)$. Supongamos que $A \in C_n(Y) - C_n(Y, E \cup D)$. Luego, $A \subset Y$ y $A \cap D = \emptyset$; es decir, $A \subset Y - D \subset X$. Además, $A \cap E = \emptyset$. Así, $A \in C_n(X) - C_n(X, E)$. Recíprocamente, si $A \in C_n(X) - C_n(X, E)$, entonces

$$\begin{aligned} A \cap (E \cup D) &= (A \cap E) \cup (A \cap D) \\ &= A \cap X \cap D \\ &= A \cap \{p\} \subset A \cap E = \emptyset, \end{aligned}$$

por lo cual $A \in C_n(Y) - C_n(Y, E \cup D)$. Esto demuestra la igualdad en (b).

Por último, mostraremos la igualdad en (c). Como $C_n(X)$ es un subconjunto cerrado de $C_n(Y)$ y $C_n(Y) - C_n(Y, E \cup D) \subset C_n(X)$, se tiene que

$$\begin{aligned} \text{cl}_{C_n(Y)}(C_n(Y) - C_n(Y, E \cup D)) &= \text{cl}_{C_n(X)}(C_n(X) - C_n(X, E)) \\ &\subset C_n(X). \end{aligned}$$

Además, $C_n(X, E)$ es un subconjunto cerrado de $C_n(X)$ y $C_n(Y, E \cup D)$ es un subconjunto cerrado de $C_n(Y)$. Por lo tanto,

$$\begin{aligned} \text{Fr}_{C_n(X)}(C_n(X, E)) &= C_n(X, E) \cap \text{cl}_{C_n(X)}(C_n(X) - C_n(X, E)) \\ &= (C_n(X) \cap C_n(Y, E \cup D)) \cap \\ &\quad \text{cl}_{C_n(Y)}(C_n(Y) - C_n(Y, E \cup D)) \\ &= C_n(Y, E \cup D) \cap \text{cl}_{C_n(Y)}(C_n(Y) - C_n(Y, E \cup D)) \\ &= \text{Fr}_{C_n(Y)}(C_n(Y, E \cup D)). \end{aligned}$$

Esto concluye la demostración. \square

Lema 6.5. Sean X, Y, D continuos, $p \in Y$ y E un subconjunto cerrado de X , tales que $Y = X \cup D$, $X \cap D = \{p\}$ y $p \in E$. Si $C_n(X, E)$ y $C_n(Y, E \cup D)$ son cubos de Hilbert, $\text{Fr}_{C_n(X)}(C_n(X, E))$ es un Z -conjunto en $C_n(X, E)$ y $\text{Fr}_{C_n(Y)}(C_n(Y, E \cup D))$ es un Z -conjunto en $C_n(Y, E \cup D)$, entonces $C_n(X)$ es homeomorfo a $C_n(Y)$.

Demostración. Por el Lema 6.4, $\text{Fr}_{C_n(X)}(C_n(X, E)) = \text{Fr}_{C_n(Y)}(C_n(Y, E \cup D))$, por lo cual la función identidad

$$id : \text{Fr}_{C_n(X)}(C_n(X, E)) \rightarrow \text{Fr}_{C_n(Y)}(C_n(Y, E \cup D))$$

es un homeomorfismo. Aplicando el Teorema 6.2 podemos extender id a un homeomorfismo $h_1 : C_n(X, E) \rightarrow C_n(Y, E \cup D)$. Consideremos la función $h : C_n(X) \rightarrow C_n(Y)$ dada por

$$A \xrightarrow{h} \begin{cases} h_1(A) & \text{si } A \in C_n(X, E), \\ Id_{C_n(X)}(A) & \text{si } A \in C_n(X) - C_n(X, E). \end{cases}$$

Obsérvese que, si $C = C_n(X, E) \cap \text{cl}_{C_n(X)}(C_n(X) - C_n(X, E))$, entonces $C = \text{Fr}_{C_n(X)}(C_n(X, E))$, de manera que $h_1|_C = id = Id_{C_n(X)}|_C$. Así, por el Lema 5.1, obtenemos que h es continua. Además, de nuevo por el Lema 6.4, $C_n(X) - C_n(X, E) = C_n(Y) - C_n(Y, E \cup D)$. Luego,

$$\begin{aligned} Id_{C_n(X)}(C_n(X) - C_n(X, E)) &= C_n(Y) - C_n(Y, E \cup D) \\ &= C_n(Y) - h_1(C_n(X, E)) \end{aligned}$$

y, como $Id_{C_n(X)}$ y h_1 son inyectivas, se tiene que h es inyectiva. Más aún,

$$\begin{aligned} h(C_n(X)) &= h_1(C_n(X, E)) \cup Id_{C_n(X)}(C_n(X) - C_n(X, E)) \\ &= C_n(Y, E \cup D) \cup (C_n(Y) - C_n(Y, E \cup D)) \\ &= C_n(Y); \end{aligned}$$

es decir, h es sobreyectiva. Así, se concluye así que h es un homeomorfismo. \square

En la siguiente demostración se utiliza de nuevo la notación B_d^X , detallada en los comentarios previos al Lema 5.5. Además, se utiliza la notación N_d^X para denotar los conjuntos N_d , pero considerando la métrica inducida en el subespacio X por d , en vez de d misma. Es decir, dados un subespacio X de algún espacio con métrica d , un punto $A \in 2^X$ y $\varepsilon > 0$, denotaremos $N_d^X(A, \varepsilon) = N_d(A, \varepsilon) \cap X$.

Teorema 6.6. *Sea X un continuo localmente conexo. Si X no es casi enrejado, entonces, para cada $n \in \mathbb{N}$, se tiene que X no tiene hiperespacio único $C_n(X)$.*

Demostración. Sea $n \in \mathbb{N}$. Como X no es casi enrejado, $X - \mathcal{P}(X)$ no es denso en X ; es decir, existe $p \in \text{int}_X(\mathcal{P}(X))$. Por el Teorema 6.3, existe un continuo localmente conexo D que no contiene arcos libres, tal que el continuo localmente conexo $Y = X \cup_p D$ no es homeomorfo a X .

Mostremos ahora que $C_n(X)$ y $C_n(Y)$ son homeomorfos. Para tal efecto, por el Teorema 4.8, podemos suponer que la métrica, d , de Y es convexa. Esto implica, por el Lema 4.9, que las métricas inducidas por d en X y D son convexas. Sea $r > 0$, tal que $B_d^X(p, 2r) \subset \mathcal{P}(X)$. Hagamos $E = C_d(\{p\}, r)$. Probaremos las siguientes tres afirmaciones.

Afirmación 1. $C_m(Y, E \cup D)$ y $C_m(X, E)$ son cubos de Hilbert.

Obsérvese que E es un subconjunto cerrado de X y $E \subset B_d^X(p, 2r) \subset \mathcal{P}(X)$; es decir, E es un subconjunto cerrado de $\mathcal{P}(X)$. Como D no contiene arcos libres, por el Lema 2.12, se cumple que $\text{cl}_D(\mathcal{G}(D)) = \text{cl}_D(\mathcal{FA}(D)) = \emptyset$. Así, $\mathcal{G}(D) = \emptyset$ y $\mathcal{P}(D) = D$. Además, $E \cup D$ es un subconjunto cerrado de Y y, por la Observación 6.1, se cumple que $E \cup D \subset \mathcal{P}(X) \cup \mathcal{P}(D) \subset \mathcal{P}(Y)$, esto es, $E \cup D$ es un subconjunto cerrado de $\mathcal{P}(Y)$. Así, por el Teorema 5.6, $C_n(X, E)$ y $C_n(Y, E \cup D)$ son cubos de Hilbert.

Afirmación 2. $\text{Fr}_{C_m(X)}(C_m(X, E))$ es un Z -conjunto en $C_m(X, E)$.

Sea $\varepsilon > 0$. Por la Observación 4.3, $C_n(X, E)$ es un hiperespacio de crecimiento. Aplicando el Lema 4.17, obtenemos que la función

$$\Phi_\varepsilon|_{C_n(X, E)} : C_n(X, E) \rightarrow C_n(X, E)$$

es continua y cumple que $(H_X)_\infty(\text{Id}_{C_n(X, E)}, \Phi_\varepsilon|_{C_n(X, E)}) < \varepsilon$. Además, dada $A \in C_n(X, E)$, como $N_d^X(A, \varepsilon) \cap \text{cl}_X(\text{int}_X(E)) \supset A \cap E \neq \emptyset$, se tiene que

$$\Phi_\varepsilon(A) \cap \text{int}_X(E) \supset N_d^X(A, \varepsilon) \cap \text{int}_X(E) \neq \emptyset.$$

Sean $p_A \in \Phi_\varepsilon(A) \cap \text{int}_X(E)$ y $r_A > 0$, tales que $B_d^X(p_A, r_A) \subset \text{int}_X(E) \subset E$. Luego, si $\{A_n\}_{n \in \mathbb{N}}$ es una sucesión en $C_n(X)$ que converge a $\Phi_\varepsilon(A)$, entonces existe $N \in \mathbb{N}$, tal que, para cada $n > N$, se tiene $E \cap A_n \supset B_d^X(p_A, r_A) \cap A_n \neq \emptyset$;

es decir, $A_n \in C_n(X, E)$. Así, $\Phi_\varepsilon(A) \notin \text{cl}_{C_n(X)}(C_n(X) - C_n(X, E))$ y, por ende, $\Phi_\varepsilon(A) \notin \text{Fr}_{C_n(X)}(C_n(X, E))$. En esta forma,

$$\Phi_\varepsilon|_{C_n(X, E)} : C_n(X, E) \rightarrow C_n(X, E) - \text{Fr}_{C_n(X)}(C_n(X, E)).$$

Por tanto, $\text{Fr}_{C_n(X)}(C_n(X, E))$ es un Z-conjunto en $C_n(X, E)$.

Afirmación 3. $\text{Fr}_{C_m(Y)}(C_m(Y, E \cup D))$ es un Z-conjunto en $C_m(Y, E \cup D)$.

Obsérvese que $p \in \text{int}_X(E) \cap \text{int}_D(D)$, por lo cual $p \in \text{int}_Y(E \cup D)$. Además, $X - \{p\} = Y - D$ es abierto en Y . Así,

$$\begin{aligned} E &= \text{cl}_X(\text{int}_X(E)) = \text{cl}_Y(\text{int}_X(E)) = \text{cl}_Y((\text{int}_X(E) - \{p\}) \cup \{p\}) \\ &\subset \text{cl}_Y(\text{int}_{X-\{p\}}(E - \{p\}) \cup \{p\}) = \text{cl}_Y(\text{int}_Y(E - \{p\})) \cup \{p\} \\ &\subset \text{cl}_Y(\text{int}_Y(E)) \cup \{p\}. \end{aligned}$$

De esta manera, y como $D - \{p\} = Y - X$ es abierto en Y , se tiene

$$\begin{aligned} E \cup D &\subset \text{cl}_Y(\text{int}_Y(E)) \cup \{p\} \cup (D - \{p\}) \\ &= \text{cl}_Y(\text{int}_Y(E) \cup \text{int}_Y(D - \{p\})) \cup \{p\} \\ &\subset \text{cl}_Y(\text{int}_Y(E \cup D)). \end{aligned}$$

Por tanto, $E \cup D = \text{cl}_Y(\text{int}_Y(E \cup D))$ (la otra contención se sigue de la contención $\text{int}_Y(E \cup D) \subset E \cup D$ y del hecho que $E \cup D$ es cerrado en Y). Con esta última igualdad y siguiendo un procedimiento idéntico al de la prueba de la Afirmación 1, se puede mostrar que $\text{Fr}_{C_n(Y)}(C_n(Y, E \cup D))$ es un Z-conjunto en $C_n(Y, E \cup D)$. Esto prueba la Afirmación 3.

Por las Afirmaciones 1, 2 y 3, y el Lema 6.5, concluimos que $C_n(X)$ es homeomorfo a $C_n(Y)$. Por lo tanto, X no tiene hiperespacio único $C_n(X)$. \square

Como un corolario casi inmediato del resultado anterior (y del Teorema 6.3), se tiene el siguiente enunciado.

Corolario 6.7. *Sea X un continuo localmente conexo que no es casi enrejado. Entonces, existe una familia no numerable \mathcal{Y} de continuos localmente conexos y no homeomorfos por pares, tal que se satisfacen las siguientes condiciones:*

- (a) *Para cada $Y \in \mathcal{Y}$, se tiene que X no es homeomorfo a Y .*
- (b) *Para cada $n \in \mathbb{N}$ y cada $Y \in \mathcal{Y}$, se cumple que $C_n(X)$ es homeomorfo a $C_n(Y)$.*

Demostración. Por el Teorema 6.3, existe una familia no numerable \mathcal{D} de dendritas que satisface las siguientes condiciones:

- (i) Para cada $D \in \mathcal{D}$, se cumple que D no contiene arcos libres.
- (ii) El continuo localmente conexo $X \cup_p D$ no es homeomorfo a X .
- (iii) Si B y D son elementos distintos de \mathcal{D} , entonces $X \cup_p B$ y $X \cup_p D$ no son homeomorfos.

Siguiendo la demostración del Teorema 6.6, se puede mostrar que para cualquier $D \in \mathcal{D}$ los hiperespacios $C_n(X)$ y $C_n(X \cup_p D)$ son homeomorfos. De esta forma, si hacemos $\mathcal{Y} = \{X \cup_p D : D \in \mathcal{D}\}$, entonces \mathcal{Y} es una familia que cumple (a) y (b). \square

7 Continuos casi enrejados sin hiperespacio único

Habiendo probado que todo continuo localmente conexo que no es casi enrejado no tiene n -ésimo hiperespacio único, es natural preguntarse si todo continuo localmente conexo que es casi enrejado tiene n -ésimo hiperespacio único. El siguiente resultado responde a este cuestionamiento negativamente. Además de dicho enunciado, este apartado presenta algunos corolarios que pueden ser útiles en la obtención de continuos sin hiperespacio único. Todos estos resultados se exponen originalmente en [8].

En la siguiente demostración se usa de nuevo la notación N_d^X , la cual se detalla en los comentarios previos al Teorema 6.6.

Teorema 7.1. *Sean X un continuo localmente conexo y $n \in \mathbb{N}$. Supongamos que existen un subconjunto cerrado R de $\mathcal{P}(X)$ y subconjuntos abiertos, no vacíos y ajenos por pares U_1, \dots, U_{n+1} de X , tales que se satisfacen las siguientes condiciones:*

$$(a) \quad X - R = \bigcup_{i=1}^{n+1} U_i \text{ y,}$$

$$(b) \quad \text{para cada } i \in \{1, \dots, n+1\}, \text{ se cumple que } R \subset \text{cl}_X(U_i).$$

Entonces, X no tiene hiperespacio único $C_m(X)$, para cada $m \leq n$.

Demostración. Sea $m \leq n$. Fijemos $p \in R$. Por el Teorema 6.3, existe un continuo localmente conexo D que no contiene arcos libres, tal que el continuo localmente conexo $Y = X \cup_p D$ no es homeomorfo a X . Demostraremos que $C_m(X)$ y $C_m(Y)$ son homeomorfos. Para tal efecto, por el Teorema 4.8, podemos suponer que la métrica de Y , denotada por d , es convexa. Esto implica, por el Lema 4.9, que las métricas inducidas sobre X y D por d son convexas. Vamos a probar las siguientes tres afirmaciones.

Afirmación 1. $C_m(Y, R \cup D)$ y $C_m(X, R)$ son cubos de Hilbert.

Obsérvese en primer lugar que $R \cup D \subset \mathcal{P}(X) \cup \mathcal{P}(D) \subset \mathcal{P}(Y)$. Como $R \cup D$ es compacto, se tiene que $R \cup D$ es un subconjunto cerrado de $\mathcal{P}(Y)$. Además, como por hipótesis R es un subconjunto cerrado de $\mathcal{P}(X)$, aplicando el Teorema 5.6 obtenemos que $C_m(Y, R \cup D)$ y $C_m(X, R)$ son cubos de Hilbert.

Afirmación 2. $\text{Fr}_{C_m(Y)}(C_m(Y, R \cup D))$ es un Z-conjunto en $C_m(Y, R \cup D)$.

Sea $\varepsilon > 0$. Por la Observación 4.3, $C_m(Y, R \cup D)$ es un hiperespacio de crecimiento. Luego, aplicando el Lema 4.17, se tiene que la función $\Phi_\varepsilon|_{C_m(Y, R \cup D)} : C_m(Y, R \cup D) \rightarrow C_m(Y, R \cup D)$ es continua y que

$$(H_X)_\infty(\text{Id}_{C_m(Y, R \cup D)}, \Phi_\varepsilon|_{C_m(Y, R \cup D)}) \leq \varepsilon.$$

Tomemos $A \in C_m(Y, R \cup D)$.

Caso 1. $A \cap R \neq \emptyset$.

Por hipótesis, para cada $i \in \{1, \dots, n+1\}$ se cumple que $R \subset \text{cl}_X(U_i)$, por lo cual $N_d^X(A, \varepsilon) \cap \text{cl}_X(U_i) \supset A \cap \text{cl}_X(U_i) \supset A \cap R \neq \emptyset$. Así, $\Phi_\varepsilon(A) \cap U_i \supset N_d^X(A, \varepsilon) \cap U_i \neq \emptyset$. Consideremos una sucesión $\{A_k\}_{k \in \mathbb{N}}$ en $C_m(Y)$, tal que $\lim A_k = \Phi_\varepsilon(A)$. Luego, existe $M \in \mathbb{N}$, tal que, para cada $i \in \{1, \dots, n+1\}$ y $j \geq M$, se cumple $U_i \cap A_j \neq \emptyset$. Tomemos $j \geq M$. Supongamos que $A_j \cap (R \cup D) = \emptyset$. Entonces, $A_j \subset Y - (R \cup D) = (Y - R) \cap (Y - D) = (Y - R) \cap (X - \{p\}) = X - R = \bigcup_{i=1}^{n+1} U_i$. De esta manera, y como los U_i son abiertos, ajenos por pares e intersectan a A_j , se cumple que A_j tiene a lo menos $n+1$ componentes. Esto contradice el hecho que $A_j \in C_m(Y)$. Por tanto, $A_j \cap (R \cup D) \neq \emptyset$; es decir, $A_j \in C_m(Y, R \cup D)$. Esto prueba que no existe sucesión alguna en $C_m(Y) - C_m(Y, R \cup D)$ que converja a $\Phi_\varepsilon(A)$. Así, $\Phi_\varepsilon(A) \notin \text{cl}_{C_m(Y)}(C_m(Y) - C_m(Y, R \cup D)) \supset \text{Fr}_{C_m(Y)}(C_m(Y, R \cup D))$.

Caso 2. $A \cap R = \emptyset$.

En este caso $\Phi_\varepsilon(A) \cap (D - \{p\}) \supset A \cap (D - \{p\}) = (A \cap R) \cup (A \cap (D - \{p\})) = A \cap (R \cup D) \neq \emptyset$. Como $D - \{p\} = Y - X$ es un abierto en Y contenido en $R \cup D$, tenemos que $\Phi_\varepsilon(A) \in \text{int}_{C_m(Y)}(C_m(Y, R \cup D))$.

Por los casos 1 y 2, tenemos que $\Phi_\varepsilon|_{C_m(Y, R \cup D)} : C_m(Y, R \cup D) \rightarrow C_m(Y, R \cup D) - \text{Fr}_{C_m(Y)}(C_m(Y, R \cup D))$. Esto prueba la Afirmación 2.

Afirmación 3. $\text{Fr}_{C_m(X)}(C_m(X, R))$ es un Z-conjunto en $C_m(X, R)$.

Sea $\varepsilon > 0$. Por la Observación 4.3, $C_m(Y, R \cup D)$ es un hiperespacio de crecimiento. Luego, aplicando el Lema 4.17, podemos restringir

$$\Phi_\varepsilon|_{C_m(X, R)} : C_m(X, R) \rightarrow C_m(X, R).$$

Obsérvese, además, que $C_m(X, R) \subset C_m(Y, R \cup D)$. Así, aplicando el Lema 6.4,

$$\begin{aligned} \Phi_\varepsilon(C_m(X, R)) &\subset C_m(Y, R \cup D) - \text{Fr}_{C_m(Y)}(C_m(Y, R \cup D)) \\ &= C_m(Y, R \cup D) - \text{Fr}_{C_m(X)}(C_m(X, R)). \end{aligned}$$

Por tanto,

$$\Phi_\varepsilon|_{C_m(X,R)} : C_m(X,R) \rightarrow C_m(X,R) - \text{Fr}_{C_m(X)}(C_m(X,R)).$$

Como $\Phi_\varepsilon|_{C_m(X,R)}$ es continua y

$$(H_X)_\infty(\text{Id}_{C_m(X,R)}, \Phi_\varepsilon|_{C_m(X,R)}) \leq \varepsilon,$$

se obtiene la Afirmación 3.

Por las Afirmaciones 1, 2 y 3 y el Lema 6.5, concluimos que h es un homeomorfismo. \square

Corolario 7.2. *Sea X un continuo localmente conexo, tal que $X - \mathcal{P}(X)$ es disconexo. Entonces, X no tiene hiperespacio único $C(X)$.*

Demostración. Sean U y V dos subconjuntos abiertos, ajenos y no vacíos de $X - \mathcal{P}(X)$, tales que

$$X - \mathcal{P}(X) = U \cup V.$$

Como $\mathcal{P}(X)$ es cerrado en X , los conjuntos U y V son abiertos en X ; es decir, los conjuntos $X - U$ y $X - V$ son cerrados en X . Así, y como $U \subset X - V$ y $V \subset X - U$, se satisface que $\text{cl}_X(U) \subset X - V$ y $\text{cl}_X(V) \subset X - U$ o, equivalentemente, $V \subset X - \text{cl}_X(U)$ y $U \subset X - \text{cl}_X(V)$.

Sea $R = \text{cl}_X(U) \cap \text{cl}_X(V)$. Obsérvese que $R \subset (X - U) \cap (X - V) = X - (U \cup V) = \mathcal{P}(X)$. Como X es casi enrejado, se tiene que $\text{int}_X(\mathcal{P}(X)) = \emptyset$ y, por consiguiente, $\text{cl}_X(U) \cup \text{cl}_X(V) = \text{cl}_X(U \cup V) = \text{cl}_X(X - \mathcal{P}(X)) = X - \text{int}_X(\mathcal{P}(X)) = X$. Como X es conexo, se sigue de esto último que $R \neq \emptyset$.

Sean $W = X - \text{cl}_X(U)$ y $Z = X - \text{cl}_X(V)$. Luego, $W \cup Z = (X - \text{cl}_X(U)) \cup (X - \text{cl}_X(V)) = X - R$. Asimismo, como $V \subset W$ y $U \subset Z$, se cumple que $R \subset \text{cl}_X(W) \cap \text{cl}_X(Z)$. De esta forma, los conjuntos R , W y Z satisfacen las condiciones del Teorema 7.1, para $n = 1$. Por tanto, X no tiene hiperespacio único $C(X)$. \square

Para la demostración del siguiente Corolario, haremos uso de la siguiente caracterización de los puntos en los que los n -ésimos hiperespacios de un continuo localmente conexo presentan dimensión finita (para una demostración detallada vea el Teorema 8.2 del Capítulo 6 de [10]).

Teorema 7.3 ([8], Teorema 4). *Sean X un continuo localmente conexo, $n \in \mathbb{N}$ y $F \in C_n(X)$. Entonces, son equivalentes las siguientes condiciones:*

(a) $\dim_F(C_n(X))$ es finita.

(b) Existe una gráfica finita D contenida en X , tal que $F \subset \text{int}_X(D)$.

(c) $F \cap \mathcal{P}(X) = \emptyset$.

Corolario 7.4. Sean X una dendrita que no es un árbol y

$$k = \sup\{\text{ord}(p, X) : p \in \mathcal{P}(X)\}.$$

Obsérvese que $k \in \mathbb{N} \cup \{\omega\}$. Entonces, para cada $m < k$ se cumple que X no tiene hiperespacio único $C_m(X)$.

Demostración. Obsérvese que, por el Teorema 6.6, podemos suponer que X es casi enrejado. Por otro lado, obsérvese que, como X no contiene curvas cerradas simples y no es un árbol, X no es una gráfica finita. Así, por el Teorema 7.3 (haciendo $F = X$), se cumple que $\mathcal{P}(X) \neq \emptyset$. En particular, k está bien definido.

Sea $m < k$. Luego, existe un punto $q \in \mathcal{P}(X)$, tal que $m < \text{ord}(q, X)$. Aplicando [14, Teorema 10.13], se tiene que $c(q, X)$ es infinito o $c(q, X)$ es finito y $c(q, X) = \text{ord}(q, X)$. En cualquiera de los dos casos, podemos tomar U_1, \dots, U_m componentes distintas de $X - \{q\}$ y U_{m+1} la unión de las componentes restantes, con $U_{m+1} \neq \emptyset$. Luego

$$X - \{q\} = \bigcup_{i=1}^{m+1} U_i.$$

Como X es localmente conexo y $X - \{p\}$ es un subconjunto abierto de X , por [4, Teorema 2.E.2], se sigue que los conjuntos U_1, \dots, U_{m+1} son abiertos en X .

Fijemos $j \in \{1, \dots, m+1\}$. Obsérvese que $\text{cl}_X(U_j) \subset X - \bigcup_{i \neq j} U_i = U_j \cup \{p\}$ (pues $U_j \subset X - \bigcup_{i \neq j} U_i$, y este último es cerrado en X). Si $p \notin \text{cl}_X(U_j)$, entonces $\text{cl}_X(U_j) = U_j$ y, por ende, U_j sería un subconjunto propio, no vacío, abierto y cerrado de X . Como esto contradice la conexidad de X , se debe tener que $p \in \text{cl}_X(U_j)$. De esta manera, los conjuntos $\{p\}$ y U_1, \dots, U_{m+1} satisfacen las condiciones del Teorema 7.1 para $n = m$. Podemos concluir entonces que X no tiene hiperespacio único $C_m(X)$. \square

Un último corolario al Teorema 7.1 establece un resultado similar al Corolario 6.7, pero reemplazando la condición de no ser casi enrejado con las condiciones del Teorema 7.1.

Corolario 7.5. Sean X un continuo localmente conexo y $n \in \mathbb{N}$, tales que satisfacen las condiciones del Teorema 7.1. Entonces, existe una familia no numerable \mathcal{Y} de continuos localmente conexos y no homeomorfos por pares, tal que las siguientes condiciones se satisfacen:

(a) para cada $Y \in \mathcal{Y}$, se tiene que X no es homeomorfo a Y , y

(b) para cada $Y \in \mathcal{Y}$ y $m \leq n$, se cumple que $C_m(X)$ es homeomorfo a $C_m(Y)$.

Demostración. Por el Teorema 6.3, existe una familia no numerable \mathcal{D} de dendritas que satisface las siguientes condiciones.

- (i) Para cada $D \in \mathcal{D}$, se cumple que D no contiene arcos libres,
- (ii) El continuo localmente conexo $X \cup_p D$ no es homeomorfo a X y
- (iii) Si B y D son elementos distintos de \mathcal{D} , entonces $X \cup_p B$ y $X \cup_p D$ no son homeomorfos.

Dados cualesquiera $D \in \mathcal{D}$ y $m \leq n$, se sigue de la demostración del Teorema 7.1, que los hiperespacios $C_m(X)$ y $C_m(X \cup_p D)$ son homeomorfos. Haciendo $\mathcal{Y} = \{X \cup_p D : D \in \mathcal{D}\}$, tenemos que \mathcal{Y} es una familia no numerable de continuos localmente conexos que satisface (a) y (b). \square

Bibliografía

- [1] G. Acosta and D. Herrera-Carrasco. Dendrites without unique hyperspace. *Houston J. Math.*, 35(2):451–467, 2009.
- [2] R. D. Anderson. On topological infinite deficiency. *Michigan Math. J.*, 14:365–383, 1967.
- [3] J. J. Charatonik and A. Illanes. Local connectednes in hyperspaces. *Rocky Mountain J. Math*, 36(3):811–856, 2006.
- [4] C. O. Christenson and L. Voxman. *Aspects of Topology*. BCS, Associates, 1998.
- [5] D. W. Curtis and R. M. Schori. 2^X and $C(X)$ are homeomorphic to the hilbert cube. *Bull. Amer. Math. Soc.*, 80:927–931, 1974.
- [6] D. W. Curtis and R. M. Schori. Hyperspaces of peano continua are hilbert cubes. *Fund. Math.*, 101(1):19–38, 1978.
- [7] R. Engelking. *General topology*. Sigma series in pure mathematics. Heldermann Verlag, 1989.
- [8] R. Hernández-Gutiérrez, A. Illanes, and V. Martínez de-la Vega. Uniqueness of hyperspaces of peano continua. *Rocky Mountain J. Math.*, 43(5):1583–1624, 2013.
- [9] A. Illanes and S. B. Nadler Jr. *Hyperspaces: Fundamentals and Recent Advances*, volume 126 of *Monographs and Textbooks in Pure and Applied Math.* Marcel Dekker, Inc., New York, Basel, 1999.

- [10] J. G. Ahuatzí-Reyes, D. Herrera-Carrasco, F. Macías-Romero. Dimensión finita en el n -ésimo hiperespacio de continuos localmente conexos (Capítulo 6), *Matemáticas y sus Aplicaciones 5*, 137–164. Textos Científicos, Dirección de Fomento Editorial de la BUAP, 2015.
- [11] S. Macías. *Topics on Continua*. Chapman & Hall/CRC Pure and Applied Mathematics. CRC Press, 2005.
- [12] E. E. Moise. Grille decomposition and convexification theorems for compact metric locally connected continua. *Bull. Amer. Math. Soc.*, 55:1111–1121, 1949.
- [13] Jr. S. B. Nadler. *Hyperspaces of Sets: A Text with Research Questions*, volume 49 of *Monographs and Textbooks in Pure and Applied Mathematics*. New York-Basel: Marcel Dekker Inc., 1978.
- [14] Jr. S. B. Nadler. *Continuum Theory: An Introduction*, volume 158 of *Monographs and Textbooks in Pure and Applied Mathematics*. New York: Marcel Dekker Inc., 1992.

Facultad de Ciencias Físico Matemáticas, BUAP
Avenida San Claudio y 18 Sur, Colonia San Manuel,
Puebla, Pue. C.P. 72570
215571041@alumnos.fcfm.buap.mx
dherrera@fcfm.buap.mx
fmacias@fcfm.buap.mx

Capítulo 10

Acerca de continuos con n -ésimo producto simétrico único

Mauricio Esteban Chacón Tirado, Luis Alberto Guerrero Méndez,
David Herrera Carrasco, Fernando Macías Romero
FCFM, BUAP

Resumen

Sea X un espacio métrico no vacío, compacto y conexo. Dado $n \in \mathbb{N}$, consideremos el n -ésimo producto simétrico de X , con la métrica de Hausdorff, denotado por $F_n(X)$. Un continuo X tiene **hiperespacio único** $F_n(X)$ si la implicación siguiente es verdadera: si Y es un continuo tal que $F_n(X)$ es homeomorfo a $F_n(Y)$, entonces X es homeomorfo a Y . En este capítulo se hace una exposición de la unicidad del n -ésimo producto simétrico para distintas clases de continuos, haciendo énfasis en la unicidad de dicho hiperespacio para los elementos de la clase de los **continuos enrejados** que es una clase para la cual se probó, recientemente, que sus elementos tienen n -ésimo producto simétrico único.

1 Introducción

La teoría de unicidad de hiperespacios es actualmente un área de intensa investigación matemática y, por ende, cuenta con una extensa lista de publicaciones (véase, por ejemplo, [1, 2, 4, 7, 10, 11, 12, 13]). En este capítulo analizamos la unicidad del n -ésimo producto simétrico. A la fecha se ha probado que para todo $n \in \mathbb{N}$, los continuos enrejados tienen n -ésimo producto simétrico único. Este hecho se dió en dos pasos: (1) Primero en 2013, los autores del artículo **Rigidez de productos simétricos**, Rodrigo Hernández Gutiérrez y Verónica Martínez de la Vega, probaron que los continuos que ellos definen como alambrados tienen n -ésimo producto simétrico cuando $n \in \mathbb{N} - \{2, 3\}$ (véase [11]). Como la clase de los continuos enrejados está contenida (de hecho propiamente) en la clase de los continuos alambrados, tenemos que los continuos enrejados tienen n -ésimo producto simétrico único cuando $n \in \mathbb{N} - \{2, 3\}$. (2) En el artículo **Los continuos enrejados tienen segundo y tercer producto simétrico único**, los tres últimos autores de este capítulo, conjuntamente con María de Jesús López Toriz, probaron que los continuos enrejados

tienen n -ésimo producto simétrico único cuando $n \in \{2, 3\}$ (véase [7]).

El caso $n = 1$ (como vemos en la sección 2) siempre se cumple porque para todo continuo X se tiene que $F_1(X)$ es homeomorfo a X .

Cabe mencionar que la clase de los continuos alambrados contiene distintas clases de continuos para las cuales se desconoce si sus elementos tienen n -ésimo producto simétrico único cuando $n = 2$ o $n = 3$.

La temática de este capítulo está involucrada en el problema general siguiente.

Problema 1.1. ¿Qué condiciones necesita tener un continuo para que se garantice la unicidad del hiperespacio $F_n(X)$?

En la segunda sección se exponen los resultados que al paso del tiempo se han ido probando sobre la unicidad del n -ésimo producto simétrico para distintas clases de continuos. En la tercera sección se presentan algunos resultados que necesitamos para hacer algunas de las demostraciones de la sección 4. De hecho en la cuarta sección se exponen resultados muy recientes, entre estos se enuncia el que afirma que los continuos enrejados tienen segundo y tercer producto simétrico único, concluyendo así como corolario que para todo número natural n , los continuos enrejados tienen n -ésimo producto simétrico único.

2 Un poco de historia

En esta sección presentamos lo más básico que se necesita para leer este capítulo y hacemos un viaje rápido por la historia de la unicidad del n -ésimo producto simétrico para distintas clases de continuos.

Para empezar aclaramos que un **continuo** es un espacio métrico no vacío, compacto y conexo.

Cualquier conjunto conexo que sea acotado y cerrado en el plano Euclidiano \mathbb{R}^2 es un ejemplo de un continuo. De hecho siguiendo la misma idea se pueden encontrar muchos ejemplos en el espacio \mathbb{R}^n con la topología usual. En la referencia [14], cuyo autor es Alejandro Illanes, se pueden ver muchos ejemplos de continuos. En trabajos nuestros como [6, Capítulo 10], [8, Capítulo 27] y [9, Capítulo 19] también se pueden ver ejemplos de continuos, incluso en [14] y el capítulo 27 de [8] se exhibe el clásico continuo universal llamado Cubo de Hilbert y se puede ver ahí una prueba de que efectivamente todo continuo se puede encajar en el Cubo de Hilbert.

Dados un continuo X y $n \in \mathbb{N}$, consideremos las siguientes familias de subconjuntos de X

$$2^X = \{A \subset X : A \text{ es no vacío y cerrado en } X\},$$

$$C_n(X) = \{A \in 2^X : A \text{ tiene a lo más } n \text{ componentes}\},$$

$$F_n(X) = \{A \in 2^X : A \text{ tiene a lo más } n \text{ puntos}\},$$

$$C(X) = C_1(X);$$

todos estas clases son metrizadas con la métrica de Hausdorff (véase [17, Definición 2.1]). De hecho, a $F_n(X)$ y a $C_n(X)$, se les conoce como el **n -ésimo producto simétrico** de X y el **n -ésimo hiperespacio** de X , respectivamente.

Definición 2.1. Sean X un continuo, $n \in \mathbb{N}$ y $\mathcal{H}(X) \in \{2^X, C_n(X), F_n(X)\}$. Se dice que X tiene **hiperespacio único** $\mathcal{H}(X)$, si la implicación siguiente es verdadera: si Y es un continuo tal que $\mathcal{H}(X)$ es homeomorfo a $\mathcal{H}(Y)$, entonces X es homeomorfo a Y .

Dado cualquier continuo X , siempre se cumple que X tiene primer producto simétrico único porque la función $h: X \rightarrow F_1(X)$, definida para cada $x \in X$ por $h(x) = \{x\}$, es continua y suprayectiva (véase [14, Lema 2.3]), además es evidente que es inyectiva y cerrada, es decir, h es un homeomorfismo. Así, cuando se quiere verificar la unicidad del n -ésimo producto simétrico para un continuo X se revisa para $n \geq 2$.

Para un continuo X , $x \in X$ y β un número cardinal, decimos que x **tiene orden en X menor o igual que β** , denotado por $ord(x, X) \leq \beta$, si x tiene una base de vecindades \mathfrak{B} en X tal que para cada $U \in \mathfrak{B}$ se cumple que la cardinalidad de la frontera de U en X es menor o igual que β . Decimos que x **tiene orden en X igual a β** , denotado por $ord(x, X) = \beta$, si $ord(x, X) \leq \beta$ y para cualquier número cardinal $\eta < \beta$ se cumple que $ord(x, X) \not\leq \eta$. Sean $E(X) = \{x \in X : ord(x, X) = 1\}$, $O(X) = \{x \in X : ord(x, X) = 2\}$ y $R(X) = \{x \in X : ord(x, X) \geq 3\}$. Los elementos de los conjuntos $E(X)$, $O(X)$ y $R(X)$ son llamados **puntos extremos**, **puntos ordinarios** y **puntos de ramificación**, respectivamente. Dado un subconjunto A de X , el **interior** de A en X es denotado por $int_X(A)$. El símbolo $cl_X(A)$ denota la **cerradura** de A en X .

Dados U_1, \dots, U_m subconjuntos de X y $n \in \mathbb{N}$, sea

$$\langle U_1, \dots, U_m \rangle_n = \{A \in F_n(X) : A \subset U_1 \cup \dots \cup U_m \text{ y } A \cap U_i \neq \emptyset, \\ \text{para cada } i \in \{1, \dots, m\}\}.$$

La familia de todos los conjuntos de la forma $\langle U_1, \dots, U_m \rangle_n$, donde $m \in \mathbb{N}$ y cada U_i es conjunto abierto en X , es una base para la topología de $F_n(X)$, (véase [17, Teorema 1.2]).

Antes de enunciar las definiciones de arco libre, arco libre maximal y ciclo, recordemos que un **arco** es un espacio topológico homeomorfo al intervalo $[0, 1]$ con la topología usual de \mathbb{R} . Si $h: [0, 1] \rightarrow J$ es un homeomorfismo y $h(0) = p$ y

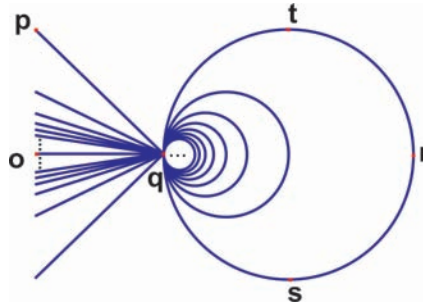


Figura 1

$h(1) = q$, entonces se dice que p y q son los puntos extremos de J . También se dice que J es un arco que va de p a q . Por una **curva cerrada simple** entendemos que es cualquier espacio topológico homeomorfo a la circunferencia unitaria S^1 con la topología usual de \mathbb{R}^2 .

Dado un continuo X , un **ciclo** en X es una curva cerrada simple S en X tal que S tiene a lo más un punto de ramificación de X . Un **arco libre** en X es un arco α con puntos extremos x y y tal que $\alpha - \{x, y\}$ es un conjunto abierto en X . Un **arco libre maximal** en X es un arco libre en X que es maximal con respecto a la inclusión.

Sea

$$\mathcal{A}_S(X) = \{J \subset X : J \text{ es un arco libre maximal o } J \text{ es un ciclo en } X\}.$$

Si $x, y \in X$, decimos que x es **adyacente** a y en X o que x y y son **adyacentes** en X si existe $J \in \mathcal{A}_S(X)$ tal que no es un ciclo y x y y son los extremos de J .

Ejemplo 2.2. Para todo $n \in \mathbb{N}$, sea A_n el segmento de recta que va de $(0, 0)$ a $(-\frac{1}{2}, \frac{1}{2n})$, B_n el segmento de recta que va de $(0, 0)$ a $(-\frac{1}{2}, -\frac{1}{2n})$. Asimismo, sean $C_n = \{(x, y) \in \mathbb{R}^2 : (x - \frac{1}{2n})^2 + y^2 = (\frac{1}{2n})^2\}$ y $B = \{(x, 0) \in \mathbb{R}^2 : -\frac{1}{2} \leq x \leq 0\}$. Consideremos

$$X = \left(\bigcup_{n \in \mathbb{N}} A_n \right) \cup \left(\bigcup_{n \in \mathbb{N}} B_n \right) \cup \left(\bigcup_{n \in \mathbb{N}} C_n \right) \cup B.$$

El espacio topológico X considerado con la topología usual de \mathbb{R}^2 es un continuo (véase la Figura 1). Para $p = (-\frac{1}{2}, \frac{1}{2})$ tenemos que $ord(p, X) = 1$. Sean $o = (-\frac{1}{2}, 0)$ y $q = (0, 0)$. Para todo x que pertenece al arco que va de o a q se cumple que $ord(x, X) = \aleph_0$. Para $r = (1, 0)$ tenemos que $ord(r, X) = 2$. Así, tenemos que p es un punto extremo de X , todos los puntos del arco que va de o a q son puntos de ramificación de X y r es un punto ordinario de X . Sean $s = (\frac{1}{2}, -\frac{1}{2})$ y $t = (\frac{1}{2}, \frac{1}{2})$.



Figura 2: Ejemplo de un 10-odo simple y de un 17-odo simple.

En X tenemos que el arco que va de s a t y pasa por q no es un arco libre en X . El arco que va de s a t y pasa por r sí es un arco libre en X . El arco que va de p a q es un arco libre maximal en X . Para cada $n \in \mathbb{N}$, tenemos que C_n es un ciclo en X .

La historia sobre la unicidad del n -ésimo producto simétrico comienza en el año 2006 cuando Enrique Castañeda y Alejandro Illanes demostraron que las gráficas finitas tienen n -ésimo producto simétrico único. Este resultado salió publicado en el artículo llamado precisamente **Las gráficas finitas tienen n -ésimo producto simétrico único** (véase [4]). Enunciamos a continuación la definición de gráfica finita, vemos algunos ejemplos y enunciamos el teorema de Castañeda e Illanes donde prueban lo dicho.

Definición 2.3. Una **gráfica finita** es un continuo que se puede escribir como unión finita de arcos tales que cualesquiera dos de ellos, o son ajenos o se intersectan solamente en uno o en ambos puntos extremos.

Ejemplo 2.4. Un **n -odo simple** X es un continuo que es unión de n arcos J_1, J_2, \dots, J_n tal que existe $p \in X$ con la propiedad de que $J_i \cap J_j = \{p\}$, cuando $i \neq j$, y p es un extremo de cada uno de los arcos J_i . Los n -odos simples son ejemplos de gráficas finitas. En la Figura 2 vemos gráficamente un 10-odo simple y un 17-odo simple.

En 2006 Enrique Castañeda y Alejandro Illanes probaron lo siguiente.

Teorema 2.5. [4, Corolario 5.9] Sean X una gráfica finita, Y un continuo y $n \in \mathbb{N}$. Si $F_n(X)$ es homeomorfo a $F_n(Y)$, entonces X es homeomorfo a Y .

A continuación definimos las dendritas, vemos algunos ejemplos y lo que se ha hecho respecto a la unicidad del n -ésimo producto simétrico para una clase especial de dendritas.

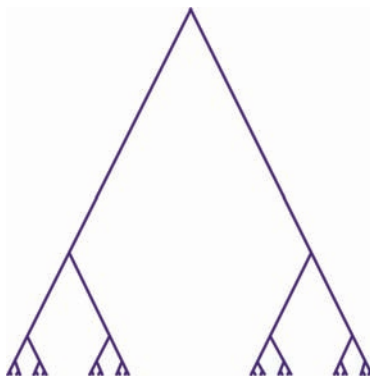


Figura 3: Dendrita de Gehman.

Definición 2.6. Una **dendrita** es un continuo localmente conexo sin curvas cerradas simples.

Ejemplo 2.7. Un ejemplo clásico e importante es la **dendrita de Gehman**. Esta dendrita se construye empezando en la parte superior, de donde salen dos segmentos de recta, en cuyos extremos inferiores se colocan otros dos segmentos de recta más pequeños. Repitiendo este proceso una infinidad de veces y finalmente se considera la cerradura de la unión de todos los segmentos de recta construidos. De hecho el conjunto de puntos que se añaden al tomar la cerradura de la unión de los segmentos de recta construidos es el conjunto de Cantor (véase la Figura 3).

Sean \mathfrak{G} la clase de las gráficas finitas y \mathfrak{D} la clase de las dendritas tales que su conjunto de puntos extremos es cerrado. Se cumple que $\mathfrak{G} \not\subset \mathfrak{D}$ y $\mathfrak{D} \not\subset \mathfrak{G}$, como se puede ver en los ejemplos siguientes.

Ejemplo 2.8. Una curva cerrada simple pertenece a la clase \mathfrak{G} y no pertenece a la clase \mathfrak{D} porque no es una dendrita.

Ejemplo 2.9. Consideremos la dendrita de Gehman (Ejemplo 2.7). La dendrita de Gehman pertenece a la clase \mathfrak{D} y no pertenece a la clase \mathfrak{G} porque no es unión finita de arcos.

Conjuntando el resultado de los autores Gerardo Acosta, Rodrigo Hernández Gutiérrez y Verónica Martínez de la Vega publicado en el artículo titulado **Dendritas y productos simétricos** (véase [1, Teorema 5.2]) y el resultado de David Herrera Carrasco, María de Jesús López Toriz y Fernando Macías Romero que apareció en el artículo llamado **Dendritas con producto simétrico único** (véase [12, Teorema 3.5]), se obtuvo en el 2009 lo siguiente para la clase \mathfrak{D} .

Teorema 2.10. [12, Teorema 3.7] Si $X \in \mathfrak{D}$ y $n \in \mathbb{N}$, entonces X tiene hiperespacio único $F_n(X)$.

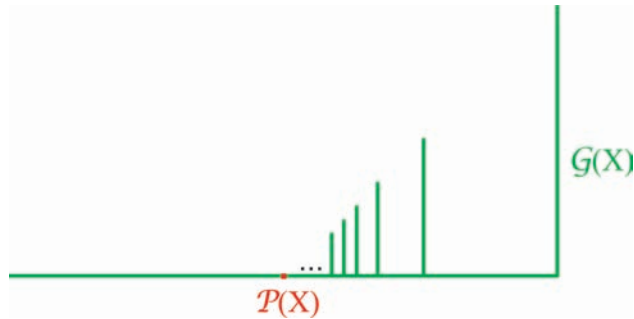


Figura 4: Ejemplo de continuo casi enrejado que no es enrejado.

Varios años más tarde Alejandro Illanes, Rodrigo Hernández Gutiérrez, Verónica Martínez de la Vega definen nuevas clases de continuos en su artículo titulado **Unicidad de hiperespacios para continuos de Peano** (véase [10]); la clase de los continuos casi enrejados y la de los continuos enrejados.

Veamos como va la definición de continuo casi enrejado y de continuo enrejado. Para esto, dado un continuo X , sean

$$\mathcal{G}(X) = \{x \in X : x \text{ tiene una vecindad } G \text{ en } X \text{ tal que } G \text{ es una gráfica finita}\}$$

y

$$\mathcal{P}(X) = X - \mathcal{G}(X).$$

Definición 2.11. Un continuo X es **casi enrejado** si $\mathcal{G}(X)$ es denso en X . Un continuo casi enrejado X es **enrejado** si X tiene una base de vecindades \mathfrak{B} tal que para cada $U \in \mathfrak{B}$ se cumple que $U - \mathcal{P}(X)$ es conexo.

Ejemplo 2.12. Sea $X = ([-1, 1] \times \{0\}) \cup \left(\bigcup_{n \in \mathbb{N}} \left(\left\{\frac{1}{n}\right\} \times \left[0, \frac{1}{n}\right]\right)\right)$. El conjunto X considerado con la topología usual de \mathbb{R}^2 es un continuo. En este continuo X tenemos que $\mathcal{P}(X) = \{(0, 0)\}$. El continuo X (véase la Figura 4) es casi enrejado pero no es enrejado porque para cualquier abierto U en X con $(0, 0)$ en U se cumple que $U - \mathcal{P}(X)$ no es conexo.

Ejemplo 2.13. El continuo de la Figura 5 es definido de la manera siguiente. Para cada $n \in \mathbb{N}$, sea $A_n = \{(x, 2^{-n+1}) \in \mathbb{R}^2 : 0 \leq x \leq 1\}$ y $A_0 = \{(x, 0) \in \mathbb{R}^2 : 0 \leq x \leq 1\}$. Ahora, para cada $n \in \mathbb{N} \cup \{0\}$ y cada entero m tal que $0 \leq m \leq 2^{n+1}$, sea $B_{n,m} = \{(m2^{-n-1}, y) \in \mathbb{R}^2 : 0 \leq y \leq 2^{-n}\}$. Finalmente, consideremos

$$X = \left(\bigcup_{n=0}^{\infty} A_n\right) \cup \left(\bigcup_{n=0}^{\infty} \left(\bigcup_{m=0}^{2^{n+1}} B_{n,m}\right)\right).$$

El continuo X es un continuo enrejado. En este espacio tenemos que $\mathcal{P}(X) = A_0$ y $\mathcal{G}(X) = X - A_0$.

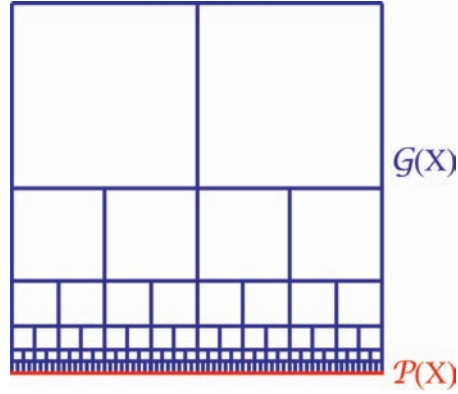


Figura 5: Ejemplo de continuo enrejado.

Ahora, antes de definir algunos conjuntos de puntos del n -ésimo producto simétrico de un continuo recordemos la siguiente definición.

Definición 2.14. Dado $n \in \mathbb{N}$, una n -celda es un espacio topológico homeomorfo a $\prod_{i=1}^n [0, 1]$ con la topología producto.

Dado un continuo X y $n \in \mathbb{N}$, sean

$$P_n(X) = \{A \in F_n(X) : A \cap \mathcal{P}(X) \neq \emptyset\},$$

$$R_n(X) = \{A \in F_n(X) : A \cap R(X) \neq \emptyset\},$$

$$\Lambda_n(X) = F_n(X) - (R_n(X) \cup P_n(X)),$$

$$\mathcal{E}_n(X) = \{A \in F_n(X) : A \text{ tiene una vecindad en } F_n(X) \text{ que es una } n\text{-celda}\}.$$

Sean \mathcal{AM} la clase de los continuos casi enrejados, \mathcal{M} la clase de los continuos enrejados y \mathcal{LC} la clase de los continuos localmente conexos.

En el citado artículo **Unicidad de hiperespacios para continuos de Peano** (véase [10]), Alejandro Illanes, Rodrigo Hernández Gutiérrez y Verónica Martínez de la Vega prueban que la clase de los continuos enrejados contiene a las clases de las gráficas finitas y de las dendritas de la clase \mathfrak{D} [10, Teorema 6]. De hecho dichas contenciones son propias, para ver esto basta considerar el continuo del Ejemplo 2.13 el cual es un continuo enrejado que ni es gráfica finita ni es dendrita, mucho menos pertenece a la clase \mathfrak{D} . También en este mismo artículo sus autores prueban que los continuos enrejados son localmente conexos [10, Lema 2]. Así, la clase de los continuos enrejados es una subclase de los continuos casi enrejados localmente conexos, es decir, $\mathcal{M} \subset \mathcal{AM} \cap \mathcal{LC}$. De hecho la contención es propia como lo muestra el ejemplo 2.15.

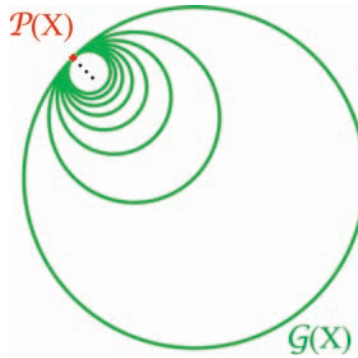


Figura 6: Arete Hawaiano Armónico.

Ejemplo 2.15. Para todo $n \in \mathbb{N}$, sea $C_n = \left\{ (x, y) \in \mathbb{R}^2 : \left(x - \frac{1}{2n}\right)^2 + y^2 = \left(\frac{1}{2n}\right)^2 \right\}$. Consideremos $X = \bigcup_{n \in \mathbb{N}} C_n$. El espacio topológico X considerado con la topología usual de \mathbb{R}^2 es un continuo conocido como el Arete Hawaiano Armónico (véase la Figura 6). Tenemos que X es un continuo casi enrejado localmente conexo que no es enrejado, pues si consideramos cualquier abierto U que contenga al único punto de $\mathcal{P}(X)$ tenemos que $U - \mathcal{P}(X)$ deja de ser un conjunto conexo, y así no se podrá encontrar la base de vecindades que pide la definición de conjunto enrejado.

En el año 2012, David Herrera Carrasco, Fernando Macías Romero y Francisco Vázquez Juárez probaron lo siguiente en el artículo titulado **Continuos de Peano con producto simétrico único** (véase [13]).

Teorema 2.16. [13, Corolario 4.4] Si $X \in \mathcal{AM} \cap \mathcal{LC}$ y $n \in \mathbb{N} - \{2, 3\}$, entonces X tiene hiperespacio único $F_n(X)$.

Así, el Teorema 2.16 es una generalización para $n \geq 4$ del Teorema 2.5 y del Teorema 2.10.

En el año 2013, en el artículo **Rigidez de productos simétricos** (véase [11]), los autores definen una clase de continuos todavía más amplia, la clase de los continuos alambrados.

Definición 2.17. Un **alambre** en un continuo X es un subconjunto α de X tal que α es homeomorfo a uno de los siguientes espacios $(0, 1)$, $[0, 1)$, $[0, 1]$ o la circunferencia unitaria S^1 en el plano Euclidiano y α es una componente de un abierto en X .

Dado un continuo X , sea $W(X) = \bigcup \{ \alpha \subset X : \alpha \text{ es un alambre en } X \}$.

Definición 2.18. Un continuo X es **alambrado** si el conjunto $W(X)$ es denso en X .

Sea \mathcal{W} la clase de los continuos alambrados. En el ya mencionado artículo **Rigidez de productos simétricos**, Rodrigo Hernández Gutiérrez y Verónica Martínez de la Vega, autores de dicho artículo probaron para la clase \mathcal{W} el resultado siguiente.

Teorema 2.19. *[11, Corolario 6] Si $X \in \mathcal{W}$ y $n \in \mathbb{N} - \{2, 3\}$, entonces X tiene hiperespacio único $F_n(X)$.*

También en el artículo **Rigidez de productos simétricos** los autores ponen en evidencia que la clase \mathcal{W} contiene a las clases siguientes de continuos (véase [11, pág. 1578]). La razón es que si α es un arco libre en un continuo X y p y q son los puntos extremos de α , entonces $\alpha - \{p, q\}$ es un alambre. Luego, un continuo X en el que se cumpla que la unión de sus arcos libres es un conjunto denso en X es continuo alambrado.

- (a) La clase de las compactaciones del rayo $[0, \infty)$.
- (b) La clase de las compactaciones de la recta real.
- (c) La clase de los continuos indescomponibles tales que sus subcontinuos propios no degenerados son arcos.
- (d) La clase de los abanicos suaves.

Por lo tanto, los continuos que pertenecen a las clases de los incisos (a)-(d) tienen n -ésimo producto simétrico único cuando $n \geq 4$.

Cabe mencionar que ya desde el año 2009 Alejandro Illanes y Jorge Marcos Martínez Montejano habían probado resultados sobre la unicidad del n -ésimo producto simétrico para compactaciones del rayo en el artículo con título **Compactaciones de $[0, \infty)$ con hiperespacio único $F_n(X)$** (véase [16]). Enunciamos estos resultados a continuación.

Teorema 2.20. *[16, Teoremas 3 y 4] Si X es una compactación del rayo $[0, \infty)$ y $n \in \mathbb{N} - \{3\}$, entonces X tiene hiperespacio único $F_n(X)$.*

Teorema 2.21. *[16, Teorema 10] Si X es una compactación del rayo $[0, \infty)$ tal que su residuo es ANR (retracto absoluto de vecindad), entonces X tiene hiperespacio único $F_3(X)$.*

También cabe mencionar que en el año 2013 José Guadalupe Anaya Ortega, Enrique Catañeda Alvarado y Alejandro Illanes, autores del artículo llamado **Continuos con producto simétrico único** (véase [2]) probaron que los continuos indescomponibles tales que sus subcontinuos propios no degenerados son arcos tienen segundo producto simétrico único.

Teorema 2.22. [2, Teorema 9] *Si X es un continuo indescomponible tal que sus subcontinuos propios no degenerados son arcos, entonces X tiene hiperespacio único $F_2(X)$.*

Dado que los continuos indescomponibles tales que sus subcontinuos propios no degenerados son arcos son continuos alambrados, tenemos como consecuencia del Teorema 2.19 y del Teorema 2.22 lo siguiente.

Corolario 2.23. *Si X es un continuo indescomponible tal que sus subcontinuos propios no degenerados son arcos y $n \in \mathbb{N} - \{3\}$, entonces X tiene hiperespacio único $F_n(X)$.*

Ejemplo 2.24. Un **solenoides** (véase [18, 2.8]) es un ejemplo de un continuo que es alambrado pero que no es casi enrejado. Un solenoide es un continuo alambrado porque es un continuo indescomponible en el que todos sus subcontinuos propios no degenerados son arcos y la clase de los continuos alambrados contiene a la clase de los continuos indescomponibles tales que todos sus subcontinuos propios no degenerados son arcos (véase la página 1578 de [11]). De hecho, un solenoide S no es un continuo casi enrejado porque en cada punto se comporta localmente como el producto cartesiano del conjunto de Cantor con el intervalo $[0,1]$, no pudiendo tener así una vecindad que sea gráfica finita, por lo que el conjunto $\mathcal{G}(S)$ no puede ser denso en S .

Como consecuencia del Corolario 2.23 y de que un solenoide es un continuo indescomponible en el que todos sus subcontinuos propios no degenerados son arcos, tenemos el siguiente.

Corolario 2.25. *Si S es un solenoide y $n \in \mathbb{N} - \{3\}$, entonces S tiene hiperespacio único $F_n(S)$.*

Como ya vimos, por [11, p. 1578], tenemos que $\mathcal{AM} \subset \mathcal{W}$, y por el Ejemplo 2.24, tenemos que $\mathcal{AM} \subsetneq \mathcal{W}$. Así, el Teorema 2.19 generaliza el Teorema 2.16.

En este trabajo exponemos un poco acerca de la prueba del resultado siguiente que ha sido demostrado recientemente por los tres últimos autores de este capítulo, conjuntamente con María de Jesús López Toriz.

Teorema 2.26. [7, Teorema 3.11] *Si $X \in \mathcal{M}$ y $n \in \{2,3\}$, entonces X tiene hiperespacio único $F_n(X)$.*

Como $\mathfrak{G} \subsetneq \mathcal{M}$ y $\mathfrak{D} \subsetneq \mathcal{M}$ tenemos que el Teorema 2.26 es una generalización, para el caso $n = 2$ y $n = 3$ del Teorema 2.5 y del Teorema 2.10.

Por otro lado, el Teorema 2.16 y el Teorema 2.26 son una respuesta parcial positiva al problema siguiente que permanece abierto.

Pregunta 2.27. [1, Pregunta 1.1] Sea X una dendrita y $n \in \mathbb{N} - \{1\}$. ¿Tiene X hiperespacio único $F_n(X)$?

3 Preliminares

En esta sección presentamos principalmente algunos resultados sobre continuos enrejados y casi enrejados, algunos de los cuales son necesarios para ciertas demostraciones de la sección 4. También recordamos lo que significa que una familia de subconjuntos de un espacio topológico X sea localmente finita en X para posteriormente usar este concepto en la sección 4.

En un continuo cualquiera, si un punto p tiene una vecindad que es gráfica finita, entonces no existe una sucesión de puntos de ramificación que tengan vecindad que sea gráfica finita tal que converga a p . Lo enunciamos enseguida de manera formal como un teorema.

Teorema 3.1. [7, Teorema 3.2] Sean X un continuo y $x \in X$. Si existe una sucesión de puntos diferentes de $R(X) \cap \mathcal{G}(X)$ que converge a x , entonces $x \in \mathcal{P}(X)$.

En los continuos enrejados es equivalente que exista la sucesión que menciona el Teorema 3.1 a que el límite de dicha sucesión no tenga una vecindad que sea gráfica finita.

Teorema 3.2. [7, Teorema 3.3] Sea X un continuo enrejado. Entonces $x \in \mathcal{P}(X)$ si y solo si existe una sucesión de puntos diferentes de $R(X) \cap \mathcal{G}(X)$ que converge a x .

En 2009 los autores de [1] probaron un resultado análogo al Teorema 3.3 pero para elementos de la clase \mathfrak{D} . Actualmente los autores de [7] tenemos la siguiente versión generalizada para continuos casi enrejados localmente conexos.

Teorema 3.3. [7, Teorema 3.4] Si X es un continuo casi enrejado localmente conexo y $n \in \{2, 3\}$, entonces $\mathcal{E}_n(X) = \Lambda_n(X)$.

Un arco libre en un continuo localmente conexo siempre forma parte de un arco libre maximal o de un ciclo en dicho continuo.

Lema 3.4. [10, Lema 10] Si X es un continuo localmente conexo y α es un arco libre en X , entonces existe $J \in \mathcal{A}_S(X)$ tal que $\alpha \subset J$.

En los continuos enrejados los puntos de los arcos libres maximales y de los ciclos siempre tienen una vecindad que es gráfica finita.

Teorema 3.5. [7, Teorema 3.1] Si X es un continuo enrejado y $J \in \mathcal{A}_S(X)$, entonces $J \cap \mathcal{P}(X) = \emptyset$.

El lema que sigue nos dice cómo son las componentes del conjunto $\Lambda_n(X)$ para continuos que son casi enrejados localmente conexos diferentes del arco y de la curva cerrada simple.

Lema 3.6. [13, Teorema 3.9] Sean X un continuo casi enrejado localmente conexo que no es un arco ni una curva cerrada simple y $n \in \mathbb{N}$. Entonces las componentes de $\Lambda_n(X)$ son los conjuntos no vacíos de la forma siguiente

$$\langle \text{int}_X(J_1), \dots, \text{int}_X(J_j) \rangle_n,$$

donde $j \leq n$, los conjuntos $\text{int}_X(J_1), \dots, \text{int}_X(J_j)$ son disjuntos dos a dos y para cada $k \in \{1, \dots, j\}$, se cumple que $J_k \in \mathcal{A}_S(X)$.

Sean X un continuo y U y W conjuntos abiertos en X , definimos $c(U, W, X)$ como el número de componentes de $U \cap W$, si la cardinalidad del conjunto de componentes de $U \cap W$ es finita, si no ocurre así definimos $c(U, W, X) = \infty$.

Para cada $p \in \text{cl}_X(W)$, definimos

$$v(p, W, Y) = \min(\{m \in \mathbb{N} : \text{existe una base de vecindades } \mathfrak{B} \text{ de } p \text{ en } Y \\ \text{tal que } c(U, W, Y) = m, \text{ para cada } U \in \mathfrak{B}\} \cup \{\infty\}).$$

Notemos que $\mathcal{E}_n(X)$ es un conjunto abierto en $F_n(X)$. Dado $A \in F_n(X)$, sea

$$v(A) = v(A, \mathcal{E}_n(X), F_n(X)).$$

Teorema 3.7. Sean X y Y continuos casi enrejados localmente conexos y $n \in \{2, 3\}$. Si $h : F_n(X) \rightarrow F_n(Y)$ es un homeomorfismo, entonces

(a) $h(R_n(X) \cup P_n(X)) = R_n(Y) \cup P_n(Y)$.

(b) Si $A \in F_n(X)$, entonces $v(A) = v(h(A))$.

(c) $h(R_n(X)) = R_n(Y)$.

(d) $h(F_1(R(X) \cap \mathcal{G}(X))) = F_1(R(Y) \cap \mathcal{G}(Y))$.

(e) $h(P_n(X)) = P_n(Y)$.

(f) Si X es un continuo enrejado, entonces $h(F_1(\mathcal{P}(X))) \subset F_1(\mathcal{P}(Y))$.

Los continuos que no son localmente conexos no pueden ser enrejados. Es lo que afirma el Teorema 3.8 probado en [10] en el año 2013.

Teorema 3.8. [10, Lema 2] *Si X es un continuo enrejado, entonces X es localmente conexo.*

Sean X y Y continuos enrejados, $n \in \{2, 3\}$ y $h : F_n(X) \rightarrow F_n(Y)$ un homeomorfismo. Por el Teorema 3.8, tenemos que X y Y son localmente conexos. Luego, por (d) del Teorema 3.7, para cada $x \in R(X) \cap \mathcal{G}(X)$, existe $y \in R(Y) \cap \mathcal{G}(Y)$ tal que $h(\{x\}) = \{y\}$. Por (f) del Teorema 3.7, para cada $z \in \mathcal{P}(X)$, existe $w \in \mathcal{P}(Y)$ tal que $h(\{z\}) = \{w\}$. Por el Teorema 3.5 tenemos que $\mathcal{G}(X) = \bigcup\{J : J \in \mathcal{A}_S(X)\}$ y $\mathcal{G}(Y) = \bigcup\{J : J \in \mathcal{A}_S(Y)\}$. Luego, existe una función biyectiva entre los conjuntos $(R(X) \cap \mathcal{G}(X)) \cup \mathcal{P}(X)$ y $(R(Y) \cap \mathcal{G}(Y)) \cup \mathcal{P}(Y)$.

Sea

$$f_1 : (R(X) \cap \mathcal{G}(X)) \cup \mathcal{P}(X) \rightarrow (R(Y) \cap \mathcal{G}(Y)) \cup \mathcal{P}(Y) \quad (*)$$

tal que si $x \in (R(X) \cap \mathcal{G}(X)) \cup \mathcal{P}(X)$, entonces $f_1(x)$ es el único elemento de $(R(Y) \cap \mathcal{G}(Y)) \cup \mathcal{P}(Y)$ tal que $h(\{x\}) = \{f_1(x)\}$. Así, $f_1(R(X) \cap \mathcal{G}(X)) = R(Y) \cap \mathcal{G}(Y)$ y $f_1(\mathcal{P}(X)) = \mathcal{P}(Y)$.

La función f_1 definida en (*) resulta ser una función continua. Dado que la prueba es un tanto engorrosa la omitiremos, dejando enunciado este hecho como un lema con su respectiva referencia.

Lema 3.9. [7, p. 8] *La función f_1 definida en (*) es continua.*

La conexidad local es una propiedad que queda invariante al pasar del espacio al hiperespacio y viceversa.

Teorema 3.10. [3, Teorema 6.3] *Sean X un continuo y $n \in \mathbb{N}$. Entonces X es localmente conexo si y solo si $F_n(X)$ es localmente conexo.*

Ahora, una definición y un teorema sobre espacios topológicos en general que son muy útiles en la demostración del Teorema 4.1.

Definición 3.11. Una familia $\{A_\alpha : \alpha \in \Lambda\}$ de subconjuntos de un espacio topológico X es **localmente finita** en un subconjunto A de X si para cada punto x de A existe una vecindad V de x en X tal que $V \cap A_\alpha \neq \emptyset$ para a lo más un número finito de índices α .

Ejemplo 3.12. Consideremos el intervalo $(0,1)$ y \mathbb{R} con la topología usual. Sea $\mathcal{F} = \{(0, \frac{1}{n}) : n \in \mathbb{N}\}$. Veamos que la familia \mathcal{F} es localmente finita en $(0,1)$. Para para cualquier $r \in (0,1)$ existe $N_r \in \mathbb{N}$ tal que $0 < \frac{1}{N_r} < r$. Sea $V_r = (r - \frac{1}{N_r}, r + \frac{1}{N_r})$. Notemos que V_r es una vecindad de r que solo intersecta a los elementos de la colección \mathcal{F} tales que $n < N_r$. Como $\{(0, \frac{1}{n}) : n < N_r\}$ es una subcolección finita de \mathcal{F} tenemos que \mathcal{F} es efectivamente localmente finita en $(0,1)$. Sin embargo \mathcal{F} no es una familia localmente finita en \mathbb{R} porque para cualquier vecindad V del 0 se tiene que la cardinalidad del conjunto de elementos de \mathcal{F} que intersectan a V no es finita.

Teorema 3.13. [5, Teorema 9.4, p. 83] Sean X y Y espacios topológicos y $\{A_\alpha : \alpha \in \Lambda\}$ una cubierta de X que satisface (a) o (b):

- (a) Todos los conjuntos A_α son abiertos en X .
- (b) Todos los conjuntos A_α son cerrados en X y forman una familia localmente finita en X .

Sea $\{f_\alpha : A_\alpha \rightarrow Y\}$ una familia de funciones continuas tal que para cada $(\alpha, \beta) \in \Lambda \times \Lambda$, se cumple que $f_\alpha|_{A_\alpha \cap A_\beta} = f_\beta|_{A_\alpha \cap A_\beta}$, entonces existe una función continua única $f : X \rightarrow Y$ la cual es una extensión de cada f_α , es decir, para todo $\alpha \in \Lambda$, se cumple que $f|_{A_\alpha} = f_\alpha$.

Para finalizar esta sección enunciamos dos resultados interesantes y necesarios para nuestro propósito sobre continuos localmente conexos.

Teorema 3.14. [13, Teorema 3.1] Si X es un continuo localmente conexo, entonces las proposiciones siguientes son equivalentes.

- (a) El continuo X es casi enrejado.
- (b) Para cada $n \in \mathbb{N}$, el conjunto $\mathcal{E}_n(X)$ es denso en $F_n(X)$.
- (c) Todo conjunto no vacío y abierto en X contiene un arco libre de X .

Lema 3.15. [10, Lema 8] Sean X un continuo localmente conexo X y $\{J_j\}_{j=1}^\infty$ una sucesión de puntos diferentes de $\mathcal{A}_S(X)$. Para cada $j \in \mathbb{N}$, sea $x_j \in J_j$. Si $\lim x_j = x$ para algún $x \in X$, entonces $\lim J_j = \{x\}$ (en $C(X)$).

4 Unicidad del n -ésimo producto simétrico en continuos enrejados

En esta sección presentamos lo correspondiente sobre la unicidad del producto simétrico para continuos enrejados. Comenzamos esta sección con un teorema que enunciamos con su correspondiente referencia y solo damos la idea de la demostración del inciso (e) que es la parte más importante de este resultado.

Teorema 4.1. [7, Teorema 3.9] Sean X y Y continuos enrejados, $n \in \{2, 3\}$ y $h : F_n(X) \rightarrow F_n(Y)$ un homeomorfismo. Si $f_1 : (R(X) \cap \mathcal{G}(X)) \cup \mathcal{P}(X) \rightarrow (R(Y) \cap \mathcal{G}(Y)) \cup \mathcal{P}(Y)$ es la función definida en (*), entonces se cumplen las proposiciones siguientes.

- (a) Si $x, z \in R(X) \cap \mathcal{G}(X)$ son adyacentes en X , entonces $f_1(x)$ y $f_1(z)$ son adyacentes en Y .
- (b) Si $x, z \in R(X) \cap \mathcal{G}(X)$ son adyacentes en X , entonces el número de arcos de $\mathcal{A}_S(X)$ que unen x y z coincide con el número de arcos de $\mathcal{A}_S(Y)$ que unen $f_1(x)$ y $f_1(z)$.
- (c) Si $x \in R(X) \cap \mathcal{G}(X)$, entonces $\text{ord}(x, X) = \text{ord}(f_1(x), Y)$.
- (d) Sea $x \in R(X) \cap \mathcal{G}(X)$. Supongamos que el número de ciclos de X (respectivamente de Y) que contienen a x (respectivamente a $f_1(x)$) es k_1 (respectivamente k'_1), el número de puntos extremos de X (respectivamente de Y) adyacentes a x (respectivamente a $f_1(x)$) es k_2 (respectivamente k'_2) y el número de arcos de X (respectivamente de Y) que unen x (respectivamente $f_1(x)$) con otro punto de ramificación de X (respectivamente de Y) es k_3 (respectivamente k'_3). Entonces $k_1 = k'_1$, $k_2 = k'_2$ y $k_3 = k'_3$.
- (e) X es homeomorfo a Y .

Bosquejo de la demostración del inciso (e) del Teorema 4.1. Para probar que X y Y son homeomorfos lo que se hace es extender la función f_1 a un homeomorfismo entre X y Y .

Para cualesquiera $x, z \in R(X) \cap \mathcal{G}(X)$ tales x y z son adyacentes en X , definimos los conjuntos

$$\mathcal{J}_{x,z} = \{J: J \text{ es un arco en } X \text{ y } J \text{ va de } x \text{ a } z\} \text{ y}$$

$$\mathcal{J}_{f_1(x),f_1(z)} = \{K: K \text{ es un arco en } Y \text{ y } K \text{ va de } f_1(x) \text{ a } f_1(z)\}.$$

Por el inciso (b), existe una función biyectiva $\varphi_{x,z}: \mathcal{J}_{x,z} \rightarrow \mathcal{J}_{f_1(x),f_1(z)}$. Para cada $J \in \mathcal{J}_{x,z}$, se define un homeomorfismo $f_{x,z}: J \rightarrow \varphi_{x,z}(J)$ que cumpla que $f_{x,z}(x) = f_1(x)$ y $f_{x,z}(z) = f_1(z)$.

Ahora, para cada $x \in R(X) \cap \mathcal{G}(X)$, sean

$$\mathcal{E}_x = \{J: J \text{ es un arco en } X \text{ y } J \text{ va de } x \text{ a algún punto extremo de } X\} \text{ y}$$

$$\mathcal{E}_{f_1(x)} = \{K: K \text{ es un arco en } Y \text{ y } K \text{ va de } f_1(x) \text{ a algún punto extremo de } Y\}.$$

Por el inciso (d), existe una función biyectiva $\varphi_x: \mathcal{E}_x \rightarrow \mathcal{E}_{f_1(x)}$. Sea $J \in \mathcal{E}_x$ y supongamos que $\{e_x\} = E(J) - \{x\}$ y $\{e_{f_1(x)}\} = E(\varphi_x(J)) - \{f_1(x)\}$. Se define un homeomorfismo $f_{x,e_x}: J \rightarrow \varphi_x(J)$ tal que $f_{x,e_x}(x) = f_1(x)$ y $f_{x,e_x}(e_x) = e_{f_1(x)}$.

También, para cada $x \in R(X) \cap \mathcal{G}(X)$, sean

$$\mathcal{C}_x = \{J: J \text{ es un ciclo en } X \text{ y } x \in J\} \text{ y}$$

$$\mathcal{C}_{f_1(x)} = \{K: K \text{ es un ciclo en } Y \text{ y } f_1(x) \in K\}.$$

Otra vez, por el inciso (d), existe una función biyectiva $g_x: \mathcal{C}_x \rightarrow \mathcal{C}_{f_1(x)}$. Para cada $J \in \mathcal{C}_x$ se define un homeomorfismo $f_x: J \rightarrow g_x(J)$ tal que $f_x(x) = f_1(x)$.

Tenemos que $\mathcal{A}_S(X)$ es una familia localmente finita en $\mathcal{G}(X)$ (véase la Definición 3.11) porque para todo $z \in \mathcal{G}(X)$ existe una gráfica finita G_z contenida en X tal que $z \in \text{int}_X(G_z)$ y para a lo más una cantidad finita de elementos J de $\mathcal{A}_S(X)$ se cumple que $\text{int}_X(G_z) \cap J \neq \emptyset$.

Por el Teorema 3.13, existe una función continua única

$$f_2: \mathcal{G}(X) \rightarrow \mathcal{G}(Y),$$

la cual es una extensión de cada una de las funciones $f_{x,z}$, f_{x,e_x} y f_x , respectivamente. Es decir, para cada $J \in \mathcal{J}_{x,z}$, se cumple que $f_2|_J = f_{x,z}$. Para cada $J' \in \mathcal{E}_x$, se cumple que $f_2|_{J'} = f_{x,e_x}$. También, para cada $J'' \in \mathcal{C}_x$, se cumple que $f_2|_{J''} = f_x$.

Notemos que para cada $x \in R(X) \cap \mathcal{G}(X)$ se cumple $f_2(x) = f_1(x)$ y para cada $f_2(J) \in \mathcal{A}_S(Y)$ se cumple que $J \in \mathcal{A}_S(X)$. Así, f_2 es biyectiva.

Ahora, sea

$$f_3: X \rightarrow Y$$

definida para cada $x \in X$ por

$$f_3(x) = \begin{cases} f_1(x) & \text{si } x \in \mathcal{P}(X) \\ f_2(x) & \text{si } x \in \mathcal{G}(X). \end{cases}$$

Para ver que f_3 es una función continua sean $x \in X$ y $\{x_m\}_{m=1}^{\infty}$ una sucesión en X que converge a x .

Caso 1. Si $x \in \mathcal{G}(X)$, entonces existe una gráfica finita G contenida en X tal que $x \in \text{int}_X(G)$. Existe $N \in \mathbb{N}$ tal que si $m \geq N$, entonces $x_m \in \text{int}_X(G)$. Podemos suponer que para cada $m \in \mathbb{N}$ se tiene que $x_m \in \mathcal{G}(X)$. Luego, $f_3(x_m) = f_2(x_m)$ y $f_2(x) = f_3(x)$. Así, la sucesión $\{f_3(x_m)\}_{m=1}^{\infty}$ converge a $f_3(x)$.

Caso 2. Si $x \in \mathcal{P}(X)$, entonces analizamos tres casos.

- (i) Para cada $m \in \mathbb{N}$, tenemos que $x_m \in \mathcal{P}(X)$. Por la continuidad de f_1 , tenemos que f_3 es continua en x .
- (ii) Para cada $m \in \mathbb{N}$, tenemos que $x_m \in X - \mathcal{P}(X)$. Existe una sucesión $\{J_m\}_{m=1}^{\infty}$ de elementos diferentes de $\mathcal{A}_S(X)$ tal que $x_m \in J_m$. Luego, por el Lema 3.15, tenemos que $\{J_m\}_{m=1}^{\infty}$ converge a $\{x\}$. Para cada $m \in \mathbb{N}$, sea $r_m \in J_m \cap R(X)$. Como la sucesión $\{r_m\}_{m=1}^{\infty}$ converge a x , tenemos que $f_3(r_m) = f_2(r_m) = f_1(r_m)$ y $f_1(x) = f_3(x)$, tenemos que $\{f_3(r_m)\}_{m=1}^{\infty}$ converge a $f_3(x)$. Como $f_3(r_m) \in f_2(J_m)$ y $\{f_2(J_m)\}_{m=1}^{\infty}$ es una sucesión de elementos diferentes de

$\mathcal{A}_S(Y)$, otra vez, por el Lema 3.15, tenemos que $\{f_2(J_m)\}_{m=1}^\infty$ converge a $\{f_3(x)\}$. Como $f_3(x_m) = f_2(x_m) \in f_2(J_m)$, tenemos que $\{f_3(x_m)\}_{m=1}^\infty$ converge a $f_3(x)$.

- (iii) Existe una cantidad infinita de elementos de $\{x_m\}_{m=1}^\infty$ en $\mathcal{G}(X)$ y existe una cantidad infinita de elementos de $\{x_m\}_{m=1}^\infty$ en $\mathcal{P}(X)$. En este caso, procedemos como en los casos (i) y (ii) al mismo tiempo.

Así, f_3 es una función continua. Además, f_3 es biyectiva. Por lo tanto, f_3 es un homeomorfismo. \square

Pregunta 4.2. Dados $n \in \mathbb{N}$, una clase de continuos \mathfrak{C} , un continuo $X \in \mathfrak{C}$ y Y un continuo cualquiera. Si los hiperespacios $F_n(X)$ y $F_n(Y)$ son homeomorfos, ¿será que Y también pertenece a la clase \mathfrak{C} ?

La respuesta es que sí puede suceder, depende de la clase \mathfrak{C} de la que estemos hablando y del hiperespacio en cuestión. Las clases que cumplen con esta propiedad reciben un nombre especial.

Definición 4.3. Sean \mathfrak{C} una clase de continuos, $n \in \mathbb{N}$ y $\mathcal{H}(X) \in \{2^X, F_n(X), C_n(X)\}$. Decimos que la clase \mathfrak{C} es \mathcal{H} -**cerrada** si la implicación siguiente es verdadera: si $X \in \mathfrak{C}$ y Y es un continuo tal que $\mathcal{H}(X)$ es homeomorfo a $\mathcal{H}(Y)$, entonces $Y \in \mathfrak{C}$.

En seguida dos ejemplos de clases de continuos que son F_n -cerradas.

Teorema 4.4. [7, Teorema 3.10] Para todo $n \in \mathbb{N}$, la clase de los continuos casi enrejados localmente conexos es F_n -cerrada.

Demostración. Sean X un continuo casi enrejado localmente conexo, $n \in \mathbb{N}$ y $h : F_n(X) \rightarrow F_n(Y)$ un homeomorfismo. Por el Teorema 3.8, tenemos que X es localmente conexo. Luego, por el Teorema 3.10, tenemos que Y es localmente conexo. Por otro lado, por el Teorema 3.14, tenemos que $cl_{F_n(X)}(\mathcal{E}_n(X)) = F_n(X)$. En [1, Teorema 4.1], se ve que $h(\mathcal{E}_n(X)) = \mathcal{E}_n(Y)$. Luego,

$$F_n(Y) = h(F_n(X)) = h(cl_{F_n(X)}(\mathcal{E}_n(X))) = cl_{F_n(Y)}(h(\mathcal{E}_n(X))) = cl_{F_n(Y)}(\mathcal{E}_n(Y)).$$

Así, por el Teorema 3.14, tenemos que Y es casi enrejado. Por lo tanto, para todo $n \in \mathbb{N}$, la clase de los continuos casi enrejados localmente conexos es F_n -cerrada. \square

Veamos un resultado parecido para los continuos enrejados.

Teorema 4.5. [7, Teorema 3.10] Para $n \in \{2, 3\}$, la clase de los continuos enrejados es F_n -cerrada.

Así, estamos listos para probar que los continuos enrejados tienen segundo y tercer producto simétrico único.

Teorema 4.6. *Si X es un continuo enrejado y $n \in \{2, 3\}$, entonces X tiene hiperespacio único $F_n(X)$.*

Demostración. Sea Y un continuo tal que $F_n(X)$ es homeomorfo a $F_n(Y)$. Por el Teorema 4.5, tenemos que Y es un continuo enrejado. Luego, por el Teorema 4.1, tenemos que X es homeomorfo a Y . \square

En resumen tenemos el resultado siguiente que es un buen avance en las investigaciones sobre la unicidad del n -ésimo producto simétrico en las distintas clases de continuos conocidas.

Corolario 4.7. *Si X es un continuo enrejado y $n \in \mathbb{N}$, entonces X tiene hiperespacio único $F_n(X)$.*

Demostración. Si $n > 3$, entonces por el Teorema 2.19, tenemos que X tiene hiperespacio único $F_n(X)$. Si $n \in \{2, 3\}$, entonces por el Teorema 4.6, tenemos que X tiene hiperespacio único $F_n(X)$. \square

En resumen, una de las subclases más grandes de \mathcal{W} para la que se ha probado hasta ahora que sus elementos tienen n -ésimo producto simétrico único, cuando n es cualquier número natural, es la clase \mathcal{M} de los continuos enrejados. La otra subclase de \mathcal{W} para la que también se ha probado que sus elementos tienen n -ésimo producto simétrico único, cuando n es cualquier número natural, es la clase de las compactaciones del rayo $[0, \infty)$ tal que su residuo es ANR (retracto absoluto de vecindad).

Hasta el día de hoy queda sin respuesta la pregunta siguiente, la cual nos invita a continuar investigando qué clases de continuos tienen o no producto simétrico único.

Pregunta 4.8. [7, Pregunta 3.13] Si X es un continuo alambrado y $n \in \{2, 3\}$, ¿tiene X hiperespacio único $F_n(X)$?

Respecto a continuos que no son necesariamente alambrados se sabe lo siguiente, que fue probado en 2013 por José Guadalupe Anaya Ortega, Enrique Catañeda Alvarado y Alejandro Illanes, en su artículo **Continuos con producto simétrico único** (véase [2]).

Teorema 4.9. [2, Teorema 12] *Si X es un continuo arco conexo tal que tiene un punto único que es vértice de un triodo simple en X , entonces X tiene hiperespacio único $F_2(X)$.*

Corolario 4.10. [2, Corolario 13] Si X pertenece a alguna de las clases de continuos siguientes, entonces X tiene hiperespacio único $F_2(X)$.

- (a) La clase de los abanicos.
- (b) La clase de los conos sobre espacios métricos compactos que no contienen arcos.
- (c) La clase de los conos sobre continuos hereditariamente indescomponibles.

Cuestiones como la de la Pregunta 4.8 quedan sin resolver para distintas subclases de \mathcal{W} . Enunciamos a continuación algunas preguntas sin importar si las clases de continuos en cuestión son o no subclases de \mathcal{W} , más bien las enunciamos con ánimo de que estas cuestiones nos induzcan a seguir investigando sobre unicidad de hiperespacios.

Pregunta 4.11. [1, Pregunta 1.1] Si X es una dendrita y $n \in \mathbb{N}$, ¿tiene X hiperespacio único $F_n(X)$?

Pregunta 4.12. [2, Pregunta 15] Si X es un continuo arco conexo tal que tiene un punto único que es vértice de un triodo simple en X , ¿tiene X hiperespacio único $F_3(X)$?

Pregunta 4.13. [16, Pregunta 11] Si X es cualquier compactación del rayo $[0, \infty)$, ¿tiene X hiperespacio único $F_3(X)$?

Pregunta 4.14. [15, Pregunta 40] Si X es un continuo encadenable y $n \in \mathbb{N}$, ¿tiene X hiperespacio único $F_n(X)$?

Pregunta 4.15. [15, Pregunta 41] Si X es cualquier abanico y $n \in \mathbb{N}$, ¿tiene X hiperespacio único $F_n(X)$?

Pregunta 4.16. [15, Pregunta 42] Si X es un continuo indescomponible tal que todos sus subcontinuos propios no degenerados son arcos, ¿tiene X hiperespacio único $F_3(X)$?

Pregunta 4.17. [15, Pregunta 43] Sean $n \in \mathbb{N}$. ¿Existirá un continuo de dimensión finita tal que no tenga hiperespacio único $F_n(X)$?

Pregunta 4.18. [15, Pregunta 44] Si X es un continuo hereditariamente indescomponible, ¿tiene X hiperespacio único $F_2(X)$?

Pregunta 4.19. [15, Pregunta 45] ¿El Pseudo Arco tendrá segundo producto simétrico único?

Bibliografía

- [1] Gerardo Acosta, Rodrigo Hernández Gutiérrez, Verónica Martínez de la Vega, *Dendrites and symmetric products*, Glasnik Math. Ser. III **44** (2009) no. 1, 195–210.
- [2] José Guadalupe Anaya Ortega, Enrique Castañeda Alvarado, Alejandro Illanes, *Continua with unique symmetric product*, Comment. Math. Univ. Carolin. **54** (2013) no. 3, 397–406.
- [3] Janusz Jerzy Charatonik, Alejandro Illanes, *Local connectedness in hyperspaces*, Rocky Mountain J. Math. **36** (2006), 811–856.
- [4] Enrique Castañeda, Alejandro Illanes, *Finite graphs have unique symmetric products*, Topology Appl. **153** (2006), 1434–1450.
- [5] James Dugundji, *Topology*. BCS Associates, 2nd ed., Moscow, Idaho, USA, 1978.
- [6] Luis Alberto Guerrero Méndez, David Herrera Carrasco, María de Jesús López Toriz, Fernando Macías Romero, *El n -ésimo hiperespacio suspensión de un continuo* (Capítulo 10), Matemáticas y sus Aplicaciones 3, Textos Científicos, Dirección de Fomento Editorial, BUAP, págs. 175–195, Primera Edición 2013, ISBN: 978 607 487 659 8.
- [7] Luis Alberto Guerrero Méndez, David Herrera Carrasco, María de Jesús López Toriz, Fernando Macías Romero, *Meshed continua have unique second and third symmetric products*, Topology Appl. **191** (2015), 16–27.
- [8] Luis Alberto Guerrero Méndez, David Herrera Carrasco, Fernando Macías Romero, *Dendritas Locales* (Capítulo 27), Matemáticas y sus Aplicaciones I, Textos Científicos, Dirección de Fomento Editorial, BUAP, Primera Edición 2011, págs. 291–300, ISBN: 978–607–487–338–2.
- [9] Luis Alberto Guerrero Méndez, David Herrera Carrasco, Fernando Macías Romero, *Encajes* (Capítulo 19), Matemáticas y sus Aplicaciones 2, Textos Científicos, Dirección de Fomento Editorial, BUAP, Primera Edición 2013, págs. 271–285, ISBN: 978–607–487–600–0.
- [10] Rodrigo Hernández Gutiérrez, Alejandro Illanes, Verónica Martínez de la Vega, *Uniqueness of hyperspaces for Peano continua*, Rocky Mountain J. Math. **43** (5)(2013), 1583–1624.

- [11] Rodrigo Hernández Gutiérrez, Verónica Martínez de la Vega, *Rigidity of symmetric products*, Topology Appl. **160** (2013), 1577–1587.
- [12] David Herrera Carrasco, María de Jesús López Toriz, Fernando Macías Romero *Dendrites with unique symmetric products*, Topology Proc. **34** (2009), 175–190.
- [13] David Herrera Carrasco, Fernando Macías Romero, Francisco Vázquez Juárez, *Peano continua with unique symmetric products*, Journal of Mathematics Research; **4**(4) (2012), 1–9.
- [14] Alejandro Illanes, *Hiperespacios de continuos*, Aportaciones Matemáticas, Serie Textos N. 28, Sociedad Matemática Mexicana, ISBN: 968–36–3594–6, 2004.
- [15] Alejandro Illanes, *Uniqueness of Hyperspaces*, Questions Answers Gen. Topology, **30** (2012), 21–44.
- [16] Alejandro Illanes, Jorge Martínez Montejano, *Compactifications of $[0, \infty)$ with unique hyperspace $F_n(X)$* , Glasnik Math., Serie III **44** (2009), 457–478.
- [17] Alejandro Illanes, Sam Bernard Nadler Jr., *Hyperspaces Fundamentals and Recent Advances*. Monographs and Textbooks in Pure and Applied Math., Vol. 216, Marcel Dekker, Inc., New York, 1999.
- [18] Sam Bernard Nadler Jr., *Continuum Theory: An Introduction*, Monographs and Textbooks in Pure and Applied Math. Vol. 158, Marcel Dekker, New York, Basel, Hong Kong, 1992.

Facultad de Ciencias Físico Matemáticas, BUAP
Avenida San Claudio y 18 Sur, Colonia San Manuel,
Puebla, Pue. C.P. 72570
maeschacon@fcfm.buap.mx
luisgm@alumnos.fcfm.buap.mx
dherrera@fcfm.buap.mx
fmacias@fcfm.buap.mx

Índice de autores

Ahuatzi Reyes, José Gerardo, 215
Angoa Amador, José Juan, 57
Aparicio Hernández, Aarón, 107, 131
Arenas Martínez, Yoanna, 153

Cervantes Gómez, Lucía, 13
Chacón Tirado, Mauricio Esteban, 241
Cocoletzi Adame, Fernando, 57
Cruz Suárez, Hugo Adán, 153, 179

González Pérez, Ana Luisa, 13
Guerrero Méndez, Luis Alberto, 241

Herrera Carrasco, David, 215, 241

Juárez Hernández, Bulmaro, 195

Lübberstedt, Thomas, 195

Macías Romero, Fernando, 215, 241
Muñoz Aguirre, Evodio, 33

Poisot Macías, Julio Erasto, 13

Remedios Santiago, Leonardo, 13
Reyes Cervantes, Hortensia, 153
Reyes Matamoros, Jenaro, 195

Salem Silva, Francisco Sergio, 33

Tajonar Sanabria, Francisco, 153

Vázquez Guevara, Víctor Hugo, 179

Matemáticas y sus aplicaciones 6
su composición, diseño y cuidado, estuvo a cargo
de Fernando Macías Romero
se terminó de imprimir el 7 de diciembre de 2015, en los talleres de
El Errante editor, S. A. de C. V., sito en privada Emiliano Zapata 5947,
Col. San Baltazar Campeche, Puebla, Pue., México.

El tiraje consta de 500 ejemplares.