

Es esta memoria la expresión de un esfuerzo por dejar constancia de la riqueza matemática de la Quinta Gran Semana Nacional de la Matemática. Aquí se recogen los trabajos de acuerdo a las sesiones de la misma, los cuales fueron sometidos a estricto arbitraje.



ISBN: 978-607-487-133-3



9 786074 871333

5<sup>a</sup> Gran Semana Nacional de la Matemática 5GSNM Memorias



5<sup>a</sup>

Gran  
Semana  
Nacional  
de la  
Matemática  
5GSNM

Memorias





5<sup>a</sup> Gran Semana Nacional de la Matemática  
5GSNM, Memorias  
Facultad de Ciencias Físico Matemáticas  
*Benemérita Universidad Autónoma de Puebla*

EDITORES:

Juan Angoa Amador,  
Fernando Macías Romero,  
Jacobo Oliveros Oliveros

BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

Enrique Agüera Ibáñez

*Rector*

José Ramón Eguíbar Cuenca

*Secretario General*

Pedro Hugo Hernández Tejeda

*Vicerrector de Investigación y Estudios de Posgrado*

Lilia Cedillo Ramírez

*Vicerrectora de Extensión y Difusión de la Cultura*

Cuapatitzio Ramírez Romero

*Director de la Facultad de Ciencias Físico Matemáticas*

Carlos Contreras Cruz

*Director de Fomento Editorial*

Facultad de Físico Matemáticas

Ciudad Universitaria

Tel. (222) 229.55.00 Ext. 7565

Primera Edición 2010.

ISBN: 978-607-487-133-3

©Benemérita Universidad Autónoma de Puebla

La QUINTA GRAN SEMANA NACIONAL DE LA MATEMÁTICA (5GSNM) se realizó del 31 agosto al 4 de septiembre de 2009. Esta actividad fue apoyada financieramente por:

- Vicerrectoría de Investigación y Estudios de Posgrado
- Vicerrectoría de Docencia
- Cuerpo Académico de Topología y Sistemas Dinámicos
- Academia de Matemáticas de la Facultad de Ciencias Físico Matemáticas

Editores:

Juan Angoa Amador, Fernando Macías Romero, Jacobo Oliveros Oliveros.



# Contenido

PRESENTACIÓN	1
<b>Análisis Matemático</b>	<b>3</b>
VARIACIONES SOBRE UN MISMO TEMA: LA DERIVADA.	5
<p><i>José Antonio Cariño Ortega</i>  <i>Juan Alberto Escamilla Reyna</i>  <i>Juan Alberto Martínez Cadena.</i></p>	
TEOREMA ESPECTRAL.	11
<p><i>Slaviša Djordjević</i>  <i>Sergio Atayan García Balán.</i></p>	
ENTROPÍA Y SISTEMAS DINÁMICOS.	15
<p><i>Juan Francisco Estrada García.</i></p>	
EL DESCENSO TOPOLÓGICO UNIFORME Y LAS RAÍCES DE OPERADORES B-FREDHOLM.	21
<p><i>Gabriel Kantún Montiel.</i></p>	
FUNCIONES ABSOLUTAMENTE CONTINUAS A PARTIR DEL CONJUNTO DE CANTOR.	27
<p><i>Francisco Javier Mendoza Torres.</i></p>	
CONTINUIDAD DEL ESPECTRO EN DESPLAZAMIENTOS UNILATERALES DE MULTIPLICIDAD FINITA.	33
<p><i>Salvador Sánchez Perales.</i></p>	
<b>Divulgación</b>	<b>39</b>
EL RAZONAMIENTO MATEMÁTICO.	41
<p><i>Eduadro Ariza Velázquez</i>  <i>Pedro García Juárez</i>  <i>Rosa García Tamayo</i>  <i>Carlos Palomino Jiménez.</i></p>	
EL REFRESCO DE COLA Y LA ESTADÍSTICA.	61
<p><i>Hortensia Reyes Cervantes</i>  <i>Sara Rodríguez Rodríguez.</i></p>	

LA DIVULGACIÓN DE LA MATEMÁTICA: UN ENFOQUE PERSONAL.	71
<i>Roberto Torres Hernández.</i>	
<b>Ecuaciones Diferenciales y Modelación Matemática</b>	<b>83</b>
EJEMPLOS NUMÉRICOS DE PROBLEMAS CON INESTABILIDAD NUMÉRICA Y SU REGULARIZACIÓN MEDIANTE EL MÉTODO DE TIJONOV.	85
<i>Félix Aquino Camacho</i> <i>Jacobo Oliveros Oliveros.</i>	
PRUEBAS MAXIMIN PARA LA CALIDAD DE ESTABILIZACIÓN ROBUSTA APLICADAS A UN ROBOT MÓVIL.	101
<i>Honorato Azucena Coyotécatl</i> <i>Javier Cabrera Martínez</i> <i>Eladio Flores Mena</i> <i>José Eligio Moisés Gutiérrez Arias.</i>	
DISEÑO DE UN CONTROL ÓPTIMO CONTINUO PARA LA ESTABILIZACIÓN DE TRAYECTORIAS EN UN ROBOT MÓVIL.	117
<i>Javier Cabrera Martínez</i> <i>José Eligio Moisés Gutiérrez Arias</i> <i>Lucio Hernández Angulo</i> <i>María Monserrat Morín Castillo.</i>	
RESOLUCION AUTOMÁTICA DE ECUACIONES DIFERENCIALES USANDO EL SISTEMA CHEBOP.	131
<i>Mario Alberto Carballo Flores</i> <i>Francisco Sergio Salem Silva.</i>	
IDENTIFICACIÓN DE FUENTES BIOELÉCTRICAS EN EL CEREBRO CONSIDERANDO UNA GEOMETRÍA SIMPLE DE LA CABEZA.	145
<i>Julio Conde Mones</i> <i>Andrés Fragueta Collar</i> <i>María Monserrat Morín Castillo</i> <i>Jacobo Oliveros Oliveros.</i>	
DISEÑO DE UN CONTROL ÓPTIMO DIGITAL PARA UN ROBOT MÓVIL.	155
<i>José Eligio Moisés Gutiérrez Arias</i> <i>María Monserrat Morín Castillo</i> <i>Gelacio Salas Ortega.</i>	

CONTROL ÓPTIMO Y UNA PROPUESTA DE SOLUCIÓN NUMÉRICA PARA UN PVF.	165
<i>Agustín Méndez Andrade</i> <i>Evodio Muñoz Aguirre.</i>	
VALIDACIÓN NUMÉRICA Y EXPERIMENTAL DE UN ALGORITMO DE IDENTIFICACIÓN DE FUENTES BIOELÉCTRICAS TIPO DIPOLO.	177
<i>María Monserrat Morín Castillo</i> <i>Jacobo Oliveros Oliveros</i> <i>Blanca Minerva Rodríguez Ramírez.</i>	
<b>Enseñanza e Historia</b>	187
UNA ESTRATEGIA DE ENSEÑANZA-APRENDIZAJE.	189
<i>Olga Leticia Fuchs Gómez</i> <i>María Guadalupe Raggi Cárdenas.</i>	
TEOREMA DE EQUIPOTENCIA DE BERNSTEIN.	195
<i>Sergio Atayan García Balán</i> <i>Manuel Ibarra Contreras.</i>	
LA SIMETRÍA DINÁMICA EN LA OBRA DE JOSÉ CLEMENTE OROZCO.	197
<i>Emma Garrido Sánchez</i> <i>Iván Ruíz</i> <i>Pedro Tolentino Eslava.</i>	
EL CÁLCULO.	207
<i>Raúl Linares Gracia.</i>	
<b>Lógica Matemática</b>	215
RESEÑA BREVE SOBRE LÓGICAS MULTIVALUADAS.	217
<i>Iván Cortés Cortés</i> <i>José Arrazola Ramírez.</i>	

<b>Probabilidad y Estadística</b>	227
KAPLAN-MEIER Y NELSON-AALEN, DOS MÉTODOS NO PARAMÉTRICOS EMPLEADOS EN SUPERVIVENCIA.	229
<i>Bulmaro Juárez Hernández</i>	
<i>Óscar Palmeros Rojas</i>	
<i>Francisco Solano Tajonar Sanabria.</i>	
VALUACIÓN DE UNA OPCIÓN CALL BARRERA DOBLE TIPO A.	237
<i>Carlos Palomino Jiménez</i>	
<i>Francisco Solano Tajonar Sanabria.</i>	
<b>Topología</b>	247
PROYECCIONES ORBITALES COMO $G$ -FIBRACIONES APROXIMATIVAS.	249
<i>Alexander Bykov.</i>	
COCIENTE DE PRODUCTOS SIMÉTRICOS DE UN CONTINUO.	267
<i>Enrique Castañeda Alvarado</i>	
<i>Javier Sánchez Martínez.</i>	
EL SEGUNDO PRODUCTO SIMÉTRICO DE DENDRITAS.	275
<i>Enrique Castañeda Alvarado</i>	
<i>Lázaro Trejo Allende.</i>	
DENDRITAS CUYO CONJUNTO DE PUNTOS EXTREMOS ES CERRADO.	279
<i>David Herrera Carrasco</i>	
<i>Fernando Macías Romero</i>	
<i>Francisco Vázquez Juárez.</i>	
ALGUNAS PROPIEDADES TOPOLÓGICAS Y LA FUNCIÓN $T$ DE JONES.	291
<i>Sergio Macías.</i>	
TODO ESPACIO MÉTRICO ES PERFECTO.	299
<i>Armando Martínez García.</i>	
EXTENSIONES SIMPLES, CONEXIDAD Y SEPARACIÓN.	303
<i>Marcelino Taxis Taxis.</i>	

## Presentación

La QUINTA GRAN SEMANA NACIONAL DE LA MATEMÁTICA (5GSNM) se realizó del 31 de agosto al 4 de septiembre de 2009.

Las Grandes Semanas Nacionales de la Matemática son un ejercicio anual del colectivo matemático nacional que, organizado bajo la dirección de la Academia de Matemáticas de la Facultad de Ciencias Físico Matemáticas, presenta una gran gama de actividades que forman parte del quehacer matemático.

Es esta memoria la expresión de un esfuerzo por dejar constancia de la riqueza matemática de esta 5GSNM. Toda obra editorial se realiza con la esperanza de tener numerosos lectores, si por lo menos los asistentes se convierten en lectores y estos propagan este volumen, estaremos satisfechos.

En este volumen se recogen las memorias de la 5GSNM, agrupando los trabajos de acuerdo a las sesiones de la misma, los cuales fueron sometidos a estricto arbitraje. Agradecemos sinceramente a todos los árbitros su dedicación y profesionalismo así como a los encargados de las mencionadas sesiones: Juan Angoa Amador, Enrique Arrazola Ramírez, Raúl Escobedo Conde, Manuel Ibarra Contreras, Francisco Javier Mendoza Torres, Jacobo Oliveros Oliveros y Hortensia Reyes Cervantes.

Expresamos nuestro más sincero agradecimiento a todas las personas que hicieron posible la publicación de estas memorias de la 5GSNM, muy particularmente a Miguel Ángel García Ariza por su cuidadosa entrega en el diseño editorial de esta obra y a Félix Aquino Camacho, Julio Conde Mones y Gelacio Salas Ortega por su apoyo.

También agradecemos el apoyo financiero de la Vicerrectoría de Investigación y Estudios de Posgrado de la BUAP para esta publicación.



# **Sesión de Análisis Matemático**

Quinta Gran Semana Nacional de la Matemática



# VARIACIONES SOBRE UN MISMO TEMA: LA DERIVADA

JOSÉ ANTONIO CARIÑO ORTEGA  
JUAN ALBERTO ESCAMILLA REYNA  
JUAN ALBERTO MARTÍNEZ CADENA  
FCFM - BUAP

RESUMEN. Presentaremos otros conceptos de derivada: derivada simétrica y derivada fuerte, y su relación con el concepto clásico de derivada.

## 1. INTRODUCCIÓN

Normalmente en los libros elementales de cálculo a nivel universitario [1], [3] y [4], se motiva el concepto de derivada de una función en un punto, a través de dos problemas:

1) Calcular la pendiente de la recta tangente a la gráfica de una función en un punto  $(x_0, f(x_0))$ .

2) Calcular la velocidad instantánea de un cuerpo puntual que se mueve en línea recta, conociendo la posición del cuerpo en cada instante de tiempo  $f = f(t)$ .

Ambos problemas nos conducen al siguiente problema general:

Dada  $f : [a, b] \rightarrow \mathbb{R}$  una función y  $x_0 \in [a, b]$ , calcular el siguiente límite:

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Si este límite existe se dice que  $f$  tiene derivada en el punto  $x_0$ . Sin embargo consideramos que con estos mismos problemas, se pueden motivar otros conceptos de derivada como el de derivada simétrica y el de derivada fuerte. Desde el punto de vista intuitivo consideramos que estos conceptos son tan legítimos como el clásico. La prevalencia del concepto de derivada clásico se debe a otras consideraciones de origen teórico como práctico.

Estos conceptos no son equivalentes al concepto de derivada clásico. Nuestro objetivo es presentar estos conceptos y ver su relación con el concepto clásico.

## 2. DERIVADA CLÁSICA

2.1. DEFINICIÓN. Una función  $f : [a, b] \rightarrow \mathbb{R}$  es derivable en  $x_0 \in [a, b]$ , si y sólo si,

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} \text{ existe.}$$

En este caso el límite se designa por  $f'(x_0)$  y recibe el nombre de derivada de  $f$  en  $x_0$ . Decimos también que  $f$  es derivable en  $[a, b]$  si es derivable en todo punto de  $[a, b]$ . Ver figura 1. Si  $x_0 = a$  o  $x_0 = b$  el límite anterior se entiende como un límite lateral, por la izquierda en el caso de que  $x_0 = a$  y por la derecha en el caso de que  $x_0 = b$ .

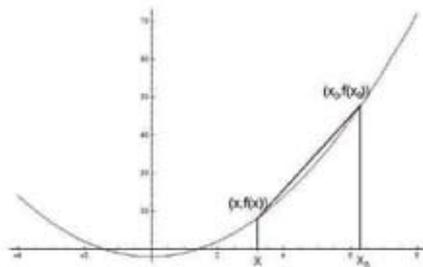


FIGURA 1. Concepto Clásico de Derivada

## 3. DERIVADA SIMÉTRICA

3.1. DEFINICIÓN. Una función  $f : [a, b] \rightarrow \mathbb{R}$  tiene derivada simétrica en  $x_0 \in (a, b)$ , si y sólo si,

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0 - h)}{2h} \text{ existe.}$$

El límite se designa por  $f'_s(x_0)$  y recibe el nombre de derivada simétrica de  $f$  en  $x_0$ . Ver figura 2.

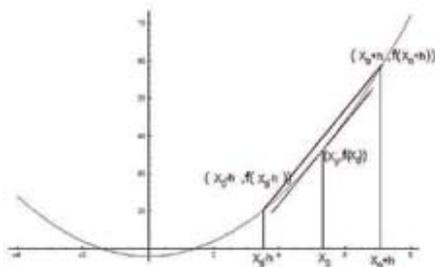


FIGURA 2. Concepto de Derivada Simétrica

3.2. TEOREMA. Sea  $f : [a, b] \rightarrow \mathbb{R}$  una función y sea  $x_0 \in (a, b)$ . Si  $f$  es derivable en  $x_0$ , entonces  $f$  tiene derivada simétrica en  $x_0$  y  $f'(x_0) = f'_s(x_0)$ .

DEMOSTRACIÓN. Sea  $h \neq 0$  tal que  $x_0 + h, x_0 - h \in (a, b)$ , entonces

$$\frac{f(x_0 + h) - f(x_0 - h)}{2h} = \frac{1}{2} \left( \frac{f(x_0 + h) - f(x_0)}{h} + \frac{f(x_0) - f(x_0 - h)}{h} \right). \quad (1)$$

Como  $f$  es diferenciable en  $x_0$

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0) \quad y \quad \lim_{h \rightarrow 0} \frac{f(x_0) - f(x_0 - h)}{h} = f'(x_0), \quad (2)$$

de (1) y (2) se cumple

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0 - h)}{2h} = f'(x_0). \quad \square$$

3.3. OBSERVACIÓN. Si  $f$  tiene derivada simétrica en un punto  $x_0$  no necesariamente tiene derivada clásica en  $x_0$ .

3.4. EJEMPLO. Sea  $f(x) = |x|$ , tomemos  $x_0 = 0$ .

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{f(0 + h) - f(0 - h)}{2h} &= \lim_{h \rightarrow 0} \frac{|0 + h| - |0 - h|}{2h} \\ &= \lim_{h \rightarrow 0} \frac{h - h}{2h} = 0. \end{aligned}$$

Por tanto su derivada simétrica en  $x_0 = 0$  es 0, pero sabemos que  $f$  en  $x_0 = 0$  no tiene derivada clásica.

#### 4. DERIVADA FUERTE

4.1. DEFINICIÓN. Sea  $f : [a, b] \rightarrow \mathbb{R}$  y sea  $x_0 \in (a, b)$ . El número  $L$  es la derivada fuerte de  $f$  en  $x_0$  si para cada  $\epsilon > 0$  existe  $\delta > 0$  tal que

$$\left| \frac{f(y) - f(x)}{y - x} - L \right| < \epsilon,$$

cada vez que  $|x - x_0| < \delta$ ,  $|y - y_0| < \delta$ ,  $x \neq y$  y  $x, y \in [a, b]$ . En el punto  $a$ , la derivada fuerte de  $f$  en  $a$  se define como la derivada clásica de  $f$  en  $a$ . De manera similar se define la derivada fuerte de  $f$  en  $b$ .

4.2. OBSERVACIÓN. Note que no es necesario que  $x_0$  pertenezca al intervalo que une  $x$  con  $y$ . Usaremos  $f'_{st}(x_0)$  para representar la derivada fuerte de la función  $f$  en el punto  $x_0$ .

4.3. TEOREMA. Sea  $f : [a, b] \rightarrow \mathbb{R}$  y sea  $x_0 \in [a, b]$ . Si  $f$  tiene derivada fuerte en  $x_0$ , entonces es derivable en  $x_0$  y  $f'(x_0) = f'_{st}(x_0)$ .

DEMOSTRACIÓN. Si  $x_0 = a$  o  $x_0 = b$  es inmediato. Sea  $x_0 \in (a, b)$ .

Sea  $\epsilon > 0$ , existe  $\delta > 0$  tal que si  $x \neq y$ ,  $|x - x_0| < \delta$  y  $|y - x_0| < \delta$ , entonces

$$\left| \frac{f(x) - f(y)}{x - y} - f'_{st}(x_0) \right| < \epsilon.$$

Sean  $0 < |x - x_0| < \delta$  y  $y = x_0$ . Entonces

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - f'_{st}(x_0) \right| < \epsilon.$$

Por lo tanto  $f$  tiene derivada clásica en  $x_0$  y  $f'(x_0) = f'_{st}(x_0)$ .  $\square$

4.4. OBSERVACIÓN. Su recíproca es falsa, definamos la función  $f : [-1, 1] \rightarrow \mathbb{R}$

$$f(x) = \begin{cases} x^2 \sin\left(\frac{\pi}{x}\right), & \text{si } x \neq 0, \\ 0, & \text{si } x = 0. \end{cases} \quad (3)$$

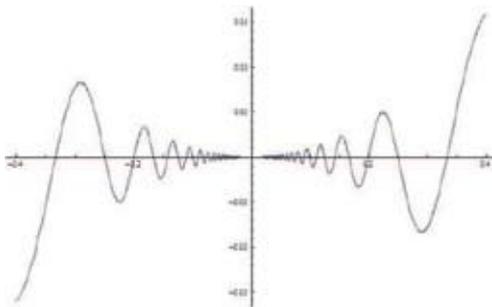


FIGURA 3. Gráfica de la función (3). El eje horizontal es el eje  $X$

La función  $f$  es diferenciable en 0 y  $f'(0) = 0$ . Para cada entero positivo tomemos:

$$x_n = \frac{1}{n + \frac{1}{2}} \quad \text{y} \quad \hat{x}_n = \frac{1}{n}$$

entonces:

$$\begin{aligned} \hat{x}_n - x_n &= \frac{1}{n(2n+1)}, \\ |f(\hat{x}_n) - f(x_n)| &= \frac{1}{(n + \frac{1}{2})^2}. \end{aligned}$$

Por lo tanto

$$\lim_{n \rightarrow \infty} \left| \frac{f(\hat{x}_n) - f(x_n)}{\hat{x}_n - x_n} \right| = \lim_{n \rightarrow \infty} \frac{4n}{2n+1} = 2.$$

Así que  $f$  no tiene derivada fuerte en 0.  $\square$

En el siguiente teorema presentamos las condiciones que garantizan que los conceptos de derivada fuerte y derivada clásica coincidan en el intervalo cerrado  $[a, b]$  [2].

4.5. TEOREMA. Una función  $f : [a, b] \rightarrow \mathbb{R}$  es fuertemente diferenciable en cada punto de  $[a, b]$ , si y solo si,  $f'$  es continua en  $[a, b]$ .

DEMOSTRACIÓN. ( $\Rightarrow$ ) Supongamos que  $f$  es fuertemente diferenciable en cada punto de  $[a, b]$ . Sea  $x_0 \in (a, b)$  y sea  $\varepsilon > 0$ , como  $f$  es fuertemente diferenciable en  $x_0$  existe  $\delta > 0$  tal que:

$$\left| \frac{f(y) - f(x)}{y - x} - f'(x_0) \right| < \varepsilon,$$

siempre que  $|x - x_0| < \delta$ ,  $|y - x_0| < \delta$ ,  $x \neq y$  y  $x, y \in [a, b]$ .

Sea  $x \in (x_0 - \delta, x_0 + \delta) \cap [a, b]$ , como  $f$  es diferenciable en  $x$ , existe  $y \in (x_0 - \delta, x_0 + \delta) \cap [a, b]$  tal que:

$$\left| \frac{f(y) - f(x)}{y - x} - f'(x) \right| < \varepsilon.$$

Entonces

$$\begin{aligned} |f'(x) - f'(x_0)| &= \left| f'(x) - \frac{f(y) - f(x)}{y - x} + \frac{f(y) - f(x)}{y - x} - f'(x_0) \right| \\ &\leq \left| f'(x) - \frac{f(y) - f(x)}{y - x} \right| + \left| \frac{f(y) - f(x)}{y - x} - f'(x_0) \right| < 2\varepsilon \end{aligned}$$

Así que la función  $f'$  es continua en  $x_0$ .

El caso  $x_0 = a$  o  $x_0 = b$  es inmediato.

( $\Leftarrow$ ) Supongamos que  $f'$  es continua en  $[a, b]$ . Sea  $x_0 \in (a, b)$  y sea  $\varepsilon > 0$ , existe  $\delta > 0$  tal que

$$|f'(d) - f'(x_0)| < \varepsilon,$$

para todo  $d \in (x_0 - \delta, x_0 + \delta) \cap [a, b]$ .

Supongamos que  $|x - x_0| < \delta$ ,  $|y - x_0| < \delta$ ,  $x < y$  y  $x, y \in [a, b]$ . Por teorema del valor medio existe  $d \in (x, y)$  tal que

$$f'(d) = \frac{f(y) - f(x)}{y - x}.$$

Entonces

$$\left| \frac{f(y) - f(x)}{y - x} - f'(x_0) \right| = |f'(d) - f'(x_0)| < \varepsilon.$$

Por lo tanto  $f$  tiene derivada fuerte en  $x_0$ .

El caso  $x_0 = a$  o  $x_0 = b$  es inmediato.  $\square$

## 5. CONCLUSIÓN

Consideramos que los conceptos de derivada simétrica y de derivada fuerte deberían, al menos, mencionarse en los textos básicos de Cálculo Diferencial a nivel universitario, como alternativas para resolver el problema de la pendiente de la recta tangente a una gráfica y al problema de la velocidad instantánea.

## REFERENCIAS

- [1] Apostol Tom M., *CALCULUS*, Inc., España, 1984.
- [2] Gordon Russell A., *The Integrals of Lebesgue, Denjoy, Perron, and Henstock*, Graduate Studies in Math., vol. 4, American Math. Soc., Providence, 1994.
- [3] Spivak Michael, *CALCULUS*, Segunda Edición, Barcelona, España, 2001.
- [4] Swokowski Earl W., *Cálculo con geometría analítica*, Wadsworth Internacional Iberoamericana, Inc. 1982.

Facultad de Cs. Físico Matemáticas, BUAP.  
18 Sur y Avenida San Claudio, Col. San Manuel,  
Puebla, Pue. C.P. 72570.  
200609414@fcfm.buap.mx  
jescami@fcfm.buap.mx  
200606306@fcfm.buap.mx

# TEOREMA ESPECTRAL

SLAVIŠA DJORDJEVIĆ  
SERGIO ATAYAN GARCÍA BALÁN  
FCFM - BUAP

RESUMEN. Sean  $V$  un espacio vectorial de dimension finita sobre el campo  $\mathbb{C}$  con producto interior y  $W$  su subespacio. Sea  $T$  un operador lineal y normal sobre  $V$ . Usando que un operador normal es diagonalizable y la unicidad de las proyecciones ortogonales, en este trabajo se va a presentar una representación de un operador normal  $T$  en un espacio  $V$  (dimensión finita), que describe a  $T$  como una combinación lineal de las proyecciones ortogonales sobre los distintos espacios propios de  $T$ . Los escalares de esta combinación lineal son precisamente los valores propios asociados a estos espacios propios. Esta representación se conoce como el Teorema Espectral.

## 1. INTRODUCCIÓN

Sean  $W_1, W_2, \dots, W_k$  subespacios de un espacio vectorial  $V$ . Diremos que  $V$  es la suma directa de los subespacios  $W_1, W_2, \dots, W_k$  y lo escribiremos  $V = W_1 \oplus W_2 \oplus \dots \oplus W_k$ , si

$$V = \sum_{i=1}^k W_i \quad \text{y} \quad W_j \cap \sum_{i \neq j} W_i = \{0\} \quad \forall j \in \{1, 2, \dots, k\}$$

En este caso,  $\dim V = \sum_{i=1}^k \dim(W_i)$  y para cada  $x \in V$  existe unico  $k$ -tuple de vectores  $(x_1, \dots, x_k)$ ,  $x_i \in W_i$ ,  $i = 1, \dots, k$ , tal que  $x = x_1 + \dots + x_k$ . Además, el operador  $P_i : V \rightarrow W_i$  definido con  $P_i(x_1 + \dots + x_k) = x_i$  se llama proyección sobre  $W_i$ .

El operador  $P$  cumple con las siguientes propiedades:

- $P$  es idempotente, es decir,  $P^2 = P$ ,
- $I - P$  y  $P^*$  son idempotentes, donde  $P^*$  es la conjugada transpuesta de  $P$ ,
- $P(I - P) = (I - P)P = O$ ,
- $Px = x$  si y sólo si  $x \in R(P)$ ,
- $N(P) = R(I - P)$ .

Sea  $S$  un subconjunto no vacío de un espacio  $V$  con producto interior. Definimos  $S^\perp$  como el conjunto de todos los vectores en  $V$  que son ortogonales a cada vector en  $S$ , es decir,  $S^\perp = \{x \in V : \langle x, y \rangle = 0 \quad \forall y \in S\}$ . Este conjunto  $S^\perp$  es llamado el complemento ortogonal de  $S$ .

1.1. TEOREMA. Sea  $V$  un espacio de dimensión finita con producto interior, si  $W$  es cualquier subespacio de  $V$ , entonces  $\dim(V) = \dim(W) + \dim(W^\perp)$ .

Si  $S$  es un subconjunto no vacío de un espacio  $V$  con producto interior, y  $M \subset S^\perp$ , entonces  $S \oplus M$  es llamada la suma directa ortogonal de  $S$ ,  $M$  y se denota por  $S \oplus^\perp M$ .

Si  $T : V \longrightarrow V$  es una proyección en un espacio con producto interior  $V$ , diremos que  $T$  es una proyección ortogonal si se cumple que  $R(T)^\perp = N(T)$  y  $N(T)^\perp = R(T)$ .

Sea  $T$  un operador lineal en un espacio vectorial  $V$ , y sea  $\lambda$  un número complejo tal que existe un vector  $x \in V \setminus \{0\}$  con  $Tx = \lambda x$ . Entonces,  $\lambda$  se llama valor propio y  $x$  vector propio de  $T$ . Sea  $E_\lambda(T) = \{x \in V : Tx = \lambda x\}$ . Diremos que  $E_\lambda(T)$  es el espacio propio de  $T$  correspondiente al valor propio  $\lambda$ .

Sea  $V$  un espacio vectorial de dimensión finita con producto interior, y sea  $T$  un operador lineal en  $V$ . Existe entonces un único operador lineal  $T^* : V \rightarrow V$  tal que  $\langle T(x), y \rangle = \langle x, T^*(y) \rangle$  para toda  $x, y \in V$ .  $T^*$  es llamado el adjunto de  $T$ .

## 2. RESULTADOS PRELIMINARES

Un operador lineal  $T$  sobre un espacio de dimensión finita  $V$  es diagonalizable, si existe una base ordenada  $\beta$  de  $V$  tal que la representación matricial de  $T$  con respecto a  $\beta$  es una matriz diagonal.

2.1. TEOREMA. Un operador lineal  $T$  en un espacio vectorial  $V$  de dimensión finita es diagonalizable si y sólo si existe una base de  $V$  cuyos elementos son los vectores propios de  $T$ .

2.2. COROLARIO. Un operador lineal  $T$  en un espacio vectorial  $V$  de dimensión finita es diagonalizable si y sólo si  $V$  es la suma directa de los espacios propios de  $T$ .

Sea  $V$  un espacio vectorial de dimensión finita con producto interior y  $T$  un operador lineal sobre  $V$ . Diremos que  $T$  es normal si  $TT^* = T^*T$ . Las siguientes son propiedades relevantes que cumplen los operadores normales.

- $\|T(x)\| = \|T^*(x)\|$  para todo  $x \in V$ ,
- $T - \lambda$  es normal para todo escalar  $\lambda$ ,
- Si  $x$  es un vector propio de  $T$ , entonces  $x$  es también un vector propio de  $T^*$ . Además, si  $T(x) = \lambda x$ , entonces  $T^*(x) = \bar{\lambda}x$ .

2.3. TEOREMA. Sea  $V$  un espacio con producto interior y  $T$  un operador normal en  $V$ , entonces si  $\lambda_1$  y  $\lambda_2$  son valores propios distintos de  $T$  correspondientes a los vectores propios  $x_1$  y  $x_2$ , entonces  $x_1$  y  $x_2$  son ortogonales.

DEMOSTRACIÓN. Sean  $\lambda_1$  y  $\lambda_2$  distintos valores propios de  $T$  con sus correspondientes vectores propios  $x_1$  y  $x_2$ . Entonces

$$\lambda_1 \langle x_1, x_2 \rangle = \langle \lambda_1 x_1, x_2 \rangle = \langle T(x_1), x_2 \rangle = \langle x_1, T^*(x_2) \rangle = \langle x_1, \overline{\lambda_2} x_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle.$$

Como  $\lambda_1 \neq \lambda_2$ , se concluye que  $\langle x_1, x_2 \rangle = 0$  □

2.4. TEOREMA. Sea  $V$  un espacio con producto interior,  $T$  un operador normal en  $V$  y  $\lambda$  su valor propio. Entonces con respecto a la decomposición  $V = E_\lambda(T) \oplus E_\lambda(T)^\perp$ , el operador  $T$  tiene representación  $T = \lambda I_{E_\lambda(T)} \oplus T_1$ , donde  $\lambda$  no es valor propio de  $T_1$ .

DEMOSTRACIÓN. Sea  $T_1 = T|_{E_\lambda(T)^\perp}$  y suponemos que  $\lambda$  es valor propio de  $T_1$  con vector propio  $v \in E_\lambda(T)^\perp$ . Es fácil ver que el vector  $v$  es vector propio correspondiente al valor propio  $\lambda$  para el operador  $T$ . Entonces,  $v \in E_\lambda(T) \cap E_\lambda(T)^\perp (= \{0\})$ . □

Finalmente, el siguiente resultado nos da una equivalencia importante de la normalidad de un vector.

2.5. TEOREMA. Si  $T$  es un operador lineal en un espacio vectorial  $V$  de dimensión finita con producto interior sobre  $\mathbb{C}$ , entonces  $T$  es normal si y sólo si existe una base ortonormal de  $V$  cuyos elementos son los vectores propios de  $T$ .

### 3. TEOREMA ESPECTRAL

3.1. TEOREMA. Sea  $T$  un operador normal sobre un espacio vectorial complejo  $V$  con dimensión finita y producto interior y sean  $\lambda_1, \lambda_2, \dots, \lambda_k$  sus distintos valores propios. Para cada  $i \in \{1, 2, \dots, k\}$ , sea  $W_i = E_{\lambda_i}(T)$  y sea  $P_i$  la proyección ortogonal de  $V$  en  $W_i$ . Entonces, se cumple lo siguiente:

- (a)  $V = W_1 \oplus W_2 \oplus \dots \oplus W_k$ .
- (b) Si  $W_i'$  denota la suma directa de los subespacios  $W_j$  para  $i \neq j$ , entonces  $W_i^\perp = W_i'$ .
- (c)  $P_i P_j = \delta_{ij} P_i$  para  $i, j \in \{1, 2, \dots, k\}$ .
- (d)  $I = P_1 + P_2 + \dots + P_k$ .
- (e)  $T = \lambda_1 P_1 + \lambda_2 P_2 + \dots + \lambda_k P_k$ .

DEMOSTRACIÓN. (a) Suponiendo que  $T$  es normal y usando Teoremas 2.5 y 2.1 y Corolario 2.2 se tiene que  $V = W_1 \oplus W_2 \oplus \dots \oplus W_k$ .

(b) Sea  $x \in W_i, y \in W_j$  con  $i \neq j$ . Entonces por Teorema 2.3 tenemos  $\langle x, y \rangle = 0$  y de aquí  $W_i' \subseteq W_i^\perp$ . De parte (a) tenemos que

$$\dim(W_i') = \sum_{j \neq i} \dim(W_j) = \dim(V) - \dim(W_i).$$

Además por Teorema 1.1  $\dim(W^\perp) = \dim(V) - \dim(W)$ , de lo anterior,  $W_i' \subseteq W_i^\perp$  y  $\dim(W_i') = \dim(W_i^\perp)$ , por lo tanto  $W_i' = W_i^\perp$ .

(c) Si  $i = j$ , entonces  $P_i P_i = P_i^2 = P_i = \delta_{ij} P_i$ . Sea  $i \neq j$ , veamos que  $P_i P_j = 0$ . En efecto, sea  $x \in V$  entonces  $P_j(x) \in W_j \subseteq W_i' = W_i^\perp$ , por lo tanto  $P_i P_j(x) = 0$ .

(d) Como  $P_i$  es la proyección ortogonal de  $V$  en  $W_i$ , se sigue de (b) que  $N(P_i) = W_i' = W_i^\perp$ . De aquí, para  $x \in V$  tal que  $x = x_1 + x_2 + \cdots + x_k$ ,  $x_i \in W_i$ , se tiene  $P_i(x) = x_i$  y de esta manera  $x = P_1(x) + P_2(x) + \cdots + P_k(x)$ . Por lo tanto  $I = P_1 + P_2 + \cdots + P_k$ .

(e) Para  $x \in V$ , tal que  $x = x_1 + x_2 + \cdots + x_k$ ,  $x_i \in W_i$ , se tiene

$$\begin{aligned} T(x) &= T(x_1) + T(x_2) + \cdots + T(x_k) \\ &= \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k \quad (\text{Teorema 2.4}) \\ &= \lambda_1 P_1(x) + \lambda_2 P_2(x) + \cdots + \lambda_k P_k(x) \\ &= (\lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_k P_k)(x) \end{aligned}$$

por lo tanto  $T = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_k P_k$ . □

El conjunto  $\{\lambda_1, \lambda_2, \dots, \lambda_k\}$  de valores propios de  $T$  es llamado el espectro de  $T$ . A la suma  $T = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_k P_k$  en (e) se le llama la descomposición espectral de  $T$ . Esta descomposición es única salvo por el orden los valores propios. Además,  $T^* = \overline{\lambda_1} P_1 + \overline{\lambda_2} P_2 + \cdots + \overline{\lambda_k} P_k$ .

#### REFERENCIAS

- [1] Friedberg, Insel, Spence; *Linear Algebra*, Fourth Edition, Prentice Hall, 2003.

Facultad de Ciencias Físico Matemáticas, BUAP.  
 Apartado Postal 1152 Puebla, Pue. C.P. 72000, México.  
 slavdj@fcfm.buap.mx  
 darkbizhop@hotmail.com

# ENTROPÍA Y SISTEMAS DINÁMICOS

JUAN FRANCISCO ESTRADA GARCÍA  
FCFM - BUAP

RESUMEN. En este artículo se hace un recuento somero sobre el surgimiento del concepto de Entropía, y algunas de sus relaciones con Sistemas Dinámicos.

## 1. INTRODUCCIÓN

Las máquinas de vapor, invención de James Watt, jugaron un gran papel en la industrialización de nuestra sociedad en el siglo XIX. ¿Cómo y hasta qué punto, se puede mejorar el rendimiento de los motores?. De los estudios que buscaron responder a tales preguntas, nacieron los principios de la termodinámica así como las nociones de energía y entropía, de modo que puede decirse que la termodinámica es la rama de la Física que trata de la energía, el calor y sus interconversiones, la cual fue iniciada por Rudolph Clausius en 1865 al enunciar sus dos primeras “leyes”. El concepto de entropía surge como consecuencia de la segunda ley, y es una medida de la energía no disponible en un sistema termodinámico. En 1872, Ludwig Boltzman logró relacionar la segunda ley de la termodinámica con la imagen atomística del universo que estuvo tan en boga durante el siglo XIX. Boltzman demostró que la entropía es función de un macroestado termodinámico que ocurre de acuerdo con la relación de muchos microestados. En otras palabras, el describía la entropía como una medida del desorden molecular (azar) y lo relacionó con el concepto de probabilidad, inaugurando así la Mecánica Estadística. La teoría de la Información creada por el matemático norteamericano C. Shannon a fines de los años cuarenta del siglo pasado, permite medir la información contenida en mensajes que tienen en principio un significado. Se define la información promedio de un mensaje como igual a la cantidad de azar contenida en la totalidad de mensajes posibles. Así la teoría de la información se ocupa, como la mecánica estadística, de medir las cantidades de azar, y esas dos teorías están por tanto estrechamente ligadas.

## 2. APLICACIÓN DE LA TEORÍA DE LA INFORMACIÓN EN EL ESTUDIO DE LOS SISTEMAS DINÁMICOS: LA ENTROPÍA DE KOLMOGOROV COMO CARACTERÍSTICA DEL CAOS.

Al inicio de los años cincuenta del siglo pasado, Kolmogorov se interesa en la teoría de la información de Shannon. Contrariamente a sus colegas norteamericanos que encontró en el Congreso de Matemáticas en Amsterdam en 1954, él pensaba que no se trataba de un dominio técnico, sino de un tema de Matemáticas puras. Con sus alumnos I. M. Gelfand y A. M. Yaglom al mismo tiempo que A. Yu Khintchine, ellos contribuyen a dar mejores bases matemáticas a la Teoría de la Información. Es entonces natural para Kolmogorov, y una sorpresa para muchos matemáticos, querer codificar las trayectorias de los Sistemas Dinámicos, en el espíritu de la teoría de la información, y de utilizar esa codificación para estudiar un problema esencial

de la Teoría Ergódica, en la cual el problema fundamental es el análisis de las causas que engendran la aparición de leyes estadísticas en los Sistemas Dinámicos. Teoría donde él había hecho sus trabajos precedentes sobre Mecánica Clásica. Ese problema es el de el isomorfismo de Sistemas Dinámicos, es decir, la posibilidad de poner en correspondencia las trayectorias de sistemas dinámicos diferentes. Ya von Neumann (1932) y von Neumann y Halmos (1942) habían demostrado que sistemas dinámicos isomorfos tenían una característica común: el espectro. Esto resultaba de la introducción de métodos del Análisis Funcional para la descripción de los sistemas dinámicos.

Es bien sabido que todas las matrices que se deducen unas de otras por medio de transformaciones ortogonales de bases, son equivalentes y tienen el mismo espectro discreto. Ese espectro (valores propios), no es otro que la diagonal de una de las matrices de la familia, aquella que no tiene elementos no nulos más que sobre la diagonal. Es por tanto natural que todos los sistemas dinámicos que tienen el mismo espectro discreto sean isomorfos. Pero esto deja de ser cierto si el espectro es continuo. Dos sistemas dinámicos teniendo el mismo espectro continuo pueden no ser isomorfos. ¿Cómo distinguirlos?. Es la respuesta a esta vieja pregunta de cerca de 30 años que Kolmogorov da en su trabajo de 1958. El introduce la Entropía como una medida de la aleatoriedad en los Sistemas Dinámicos y muestra que ciertos sistemas dinámicos pueden tener el mismo espectro y entropías diferentes. La entropía permite por tanto distinguir esos sistemas no isomorfos. Cabe señalar además, que fue Kolmogorov quien estableció los fundamentos axiomáticos de la Teoría de la Probabilidad, en donde él hace jugar un papel central a la Teoría de la Medida.

H. Poincaré, es considerado como el padre de la idea de la conexión entre la inestabilidad y estadística en sistemas dinámicos, si bien uno puede encontrar afirmaciones parecidas en los trabajos de otros matemáticos como J. Hadamard, G. Birkhoff, M. Born, y E. Hopf. Esta idea fué expresada con claridad por el físico soviético N. S. Krylov en su obra Trabajos sobre los fundamentos de la Física Estadística publicado en 1950. En la comprensión contemporánea de la conexión entre inestabilidad y la estadística así como en la transición al caos en sistemas dinámicos, el trabajo susodicho de Kolmogorov ha tenido una importancia fundamental.

### 3. SISTEMAS DINÁMICOS

Sea  $M$  un conjunto y  $f : M \rightarrow M$  una transformación. La órbita de un punto  $x \in M$  se define por  $O_f^+(x) = \{x, f(x), f^2(x), \dots\}$  donde  $f^n$  denota la  $n$ -ésima iterada de  $f$ , es decir,  $f^n = f \circ f^{n-1}$ ,  $f^0$  es la identidad en  $M$ . Supongamos que existe un punto  $x_0 \in M$  tal que  $f^n(x_0) = x_0$ . Entonces existe el más pequeño  $n$ , tal que  $p = \min \{n > 0 \text{ tal que } f^n(x_0) = x_0\}$ , y decimos que  $x_0$  es un punto periódico de  $f$  de período  $p$  y al conjunto  $O_f^+(x_0)$  se le llama órbita periódica de  $f$ .

A la pareja  $(M, f)$  se le llama Sistema Dinámico y, la intención de la teoría de los Sistemas Dinámicos es de entender la estructura del conjunto de todas las órbitas de  $(M, f)$ . Si queremos más información acerca de la estructura orbital, necesitamos especificar las características de la pareja  $(M, f)$ .

Esta especificación puede hacerse en muchas maneras diferentes por ejemplo:

- (1) Podemos suponer que  $M$  es un espacio topológico y que  $f$  es continua. El estudio de tales sistemas dinámicos es llamado **Dinámica Topológica**.
- (2) Podemos suponer que  $M$  es una variedad diferenciable real y que  $f$  sea diferenciable. Esta es la **Dinámica Diferenciable**. Además podemos especificar la dimensión de  $M$  y la suavidad de  $M$  y  $f$ .
- (3) Podemos suponer que  $M$  es una variedad diferenciable compleja y  $f$  una transformación holomorfa. Esta es la **Dinámica Holomorfa**.
- (4) Podemos suponer que  $M$  está equipada con una  $\sigma$ -álgebra y una medida  $\mu$  y a  $f$  le pedimos que mantenga invariante a  $\mu$ . Esta es la **Teoría Ergódica**.

Cada una de estas especificaciones tiene múltiples variantes y esos variados dominios están profundamente relacionados. Las teorías correspondientes siguen teniendo más preguntas que respuestas.

**3.1. Entropía Topológica.** La *entropía topológica* es un importante invariante topológico que fué introducido por Adler, Konheim, y MacAndrew en 1965 para funciones continuas de espacios métricos compactos y es una medida de la complejidad dinámica de la función. Más precisamente, mide el promedio de crecimiento cuando  $n$  tiende a infinito del número de órbitas diferentes de longitud  $n$  si usamos una precisión  $\varepsilon$  para distinguir puntos.

3.1.1. *Definición y propiedades elementales.*

3.1. DEFINICIÓN. Sea  $(E, d)$  un espacio métrico compacto  $d$  denota la métrica y  $f : E \rightarrow E$  continua. Un subconjunto  $S \subset E$  se dice  $(n, \varepsilon)$ -separado si para cualesquiera  $x, y \in S$ , existe  $m \in \mathbb{N}$  con  $m \leq n$  tal que  $d(f^m(x), f^m(y)) \geq \varepsilon$ .

Se establece

$$\delta(n, \varepsilon) = \sup_{S \subset E \text{ y } S \text{ } (n, \varepsilon)\text{-separado}} \#S$$

( $\#S$  denota el número de elementos de  $S$  el cual es finito ya que  $E$  es compacto) y, se define la **entropía topológica** de  $f$  denotada  $h(f)$  por la fórmula:

$$h(f) = \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \delta(n, \varepsilon).$$

Algunas propiedades de  $h(f)$  son:

- (1) Se puede demostrar que  $h(f)$  no depende de la métrica  $d$ .
- (2) Es un invariante topológico.
- (3) Para toda  $n \geq 0$ ,  $h(f^n) = nh(f)$  y si  $f$  es invertible,  $h(f^{-1}) = h(f)$ .
- (4) Si  $f : M \rightarrow M$  y  $g : N \rightarrow N$  donde  $M$  y  $N$  son espacios métricos compactos y  $f$  y  $g$  son continuas y semiconjugadas, es decir existe un diagrama conmutativo de la forma

$$\begin{array}{ccc} M & \xrightarrow{f} & M \\ \downarrow l & & \downarrow l \\ N & \xrightarrow{g} & N \end{array}$$

donde  $l$  es continua, entonces se tiene  $h(f) \leq h(g)$  y si  $l$  es un homeomorfismo entonces  $h(f) = h(g)$ .

(5) La entropía es semicontinua inferiormente con la topología  $C^p$ .

3.2. EJEMPLO. (1) Sea  $f : S^1 \rightarrow S^1$ , definido por  $x \mapsto x + \alpha \pmod{1}$ , la rotación en la circunferencia  $S^1$ .

Entonces  $h(f) = 0$ , es decir, los puntos iterados por  $f$  no se separan.

(2) Sea  $f : S^1 \rightarrow S^1$ , definido por  $x \mapsto dx \pmod{1}$  donde  $d \in \mathbb{N} \setminus \{1\}$  entonces  $h(f) = \log d$ .

3.3. OBSERVACIÓN. Hay varias definiciones de “comportamiento caótico” en la literatura, las cuales tratan de acoplarse con la noción intuitiva que tenemos de caos. Diremos que  $f$  es caótica si  $h(f) > 0$  (una manera razonablemente débil de definir el caos de  $f$ )

Misiurewicz describió de una manera muy elegante lo que significa caótico para transformaciones sobre la circunferencia o sobre un intervalo:

3.4. TEOREMA. (Misiurewicz 1980). Sea  $E$  la circunferencia o un intervalo compacto y  $f : E \rightarrow E$  continua, entonces  $f$  es caótica si y solo si existe:

- (1) Un natural  $n$ ,
- (2) un subintervalo  $J$  de  $E$  con interior no vacío,
- (3) Dos subintervalos  $J_0$  y  $J_1$  de  $J$ , con interiores ajenos, tales que ambos  $f^n(J_0)$  y  $f^n(J_1)$  contienen a  $J$ .

Este teorema implica en particular que si  $h(f) > 0$ , entonces  $f$  tiene una infinidad de órbitas periódicas y que el número de órbitas periódicas aumenta exponencialmente con  $n$ .

Otro interesante resultado sobre la entropía topológica es

3.5. TEOREMA. (Sullivan, Milnor, Douady, Hubbard 1993) Sea  $I$  un intervalo compacto y  $f_\mu : I \rightarrow I$  la familia cuadrática  $f_\mu(x) = \mu x(1-x)$ , entonces la transformación  $\mu \mapsto h(f_\mu)$  es continua y (no estrictamente) monótona.

**3.2. Una revisión sobre Teoría Ergódica básica.** Sea  $(X, B, \mu)$  un espacio de probabilidad, es decir,  $X$  es un conjunto,  $B$  es una  $\sigma$ -álgebra de subconjuntos de  $X$ , y  $\mu$  es una medida sobre  $(X, B)$  tal que  $\mu(X) = 1$ . Si  $(X_i, B_i, \mu_i)$ ,  $i = 1, 2$ , son espacios de probabilidad, una transformación  $T : X_1 \rightarrow X_2$  se dice medible si para todo  $A \in B_2, T^{-1}A \in B_1$ . Una transformación medible  $T$  se dice que preserva medida si para todo  $A \in B_2, \mu_1(T^{-1}A) = \mu_2(A)$ . Decimos que  $T$  es una transformación invertible que preserva medida si  $T$  es biyectiva y ambos  $T$  y  $T^{-1}$  preservan medida.

Usamos la notación  $T : (X, B, \mu) \circlearrowleft$  para denotar una transformación que preserva medida (abreviado tpm) de un espacio de probabilidad en sí mismo. Se omite la mención de  $B$  cuando la  $\sigma$ -álgebra es sobre entendida. Por ejemplo, si  $X$  es un espacio topológico, entonces  $B$  siempre es la  $\sigma$ -álgebra de Borel (es decir, la  $\sigma$ -álgebra generada por los conjuntos abiertos) o la  $\sigma$ -álgebra de Borel completada respecto a  $\mu$ . Nos referimos también a  $\mu$  como una medida invariante para  $T$ .

Dos ejemplos conocidos de tpm's o flujos son los sistemas Hamiltonianos y los procesos estocásticos estacionarios. Mencionaremos también los siguientes:

3.6. EJEMPLO. (1) Sea  $S^1 \simeq \{z \in \mathbb{C} : |z| = 1\}$  y sea  $\mu$  la medida de Lebesgue en  $S^1$ . Entonces para  $\alpha \in \mathbb{R}$  fija,  $T : z \mapsto \exp(2\pi i \alpha) z$ , y  $T : z \mapsto z^2$  son tpm's.

- (2) Sea  $T : [-1, 1] \circlearrowleft$  dada por  $Tx = 1 - 2x^2$ . Entonces la medida con densidad  $1/\pi \cos(\arcsin x)$  es una tpm. Una forma de ver esto es vía el cambio de coordenadas  $x = h(\theta) = \sin \frac{\pi}{2}\theta$ . Es suficiente verificar que  $g := h^{-1} \circ f \circ h$  satisface  $|g'(\theta)| = 2$  y por tanto preserva la medida de Lebesgue.
- (3) Sea  $\mu_0$  la medida de probabilidad en  $X_0 := \{1, \dots, k\}$  con  $\mu_0\{i\} = p_i$ , y sea  $\mu$  la medida producto de  $\mu_0$  en  $X := \prod_{-\infty}^{\infty} X_0$ . Sea  $T : X \circlearrowleft$  el operador de corrimiento, es decir, si  $(\underline{x})_i$  denota la  $i$ -ésima coordenada de  $\underline{x} \in X$ , entonces  $(T\underline{x})_i = (\underline{x})_{i+1}$ . Esto define una tpm en  $X$  llamado el  $(p_1, \dots, p_k)$ -corrimiento de Bernoulli.

Como en cada categoría, no distinguimos entre ciertos objetos. Decimos que  $T_1 : (X_1, B_1, \mu_1) \circlearrowleft$  es isomorfo a  $T_2 : (X_2, B_2, \mu_2) \circlearrowleft$  si existe  $\overline{X}_i \subset X_i$  con  $\mu_i \overline{X}_i = 1, i = 1, 2$ , y una tpm invertible  $S : \overline{X}_1 \rightarrow \overline{X}_2$  tal que  $S \circ T_1 = T_2 \circ S$ .

**3.7. EJEMPLO. La transformación del panadero** Sea  $X = [0, 1] \times [0, 1]$ , y sea  $\mu$  la medida de Lebesgue en  $X$ .

$$\text{Sea } T(x) = \begin{cases} (2x, \frac{1}{2}y) & \text{si } x < \frac{1}{2}, \\ (2x - 1, \frac{1}{2}y + \frac{1}{2}) & \text{si } x > \frac{1}{2}. \end{cases}$$

Entonces  $T$  es isomorfo a el  $(\frac{1}{2}, \frac{1}{2})$ -corrimiento de Bernoulli.

Quizá una de las primeras y más importantes cuestiones fundamentales que uno se puede preguntar acerca de tpm's es, cuando se puede descomponer en más pequeños subsistemas que no interactúen.

**3.8. DEFINICIÓN.**  $T : (X, B, \mu) \circlearrowleft$  es llamada **ergódica** si para todo  $A \in B, T^{-1}A = A$  entonces  $\mu A = 0$  o  $1$ .

Finalizo enunciando el siguiente resultado.

**3.9. TEOREMA.** (Shannon-Breiman-McMillan) Sea  $T : (X, B, \mu) \circlearrowleft$  una tpm ergódica, y sea  $\alpha$  una partición finita. Entonces

$$(1) \quad c \lim_{n \rightarrow \infty} -\frac{1}{n} \log \mu \alpha^n(x) = h(T; \alpha)$$

para casi toda  $x$ .

Este teorema tiene la siguiente interpretación, la cual puede casi ser tomada como la definición de entropía. Dice que para cada tpm ergódica, existe un número  $h$  tal que se cumple lo siguiente: para toda  $\varepsilon$ , si  $\alpha$  es una partición suficientemente fina, entonces existe  $N$  tal que para toda  $n \geq N$ , existe un conjunto  $X_n \subset X$  con  $\mu X_n > 1 - \varepsilon$  tal que  $X_n$  consiste de aproximadamente  $\exp n \left( h \pm \varepsilon \right)$  elementos de  $\alpha^n$ .

## REFERENCIAS

- [1] A. Douady, *Topological Entropy of Unimodal Maps* p. 65-87 in *Real and Complex Dynamical Systems*, NATO ASI Series C: Mathematical and Physical Sciences Vol. 464 1995.
- [2] W. de Melo and S. van Strien, *One dimensional dynamics*, Springer Verlag, Berlin. 1993.
- [3] L-S. Young, *Ergodic Theory of Differentiable Dynamical Systems* p. 293-336 in *Real and Complex Dynamical Systems*, NATO ASI Series C: Mathematical and Physical Sciences Vol. 464 1995

Facultad de Cs. Físico Matemáticas, BUAP.  
Rio Verde y San Claudio, Col. San Manuel,  
Puebla, Pue. C.P. 72570.  
festrada@cfm.buap.mx

# EL DESCENSO TOPOLÓGICO UNIFORME Y LAS RAÍCES DE OPERADORES B-FREDHOLM

GABRIEL KANTÚN MONTIEL  
FCFM - BUAP

RESUMEN. Se estudian algunas propiedades del descenso topológico uniforme y su relación con el hecho de que los operadores B-Fredholm tengan raíces.

## 1. INTRODUCCIÓN

Sea  $X$  un espacio de Banach. Denotaremos por  $B(X)$  el conjunto de todas las transformaciones lineales acotadas  $T : X \rightarrow X$ . Es conocido que  $B(X)$  es un álgebra de Banach. A los elementos de  $B(X)$  les llamaremos operadores.

Para  $T \in B(X)$  y  $M \subseteq X$  un subespacio, escribimos

$$T(M) = \{y \in X : \exists x \in M, Tx = y\}.$$

El rango de  $T$  es el subespacio  $R(T) = T(X)$ . El núcleo de  $T$  es el subespacio  $N(T) = \{x \in X : Tx = 0\}$ .

Para  $T, S \in B(X)$  se cumplen

$$R(TS) \subseteq R(T),$$

$$N(S) \subseteq N(TS).$$

Definamos  $T^0 = I$ , donde  $I \in B(X)$  es la identidad ( $Ix = x$  para todo  $x \in X$ ). Entonces

$$X = R(T^0) \supseteq R(T) \supseteq R(T^2) \supseteq \dots \supseteq R(T^n) \supseteq R(T^{n+1}) \supseteq \dots$$

y también

$$\{0\} = N(T^0) \subseteq N(T) \subseteq N(T^2) \subseteq \dots \subseteq N(T^n) \subseteq N(T^{n+1}) \subseteq \dots$$

El ascenso de  $T$ , denotado por  $a(T)$ , y el descenso de  $T$ , denotado por  $d(T)$ , se definen respectivamente como:

$$a(T) = \inf \{n \in \mathbb{N}_0 : N(T^n) = N(T^{n+1})\},$$

$$d(T) = \inf \{n \in \mathbb{N}_0 : R(T^n) = R(T^{n+1})\},$$

con  $\inf \emptyset = \infty$  y  $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$ .

Sea  $M$  un subespacio de  $X$ , ahora consideremos el espacio cociente  $X/M$ , a la dimensión de  $X/M$  le llamamos la codimensión de  $M$ .

Para nuestro estudio presente, la dimensión del núcleo y la codimensión del rango juegan un papel importante, denotemos por  $\alpha(T)$  la dimensión de  $N(T)$  y por  $\beta(T)$  la dimensión de  $X/R(T)$ .

Decimos que un operador perteneciente a  $B(X)$ , es un operador de Fredholm si:

- (1)  $\alpha(T) < \infty$ ,
- (2)  $\beta(T) < \infty$ .

Para un operador de Fredholm tiene sentido considerar  $\alpha(T) - \beta(T)$ , a esta cantidad se le conoce como el índice del operador de Fredholm  $T$  y se denota por  $i(T)$ . Denotaremos al conjunto de operadores de Fredholm en  $X$  mediante  $\Phi(X)$ .

Hay varias maneras de “generalizar” los operadores de Fredholm. La que estudiaremos fue introducida en 1999 por M. Berkani([2]):

1.1. DEFINICIÓN. Sean  $T \in B(X)$  y  $T_n : R(T^n) \rightarrow R(T^n)$  el operador dado por  $T_n x = Tx$ . Decimos que  $T$  es un operador B-Fredholm si para algún  $n \in \mathbb{N}$  el rango de  $T^n$  es cerrado y  $T_n$  es un operador de Fredholm.

Los operadores B-Fredholm tienen la propiedad muy buena de que si  $T_n$  es un operador de Fredholm entonces  $T_m$  es un operador de Fredholm e  $i(T_n) = i(T_m)$  para todo  $m \geq n$  (véase [2]). Luego, tiene sentido definir el índice del operador B-Fredholm  $T$  como el índice del operador de Fredholm  $T_n$ . Denotaremos al conjunto de los operadores B-Fredholm en  $X$  mediante  $BF(X)$ .

Si  $T$  es un operador B-Fredholm entonces  $i(T^n) = ni(T)$  para  $n \in \mathbb{N}$ . Si  $S, T \in \Phi(X)$  entonces  $TS \in \Phi(X)$  e  $i(TS) = i(T) + i(S)$ , sin embargo, esto no se cumple para  $S, T \in BF(X)$  (véase [3]).

## 2. DESCENSO TOPOLÓGICO UNIFORME

Si  $T$  es un operador B-Fredholm, entonces  $T_n$  es un operador de Fredholm para algún  $n \in \mathbb{N}$ , nos interesa saber, entonces, los valores de  $\alpha(T_n)$  y  $\beta(T_n)$  con respecto de  $R(T^j)$  y  $N(T^k)$  para  $j, k = 0, 1, 2, \dots$

Si el ascenso y descenso de un operador  $T \in B(X)$  son finitos, entonces coinciden. Además, estas cantidades están relacionadas con  $\alpha(T)$  y  $\beta(T)$  como se muestra a continuación ([1, Teorema 3.4]):

- Si  $a(T) < \infty$  entonces  $\alpha(T) \leq \beta(T)$ .
- Si  $d(T) < \infty$  entonces  $\beta(T) \leq \alpha(T)$ .
- Si  $a(T) = d(T) < \infty$  entonces  $\alpha(T) = \beta(T)$ , pudiendo ser infinito.
- Si  $\alpha(T) = \beta(T) < \infty$  y si  $a(T)$  ó  $d(T)$  es finito entonces  $a(T) = d(T)$ .

Además de estas cantidades, en la literatura podemos encontrar  $\alpha_n(T)$  y  $\beta_n(T)$  definidas de la siguiente manera:

- $\alpha_n(T) = \dim \frac{N(T^{n+1})}{N(T^n)}$ ,
- $\beta_n(T) = \dim \frac{R(T^n)}{R(T^{n+1})}$ .

Notemos que  $\alpha(T) = \alpha_0(T)$  y  $\beta(T) = \beta_0(T)$ .

Usando [5, Lema 3.1] obtenemos:

$$\alpha_n(T) = \dim[N(T) \cap R(T^n)] = \alpha(T_n),$$

y por [5, Lema 3.2]:

$$\beta_n(T) = \dim \frac{X}{R(T) + N(T^n)} = \beta(T_n).$$

Notemos que si el ascenso de  $T$  es finito, digamos  $a(T) = n$ , entonces  $\alpha_n(T) = 0$ . Similarmente, si el descenso de  $T$  es finito,  $\beta_n(T) = 0$ . Inspirado en lo anterior, S. Grabiner estudió en [4] el caso en que  $\alpha_n(T)$  se vuelve constante, y no sólo cero. Enseguida presentamos su definición.

2.1. DEFINICIÓN (S. Grabiner). Un operador  $T \in B(X)$  es de descenso topológico uniforme si existe  $d \in \mathbb{N}_0$  tal que  $N(T) \cap R(T^n) = N(T) \cap R(T^{n+1})$  y  $R(T^n)$  es cerrado para todo  $n \geq d$ . Equivalentemente, si  $\alpha_n(T) = \alpha_{n+1}(T)$  para todo  $n \geq d$ .

Es conocido que los operadores B-Fredholm son de descenso topológico uniforme (véase [2]). A continuación presentamos una demostración para hacer más clara la exposición.

2.2. TEOREMA. Sea  $T \in B(X)$  un operador B-Fredholm. Entonces  $T$  es de descenso topológico uniforme.

DEMOSTRACIÓN. Como  $R(T^m) \supseteq R(T^{m+1})$  sabemos que

$$N(T) \cap R(T^{m+1}) \subseteq N(T) \cap R(T^m).$$

Como  $T$  es B-Fredholm,  $\dim N(T) \cap R(T^m) = \dim N(T_m) < \infty$ . Por lo tanto, existe  $d \in \mathbb{N}$  tal que  $N(T) \cap R(T^m) = N(T) \cap R(T^d)$  para toda  $m \geq d$ . Esto es,  $\alpha_m(T) = \alpha_d(T)$  para  $m \geq d$ . Así,  $T$  es de descenso topológico uniforme.  $\square$

El siguiente teorema nos ayudará a estudiar las “raíces” de los operadores B-Fredholm.

2.3. TEOREMA. Sea  $T \in BF(X)$  tal que  $T$  es de descenso topológico uniforme para  $n \geq d$ . Entonces  $\alpha_d(T^n) = n\alpha_d(T)$ .

DEMOSTRACIÓN. Usando repetidamente [6, Teorema 2.1] resulta que si  $M_0 \subset M_1 \subset M_2 \subset \dots \subset M_n$  son subespacios de  $X$ , entonces

$$\dim \frac{M_n}{M_0} = \sum_{k=1}^n \dim \frac{M_k}{M_{k-1}}.$$

Como

$$\alpha_d(T^n) = \dim \frac{N(T^{n(d+1)})}{N(T^{nd})} = \dim \frac{N(T^{nd+n})}{N(T^{nd})},$$

y

$$N(T^{nd}) \subseteq N(T^{nd+1}) \subseteq \dots \subseteq N(T^{nd+n}),$$

entonces

$$\begin{aligned} \alpha_d(T^n) &= \sum_{j=1}^n \dim \frac{N(T^{nd+j})}{N(T^{nd+j-1})} \\ &= \sum_{j=1}^n \alpha_{nd+j-1}(T) \\ &= \sum_{j=1}^n \alpha_d(T) \\ &= n\alpha_d(T). \end{aligned}$$

$\square$

### 3. RAÍCES

A continuación aplicaremos lo desarrollado en la sección anterior a la investigación de las “raíces” de un operador.

3.1. DEFINICIÓN. Sea  $n \in \mathbb{N}$ . Decimos que  $T \in B(X)$  tiene raíz  $n$ -ésima si existe  $S \in B(X)$  tal que  $T = S^n$ .

El estudio de las raíces de un operador está íntimamente ligado de forma natural a las propiedades de ascenso, descenso, dimensión del núcleo y codimensión del rango de que hemos venido hablando en las secciones anteriores.

Si  $T \in B(X)$  es un operador de Fredholm y  $S \in B(X)$  es raíz para  $T$ , entonces es conocido que  $S$  es un operador de Fredholm (véase, por ejemplo, [1]).

El siguiente teorema es un resultado conocido sobre operadores de Fredholm que obtenemos ahora a partir de lo que hemos estado discutiendo.

**3.2. TEOREMA.** Sea  $T \in \Phi(X)$ . Si  $\alpha(T) = 1$  y  $T$  es de ascenso infinito, entonces  $\alpha(T^n) = n$ .

DEMOSTRACIÓN. Consideremos

$$\alpha_n(T) = \dim[R(T^n) \cap N(T)].$$

Como  $R(T^m) \supseteq R(T^{m+1})$  para toda  $m \in \mathbb{N}$ , resulta

$$\alpha_0(T) \geq \alpha_1(T) \geq \dots \geq \alpha_d(T) = \alpha_{d+1}(T) = \dots,$$

donde  $d$  corresponde al descenso topológico uniforme.

Ya que  $\alpha_0(T) = \alpha(T) = 1$ , entonces  $\alpha_d(T) = 0$  ó  $\alpha_d(T) = 1$ . Pero  $\alpha(T)$  es infinito, por lo que  $N(T^n) \neq N(T^{n+1})$  para toda  $n$ , y

$$\dim \frac{N(T^{d+1})}{N(T^d)} \neq 0.$$

Por lo tanto,  $\alpha_d(T) = 1$ .

Ahora, como  $\alpha_0(T) = \alpha_d(T)$ , entonces

$$\alpha(T) = \alpha_d(T).$$

Así, usando el Teorema 2.3,

$$\alpha(T^n) = \alpha_d(T^n) = n\alpha_d(T) = n\alpha(T) = n.$$

□

Ahora estamos en posición de obtener el siguiente teorema, el cual nos da condiciones bajo las cuales un operador B-Fredholm no tiene raíces.

**3.3. TEOREMA.** Sea  $T \in BF(X)$ . Si  $i(T) = \pm 1$  o si  $\alpha_d(T) = 1$ , entonces  $T$  no tiene raíces.

DEMOSTRACIÓN. Supongamos que  $i(T) = \pm 1$  y existe  $S \in B(X)$  y  $n \in \mathbb{N}$  tal que  $T = S^n$ . Como  $T \in BF(X)$ , resulta

$$i(T) = i(S^n) = ni(S).$$

Entonces  $ni(S) = \pm 1$ , por lo que  $n = 1$  y así  $T = S$ . Por lo tanto,  $T$  no tiene raíces.

Ahora supongamos que  $\alpha_d(T) = 1$  y  $T = S^n$ , para  $S \in B(X)$ ,  $n \in \mathbb{N}$ . Como  $T \in BF(X)$  se cumple

$$\alpha_d(T) = \alpha_d(S^n) = n\alpha_d(S).$$

Entonces  $n\alpha_d(S) = \pm 1$ , por lo que  $n = 1$  y resulta que  $T = S$ . Por lo tanto,  $T$  no tiene raíces. □

Como los operadores de Fredholm son B-Fredholm, ahora recuperamos un resultado que podemos encontrar en [7].

**3.4. COROLARIO.** Sea  $T \in \Phi(X)$ . Si  $\alpha(T) = 1$  y  $T$  es de ascenso infinito, entonces  $T$  no tiene raíces.

DEMOSTRACIÓN. Por la demostración del Teorema 3.2, vemos que  $\alpha_d(T) = 1$ , y por el Teorema 3.3 resulta que  $T$  no tiene raíces.  $\square$

3.5. EJEMPLO. Sea  $X = \ell_2$ , el espacio de las sucesiones cuadrado sumables, y consideremos el operador  $T \in B(X)$  definido por

$$T(x_1, x_2, x_3, \dots) = (x_1, x_3, x_4, x_5, \dots).$$

Entonces  $N(T)$  es el subespacio generado por  $(0, x_2, 0, 0, \dots)$ , por lo que  $\alpha(T) = 1$ . Además  $R(T) = X$ , y  $\beta(T) = 0$ . Resulta que  $i(T) = 1$  y por lo tanto  $T$  no tiene raíces.

3.6. EJEMPLO. Aunque hemos considerado operadores sobre espacios de dimensión infinita, también existen operadores en espacios de dimensión finita que no tienen raíces. Por ejemplo, las matrices de la forma

$$T = \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix}$$

no tienen raíces. En efecto, si  $S$  es una matriz tal que  $T = S^n$ , entonces

$$S = \begin{pmatrix} 0 & y \\ 0 & 0 \end{pmatrix},$$

por lo que  $S^2 = 0$ . Luego,  $T$  no tiene raíces.

#### REFERENCIAS

- [1] P. Aiena, *Fredholm and local spectral theory with applications to multipliers*, Kluwer Academic Publishers, Dordrecht, 2004.
- [2] M. Berkani, *On a class of quasi-Fredholm operators*, Integral Equations Oper. Theory, 34 (1999), 244–249.
- [3] M. Berkani, D. Medková, *A note on the index of B-Fredholm operators*, Math. Bohem., 129 (2004), 177–180.
- [4] S. Grabiner, *Uniform ascent and descent of bounded operators*, J. Math. Soc. Japan, 34 (1982), 317–337.
- [5] M.A. Kaashoek, *Ascent, descent, nullity and defect, a note on a paper by A.E. Taylor*, Math. Annalen, 172 (1967), 105–115.
- [6] A.E. Taylor, *Theorems on Ascent, descent, nullity and defect of linear operators*, Math. Ann., 163 (1966), 18–49.
- [7] B. Yood, *Ascent, descent and roots of Fredholm operators*, Studia Math., 158(3) (2003), 219–226.

Facultad de Cs. Físico Matemáticas  
 San Claudio y 18 sur, Ciudad Universitaria, Puebla, México.  
 gkantun@mexico.com



# FUNCIONES ABSOLUTAMENTE CONTINUAS A PARTIR DEL CONJUNTO DE CANTOR

FRANCISCO JAVIER MENDOZA TORRES  
FCFM - BUAP

RESUMEN. Empleando el conjunto de Cantor y conjuntos semejantes al de Cantor se construyen funciones de  $[0, 1]$  sobre  $[0, 1]$ , las cuales son estrictamente crecientes y absolutamente continuas, pero cuyas funciones inversas no son absolutamente continuas.

## 1. INTRODUCCIÓN

La construcción de las funciones que trataremos en este artículo se proponen como ejercicio en la página 56 de [7]. Este ejercicio tiene la intención de mostrar la existencia de funciones que son medibles si el contradominio se considera con la  $\sigma$ -álgebra de Borel, pero que deja de ser medible si la  $\sigma$ -álgebra es la de Lebesgue. Estas funciones tienen otras propiedades interesantes, que no se mencionan en ese ejercicio, son funciones estrictamente crecientes, absolutamente continuas y sus funciones inversas no son absolutamente continuas. Silvia Spataru, matemática rumana, en [6] emplea un *conjunto semejante al de Cantor* para construir una función, diferente a las del ejercicio mencionado, que tiene las propiedades anteriores. Ella señala que no es común encontrar este tipo de funciones en la literatura matemática. Las funciones de este artículo también se construyen a partir del conjunto de Cantor y de conjuntos *semejantes al de Cantor*.

Una función  $f : [a, b] \rightarrow [c, d]$  es *absolutamente continua* si para cada  $\varepsilon > 0$  existe un  $\delta > 0$  tal que para cada colección de subintervalos no traslapados  $\{[t_{i-1}, t_i] : i = 1, 2, \dots, n\}$  de  $[a, b]$ , para la que  $\sum_{i=1}^{i=n} (t_i - t_{i-1}) < \delta$ , se tiene que  $\sum_{i=1}^{i=n} |f(t_i) - f(t_{i-1})| < \varepsilon$ . Existen algunos teoremas que caracterizan a las funciones absolutamente continuas, entre estos tenemos el teorema de Banach-Zaretsky, la versión mencionada abajo se encontrará en [1]. Por otro lado, en [2] se estudian algunas variaciones sobre este teorema, el Teorema A se encuentra en ese artículo. Estos teoremas son empleados en esta exposición.

**Teorema de Banach-Zaretsky.** *Sea  $f : [a, b] \rightarrow [c, d]$  una función continua de variación acotada, sea  $m$  la medida de Lebesgue. Entonces  $f$  es absolutamente continua sobre  $[a, b]$  si y sólo si  $m(f(A)) = 0$  para todo  $A$  medible tal que  $m(A) = 0$ .*

**Teorema A.** *Sea  $f$  una función continua sobre  $[a, b]$ , y sea  $P$  un subconjunto medible Lebesgue de  $[a, b]$  tal que el conjunto  $f([a, b] \setminus P)$  tiene medida cero y  $f$  es diferenciable en cada punto de  $P$ . Entonces una condición necesaria y suficiente para que  $f$  sea absolutamente continua sobre  $[a, b]$  es que exista una función positiva medible Lebesgue  $g$  sobre  $[a, b]$  tal que  $g \geq f'$  en  $P$ .*

En lo que sigue construiremos las funciones señaladas. Denotaremos la medida de Lebesgue como  $m$ .

## 2. CONJUNTOS SEMEJANTES AL DE CANTOR

Construiremos lo que se conoce como *conjuntos semejantes al de Cantor*, esto es, conjuntos  $D^a \subset [0, 1]$  con las mismas propiedades que el conjunto de Cantor, pero con medida positiva.

Recordemos que  $C$ , el conjunto de Cantor, se construye de la siguiente forma:

Sean

$$(1) \quad \begin{aligned} C_0 &= [0, 1] \setminus \left(\frac{1}{3}, \frac{2}{3}\right) \\ C_1 &= C_0 \setminus \left\{ \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right) \right\} \\ &\vdots \\ &\vdots \\ &\vdots \\ C_n &= C_{n-1} \setminus \left\{ \left(\frac{1}{3^{n+1}}, \frac{2}{3^{n+1}}\right) \cup \dots \cup \left(\frac{3^{n+1}-2}{3^{n+1}}, \frac{3^{n+1}-1}{3^{n+1}}\right) \right\} \\ &\vdots \\ &\vdots \end{aligned}$$

Cada  $C_n$  resulta de remover los intervalos medios de longitud  $[1/3]^{n+1}$  de cada subintervalo de  $C_{n-1}$ .

El conjunto de Cantor en  $[0, 1]$  se define como:

$$C = \bigcap_{n=0}^{\infty} C_n,$$

y tiene las siguientes propiedades:

- i) **Es no numerable y medible Lebesgue con medida 0.**
- ii) **Es perfecto.**
- iii) **Es denso en ninguna parte.**

Las demostraciones de estas propiedades se pueden consultar en [3] o [7].

De inicio supondremos que  $a > 0$ . Para construir  $D^a$  seguimos un proceso similar al realizado con el conjunto de Cantor.  $D_0^a$  se obtiene de quitarle a  $[0, 1]$  un intervalo abierto  $I_{01}$  centrado en  $1/2$  de longitud  $a^{-1}$ .  $D_1^a$  es  $D_0^a$  menos los intervalos abiertos ajenos  $I_{11}, I_{12}$  de longitud  $a^{-2}$ , respectivamente centrados en los puntos medios de los intervalos cerrados que componen  $D_0^a$ . Secuencialmente, para construir  $D_n^a$ , le quitamos a  $D_{n-1}^a$  la cantidad de  $2^n$  intervalos abiertos ajenos  $I_{ni}$ ,  $i = 1, 2, \dots, 2^n$ , de longitud  $a^{-(n+1)}$  cada uno, centrados en los puntos medios de los intervalos componentes de  $D_{n-1}^a$ . La idea es elegir  $a$  de tal forma que la longitud total de los intervalos quitados a  $[0, 1]$  sea menor que 1. Esto es:

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{2^n}{a^{n+1}} &= \frac{1}{a} \left( \frac{1}{1 - \frac{2}{a}} \right) \\ &= \frac{1}{a - 2} < 1. \end{aligned}$$

Entonces, para cada  $a > 3$ , tendremos un  $D^a$ , el cual será el complemento de la unión de todos los  $I_{ni}$ ,  $n = 0, 1, 2, \dots$ ;  $i = 1, 2, \dots, 2^n$ . Por las leyes de De Morgan, se tiene que

$$D^a = \bigcap_{n=0}^{\infty} D_n^a.$$

y

$$m(D^a) = 1 - \frac{1}{a-2} > 0.$$

Cada  $D^a$  tiene las mismas propiedades que el conjunto de Cantor, pero con medida positiva. La propiedad de densidad en ninguna parte de  $[0, 1]$  es importante en nuestra exposición. Como ejemplo, la construcción de los  $D_n^4$  es como sigue

$$\begin{aligned} D_0^4 &= [0, 1] \setminus \left(\frac{3}{2^3}, \frac{5}{2^3}\right) \\ D_1^4 &= D_0 \setminus \left\{ \left(\frac{5}{2^5}, \frac{7}{2^5}\right) \cup \left(\frac{25}{2^5}, \frac{27}{2^5}\right) \right\} \\ D_2^4 &= D_1 \setminus \left\{ \left(\frac{9}{2^7}, \frac{11}{2^7}\right) \cup \left(\frac{37}{2^7}, \frac{39}{2^7}\right) \cup \left(\frac{89}{2^7}, \frac{91}{2^7}\right) \cup \left(\frac{117}{2^7}, \frac{119}{2^7}\right) \right\} \\ &\dots \\ &\dots \end{aligned}$$

### 3. LA CONSTRUCCIÓN

Construiremos nuestras funciones a partir de  $C$  y de cada  $D^a$ ,  $a > 3$ . Para facilitar la exposición, consideraremos  $a$  fijo, lo que nos permite tomar a los  $I_{ni}$  no explícitamente dependientes de  $a$ . Esta aparente falta de dependencia también se hace para la función  $g$ , la cual se define a continuación y que será la restricción en  $[0, 1] \setminus D^a$  de las funciones que buscamos.

Sea  $g : [0, 1] \setminus D^a \rightarrow [0, 1] \setminus C$  definida de tal forma que hagamos corresponder linealmente los intervalos que se quitan en la construcción de cada  $D_n^a$  sobre los intervalos que se quitan en  $C_n$ . En la Figura 1 se muestra el caso para  $a = 4$  y cuando  $n = 0, 1$ . Debido a que en la construcción de  $D^a$ ,  $D_n^a$  resulta de quitarle a  $D_{n-1}^a$  la cantidad de  $2^n$  intervalos abiertos disjuntos  $I_{ni}$ ,  $i = 1, 2, \dots, 2^n$ , de longitud igual a  $1/a^{n+1}$  cada uno, tenemos que

$$(2) \quad [0, 1] \setminus D^a = \bigcup_{n=0}^{\infty} \left( \bigcup_{i=1}^{2^n} I_{ni} \right).$$

Sobre cada  $I_{ni}$  la función  $g$  tiene derivada igual a  $(a/3)^{n+1}$ , por lo que la función es estrictamente creciente y biyectiva.

Si  $x_0 \in D^a$ , como  $D^a$  es denso en ninguna parte en  $[0, 1]$ , entonces para toda vecindad  $V(x_0)$  de  $x_0$ , se tiene que  $V(x_0) \cap ([0, 1] \setminus D^a) \neq \emptyset$ , por lo tanto  $x_0$  es punto de acumulación de  $[0, 1] \setminus D^a$ . Por lo anterior, tiene sentido considerar  $g(x-)$  y  $g(x+)$ , respectivamente los límites por la izquierda y por la derecha en  $x_0$ . En el siguiente lema demostraremos su existencia y encontraremos sus valores.

3.1. LEMA. Para cada  $x_0 \in D^a$ ,  $g(x_0-)$  y  $g(x_0+)$  existen y además se tiene que:

- i)  $g(x_0-) = \sup \{g(x) \mid x \in [0, 1] \setminus D^a, x < x_0\}$ ,
- ii)  $g(x_0+) = \inf \{g(x) \mid x \in [0, 1] \setminus D^a, x_0 < x\}$ .

DEMOSTRACIÓN. *i)* Sean  $T = \{x \in [0, 1] \setminus D^a : x < x_0\}$  y  $Z = \{g(x) : x \in T\}$ .

Sea  $x_1 \in (x_0, 1) \cap ([0, 1] \setminus D^a)$ . Se tiene que  $g(x) \leq g(x_1)$  para todo  $x \in T$ . Esto nos dice que  $\sup Z$  existe, llamémosle a éste  $\alpha$ .

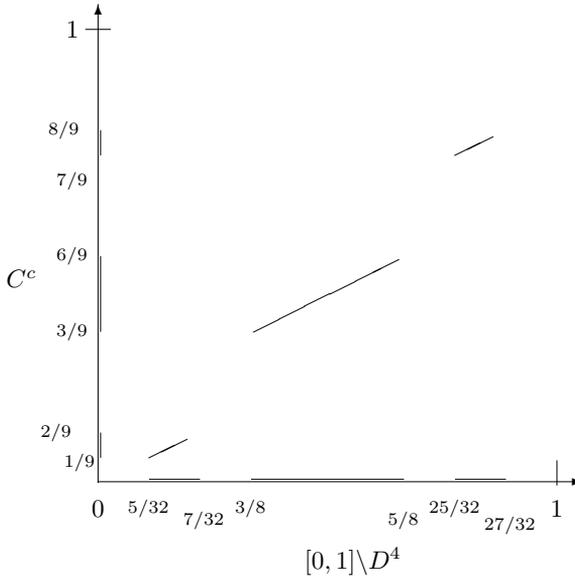


FIGURA 1. Construcción de  $f_4$ , para  $n = 0, 1$ .

Dado  $\varepsilon > 0$ , existe  $t_0 \in T$  tal que

$$\alpha - \varepsilon \leq g(t_0) < \alpha.$$

Como  $g$  es creciente en  $T$ , entonces para cada  $t \in (t_0, x_0) \cap T$  se tiene

$$\alpha - \varepsilon \leq g(t_0) \leq g(t) < \alpha.$$

Esto es;  $|g(t) - \alpha| < \varepsilon$  si  $t \in (t_0, x_0) \cap ([0, 1] \setminus D^a)$ .

Por lo tanto

$$\alpha = g(x_0^-).$$

De forma similar se prueba la igualdad *ii*. □

**3.2. LEMA.** Para cada  $x_0$  en  $D^a$ , existe el límite de  $g(x)$  cuando  $x$  tiende a  $x_0$ .

**DEMOSTRACIÓN.** Sea  $x_0 \in D^a$ . Como  $g(x) \leq g(x')$  para todas  $x, x' \in [0, 1] \setminus D^a$  tales que  $x < x_0 < x'$ , entonces por el Lema 1 se tiene que  $g(x_0^-) \leq g(x_0^+)$ . Supongamos que  $g(x_0^-) < g(x_0^+)$ . Por la construcción de  $g$ , tenemos que  $g(x_0^-)$  y  $g(x_0^+)$  están en  $C$ . Como  $C$  es denso en ninguna parte, entonces existe  $g(x^*) \notin C$ , con  $x^* \in [0, 1] \setminus D^a$  tal que  $g(x_0^-) < g(x^*) < g(x_0^+)$ . Pero por la monotonía de  $g$ ,  $x^*$  debe estar a la izquierda y, a la vez, a la derecha de  $x_0$ , lo cual no puede ser. Por lo tanto  $g(x_0^-) = g(x_0^+)$ . □

**3.3. TEOREMA.** Para cada  $a > 3$  existe una función  $f_a : [0, 1] \rightarrow [0, 1]$  continua y estrictamente creciente tal que  $f_a(D^a) = C$ .

DEMOSTRACIÓN. Sea  $f_a : [0, 1] \rightarrow [0, 1]$  definida como

$$f_a(x) = \begin{cases} g(x) & \text{si } x \in [0, 1] \setminus D \\ g(x_{0-}) & \text{si } x_0 \in D. \end{cases}$$

Por los Lemas 1 y 2,  $f_a$  es continua.

Ya sabemos que  $f_a$  es estrictamente creciente en  $X \setminus D^a$ , veamos que lo es en  $[0, 1]$ . Sean  $x_1 < x_2$  en  $[0, 1]$ . Si  $x_1 \in D^a$  y  $x_2 \in X \setminus D^a$ , entonces por la igualdad (2), existe  $I_{ni}$  para el que  $x_2 \in I_{ni}$ . Luego, por el Lema 1 se tiene que  $f_a(x_1) < f_a(x_2)$ . Cuando  $x_1 \in X \setminus D^a$  y  $x_2 \in D^a$  se hace una prueba similar. Si  $x_1$  y  $x_2$  están en  $D^a$ , como  $D^a$  es denso en ninguna parte, existe  $x_3 \in X \setminus D^a$  tal que  $x_1 < x_3 < x_2$ , de los casos anteriores se tiene que  $f_a(x_1) < f_a(x_2)$ . Por lo tanto  $f_a$  es estrictamente creciente en  $[0, 1]$ . Además observamos que  $f_a(0) = 0$  y  $f_a(1) = 1$ .  $\square$

3.4. TEOREMA.  $f_a$  es absolutamente continua y su función inversa no.

DEMOSTRACIÓN. Sabemos que  $f_a(D^a) = C$  tiene medida de Lebesgue cero. Como el valor de  $f'_a$  sobre cada punto de  $[0, 1] \setminus D^a$  es  $(a/3)^{n+1}$  para algún  $n \in \mathbb{N}$ , entonces, haciendo  $P = [0, 1] \setminus D^a$  en el Teorema A,  $f_a$  es absolutamente continua.

Como  $f_a^{-1} : [0, 1] \rightarrow [0, 1]$  es continua y monótona,  $C \subset [0, 1]$  es tal que  $m(C) = 0$ , y  $m(f_a^{-1}(C)) = m(D^a) = 1 - \frac{1}{a-2} > 0$ , entonces, por el Teorema de Banach-Zaretsky, se sigue que  $f_a^{-1}$  no es absolutamente continua.  $\square$

Tomando en cuenta que el espectro de un operador no acotado, definido en un espacio de Hilbert, se representa como unión de partes, una de las cuales es el espectro absolutamente continuo, el cual a su vez está relacionado a una medida y una función absolutamente continua, el ejemplo aquí estudiado nos puede servir para caracterizar algunos aspectos del espectro de dicho operador.

#### REFERENCIAS

- [1] Benedetto J. y Czaja W., *Integration and Modern Analysis*, Birkhauser, Boston, 2009.
- [2] Cater F., *Some Variations on the Banach Zarecki Theorem*, Real Analysis Exchange, 32, n 2, 2006-2007, pp. 547-552.
- [3] Kannan R. y King Krueger C., *Advanced Analysis on the Real Line*, Springer Verlag, New York, 1996.
- [4] Mendoza Torres F. J. y Mendoza Torres G., *Una función absolutamente continua con inversa no absolutamente continua y preimagen no medible Lebesgue*, Lecturas Matemáticas, 29, n 2, 2008.
- [5] Royden H.L., *Real Analysis, second edition*, Macmillan Publishing, New York, 1968.
- [6] Spataru S., *An absolutely continuous function whose inverse function is not absolutely continuous*, Noti di Matematica. 23, n 1, 2004, 47-49.
- [7] Wilcok H.J. and Myers D.L. *An Introduction to Lebesgue Integration and Fourier Series*, Dover Publications, Inc., New York, 1978.

Facultad de Cs. Físico Matemáticas, BUAP.  
 Río Verde y San Claudio, Col. San Manuel,  
 Puebla, Pue. C.P. 72570.  
 jmendoza@fcfm.buap.mx



# CONTINUIDAD DEL ESPECTRO EN DESPLAZAMIENTOS UNILATERALES DE MULTIPLICIDAD FINITA

SALVADOR SÁNCHEZ PERALES  
FCFM - BUAP

RESUMEN. En este artículo se prueba la continuidad del espectro y del espectro sobreyectivo sobre la clase de operadores que son desplazamientos unilaterales de multiplicidad finita.

Sea  $H$  un espacio de Hilbert de dimensión infinita. Denótese por  $B(H)$  al conjunto de operadores lineales y acotados definidos sobre  $H$  en sí mismo. Este conjunto dotado con la norma  $\|T\| = \inf\{\beta \geq 0 : \|Tx\| \leq \beta\|x\| \text{ para todo } x \in H\}$  es un álgebra de Banach. Sea  $K(H)$  el conjunto de los operadores en  $B(H)$  que son compactos,  $K(H)$  es un ideal bilateral cerrado en  $B(H)$ . Luego el espacio cociente  $B(H)/K(H)$  es un álgebra de Banach con la norma  $\|T+K(H)\| = \inf_{U \in K(H)} \|T+U\|$ .

Este espacio cociente es llamado álgebra de Calkin.

Para  $T \in B(H)$  denotemos por  $N(T)$  al núcleo de  $T$  y  $R(T)$  al rango de  $T$ . Sea  $\alpha(T)$  la dimensión  $N(T)$  y  $\beta(T)$  la dimensión de  $H/R(T)$ . Un operador  $T \in B(H)$  es semi-Fredholm si  $R(T)$  es cerrado, y  $\alpha(T) < \infty$  o  $\beta(T) < \infty$ , cuando se tiene que ambos son finitos se dice que  $T$  es un operador de Fredholm. El índice de un operador semi-Fredholm  $T \in B(H)$  es definido como  $i(T) = \alpha(T) - \beta(T)$ . Dado  $T \in B(H)$ , defínase  $\Phi_-(T) = \{\lambda \in \mathbb{C} \mid \lambda - T \text{ es semi-Fredholm e } i(\lambda - T) < 0\}$  y defínase respectivamente el espectro, el espectro sobreyectivo y el espectro punto aproximado de  $T$  como

$$\begin{aligned}\sigma(T) &= \{\lambda \in \mathbb{C} \mid \lambda - T \text{ no es invertible en } B(H)\}; \\ \sigma_{su}(T) &= \{\lambda \in \mathbb{C} \mid \lambda - T \text{ no es sobreyectiva}\}\end{aligned}$$

y

$$\sigma_{ap}(T) = \{\lambda \in \mathbb{C} \mid \text{ existe } \{x_n\}_{n \in \mathbb{N}} \subseteq H \text{ con } \|x_n\| = 1 \text{ tal que } \|(\lambda - T)x_n\| \rightarrow 0\}.$$

Si  $\pi : B(H) \rightarrow B(H)/K(H)$  es el homomorfismo canónico, entonces el espectro esencial de  $T \in B(H)$ ,  $\sigma_e(T)$ , es el espectro de  $\pi(T)$  en el álgebra de Calkin  $B(H)/K(H)$ .

El espectro  $\sigma$  puede ser visto como una función cuyo dominio es el conjunto de operadores lineales y acotados y cuyo codominio es el conjunto de subconjuntos compactos y no vacíos del plano complejo  $\mathbb{C}$ . Dotado su codominio con la métrica de Hausdorff tenemos que el espectro en general no es continuo. En este artículo examinamos la continuidad del espectro en operadores que son desplazamientos unilaterales de multiplicidad finita.

Un operador  $U \in B(H)$  es un desplazamiento unilateral, si existe una sucesión  $\{H_k\}_{k \in \mathbb{N}}$  de subespacios de  $H$  distintos del espacio trivial  $\{0\}$ , con  $H_j \perp H_k$  si  $j \neq k$ , tal que  $H = \bigoplus_{k=1}^{\infty} H_k$  y para cada  $k \in \mathbb{N}$ ,  $U|_{H_k} : H_k \rightarrow H_{k+1}$  es una isometría

sobreyectiva. Observe que  $\dim H_k = \dim H_{k-1}$  para todo  $k \geq 2$ . Tal dimensión común será llamada la multiplicidad de  $U$ .

1. EJEMPLO. Sea  $\ell^2(\mathbb{N}) = \{\{x_n\}_{n \in \mathbb{N}} \subseteq \mathbb{C} \mid \sum_{n \in \mathbb{N}} |x_n|^2 < \infty\}$ . Éste es un espacio de Hilbert con el producto interior  $\langle \{x_n\}; \{y_n\} \rangle = \sum_{n \in \mathbb{N}} x_n \overline{y_n}$ . Para cada  $k \in \mathbb{N}$ , sea  $e_k = \{\delta_{kj}\}_{j \in \mathbb{N}}$  donde  $\delta_{kj} = 1$  si  $j = k$  y  $\delta_{kj} = 0$  si  $j \neq k$ . El conjunto  $\{e_k\}_{k \in \mathbb{N}}$  es una base ortonormal para  $\ell^2(\mathbb{N})$ .

Considere  $U : \ell^2(\mathbb{N}) \rightarrow \ell^2(\mathbb{N})$  definido por

$$U(x) = (0, x_1, x_2, x_3, \dots), \quad x = (x_1, x_2, x_3, \dots).$$

Entonces  $U$  es un desplazamiento unilateral de multiplicidad 1. En efecto:

- (a)  $\ell^2(\mathbb{N}) = \bigoplus_{k=1}^{\infty} \langle e_k \rangle$  y  $\{\langle e_k \rangle\}_{k \in \mathbb{N}}$  es una sucesión de subespacios de  $\ell^2(\mathbb{N})$  distintos del espacio trivial  $\{0\}$  tales que son ortogonales entre sí.
- (b)  $U|_{\langle e_k \rangle} : \langle e_k \rangle \rightarrow \langle e_{k+1} \rangle$  es una isometría sobreyectiva :  $\|U(xe_k) - U(ye_k)\| = \|U((x-y)e_k)\| = \|(x-y)e_{k+1}\| = |x-y| = \|xe_k - ye_k\|$ .

Una forma de trabajar la convergencia con la métrica de Hausdorff es a través de los conceptos límite inferior y límite superior de una sucesión. Sea  $\{K_n\}$  una sucesión de subconjuntos compactos y no vacíos de  $\mathbb{C}$ , defínase el límite inferior,  $\liminf K_n$ , como el conjunto de los  $\lambda \in \mathbb{C}$  para los cuales existe  $\{\lambda_n\} \subseteq \mathbb{C}$  de manera que  $\lim \lambda_n = \lambda$  y  $\lambda_n \in K_n$  para cualquier  $n \in \mathbb{N}$ , y el límite superior,  $\limsup K_n$ , como el conjunto de los  $\lambda \in \mathbb{C}$  de manera que exista una sucesión creciente de números naturales  $n_1 < n_2 < n_3 < \dots$  y puntos  $\lambda_{n_k} \in K_{n_k}$  tales que  $\lim \lambda_{n_k} = \lambda$ .

Una función  $\tau$ , definida sobre  $B(H)$ , cuyos valores son subconjuntos compactos y no vacíos de  $\mathbb{C}$  se dice que es semicontinua superiormente (inferiormente) en  $A$ , si para cada  $A_n \rightarrow A$ , donde esto significa que  $\lim_{n \rightarrow \infty} \|A_n - A\| = 0$ , se cumple que  $\limsup \tau(A_n) \subseteq \tau(A)$  ( $\tau(A) \subseteq \liminf \tau(A_n)$ ).

Es bien sabido que si  $\tau$  es uniformemente acotada sobre sucesiones convergentes en  $B(H)$ , entonces  $\tau$  es continua con respecto a la métrica de Hausdorff en  $A$  si y sólo si  $\tau$  es semicontinua superiormente e inferiormente en  $A$ .

El resultado principal que exponemos es el siguiente:

**Teorema.** El espectro  $\sigma$  es continua sobre cada desplazamiento unilateral de multiplicidad finita.

Cualquier desplazamiento unilateral  $U$  y su adjunto  $U^*$  están dados por las formulas

$$U(x) = 0 \oplus \bigoplus_{k=2}^{\infty} U_k(x_{k-1}) \quad \text{y} \quad U^*(x) = \bigoplus_{k=1}^{\infty} U_{k+1}^*(x_{k+1})$$

para todo  $x = \bigoplus_{k=1}^{\infty} x_k$  en  $H = \bigoplus_{k=1}^{\infty} H_k$ , con 0 el origen de  $H_1$  y para cada  $k \geq 2$ ,  $U_k = U|_{H_{k-1}}$ . Puesto que  $U_k$  es una isometría, se sigue por [1, Proposición 4.37] que  $\|U_k(x)\| = \|x\|$  para cualquier  $x \in H_{k-1}$ , más aún, siendo  $U_k$  sobreyectiva, tenemos por [1, Proposición 5.73] que  $U_k$  es invertible en  $B(H_{k-1}, H_k)$  y  $U_k^{-1} = U_k^*$ .

Tomemos un desplazamiento unilateral  $U$ . Sea  $\|\cdot\|_{\oplus}$  la norma en el espacio suma  $\bigoplus_{k=1}^{\infty} H_k$ . Como

$$\begin{aligned} \|U(x)\|_{\oplus}^2 &= \|0 \oplus_{k=2}^{\infty} U_k(x_{k-1})\|_{\oplus}^2 \\ &= \sum_{k=2}^{\infty} \|U_k(x_{k-1})\|^2 \\ &= \sum_{k=2}^{\infty} \|x_{k-1}\|^2 \\ &= \|\bigoplus_{k=1}^{\infty} x_k\|_{\oplus}^2 = \|x\|^2. \end{aligned}$$

Entonces  $\|U(x)\| = \|x\|$  para cualquier  $x \in H$ . Sabiendo que la norma de un operador  $T \in B(H)$  también se puede definir como  $\|T\| = \sup \{\|T(x)\| : \|x\| = 1\}$ , hallamos que  $\|U^*\| = \|U\| = 1$ .

2. PROPOSICIÓN. Si  $U$  es un desplazamiento unilateral de multiplicidad finita, entonces  $\pi(U^*)\pi(U) = \pi(U)\pi(U^*) = I$ .

DEMOSTRACIÓN. Para cada  $\bigoplus_{k=1}^{\infty} x_k \in \bigoplus_{k=1}^{\infty} H_k$ ,

$$\begin{aligned} UU^*(\bigoplus_{k=1}^{\infty} x_k) &= U(\bigoplus_{k=1}^{\infty} U_{k+1}^*(x_{k+1})) = 0 \oplus \bigoplus_{k=2}^{\infty} U_k U_k^*(x_k) = \\ &= \bigoplus_{k=2}^{\infty} x_k = I(\bigoplus_{k=1}^{\infty} x_k) - P_{H_1}(\bigoplus_{k=1}^{\infty} x_k), \end{aligned}$$

donde  $P_{H_1}$  es la proyección de  $H$  sobre  $H_1$ . Así  $UU^* = I - P_{H_1}$ . Ahora, dado que el rango de  $P_{H_1}$  es el espacio  $H_1$  y puesto que éste tiene dimensión finita, pues  $U$  es de multiplicidad finita, se sigue que  $P_{H_1}$  es un operador compacto. Así  $\pi(U)\pi(U^*) = \pi(UU^*) = I$ . Por otra parte, como  $U^*U = I$ , obtenemos que  $\pi(U^*)\pi(U) = \pi(U^*U) = I$ .  $\square$

Considere un desplazamiento unilateral  $U$ . Es fácil de probar, usando las formulas de  $U$  dadas anteriormente, que para cada  $\lambda \in \mathbb{C}$ ,  $\lambda - U$  es una transformación lineal inyectiva, de esta forma  $\sigma_{su}(U) = \sigma(U)$ . Más aún, en [1, Ejemplo 6D], hallamos que  $\sigma(U) = \sigma(U^*) = \{\lambda \in \mathbb{C} \mid |\lambda| \leq 1\}$ .

Es bien conocido que si  $T \in B(H)$ , entonces  $\sigma(T) \subseteq B(0, \|T\|)$ , donde  $B(0, \|T\|)$  es el disco en  $\mathbb{C}$  centrado en 0 y con radio  $\|T\|$ .

3. LEMA. Si  $U$  es un desplazamiento unilateral de multiplicidad finita, entonces  $\sigma_e(U) \subseteq \{\lambda \in \mathbb{C} \mid |\lambda| = 1\}$ .

DEMOSTRACIÓN. Supóngase que  $\lambda \in \sigma(\pi(U))$  y que  $|\lambda| \neq 1$ . Por [4, Teorema 3.2.8],  $\lambda - U$  no es un operador de Fredholm, luego  $\lambda - U$  no es invertible. Así,  $\lambda \in \sigma(U) = \{\lambda \in \mathbb{C} \mid |\lambda| \leq 1\}$ . Siendo  $|\lambda| \neq 1$  obtenemos que  $|\lambda| < 1$ . Por otra parte, como  $\|U^*\| = 1$  se tiene que

$$\|\lambda\pi(U^*)\| \leq |\lambda|\|\pi(U^*)\| \leq |\lambda|\|U^*\| < 1.$$

Ahora, dado que  $\sigma(\lambda\pi(U^*)) \subseteq B(0, \|\lambda\pi(U^*)\|)$ , se sigue que  $1 \notin \sigma(\lambda\pi(U^*))$ . De esta forma  $I - \lambda\pi(U^*)$  es invertible. Por otro lado, la Proposición 2 nos dice que  $\pi(U)$  es invertible. Así  $\pi(U)(I - \lambda\pi(U^*))$  es invertible, pero  $\pi(U) - \lambda I = \pi(U)(I - \lambda\pi(U^*))$ . Por lo tanto,  $\lambda I - \pi(U)$  es invertible, lo cual contradice el hecho de que  $\lambda \in \sigma(\pi(U))$ .  $\square$

4. COROLARIO. Si  $U$  es un desplazamiento unilateral de multiplicidad finita, entonces  $\{\lambda \in \mathbb{C} \mid |\lambda| \leq 1\} \subseteq \overline{\Phi_-(U)}$ .

DEMOSTRACIÓN. Sea  $\lambda \in \mathbb{C}$  tal que  $|\lambda| < 1$ . Dado que  $\sigma(U) = \{\lambda \in \mathbb{C} \mid |\lambda| \leq 1\}$ ,  $\lambda - U$  no es invertible. Si  $R(\lambda - U)$  no es cerrado, entonces  $\lambda - U$  no es un operador de Fredholm, así por [4, Theorem 3.2.8],  $\lambda - \pi(U)$  no es invertible. De esta forma,  $\lambda \in \sigma_e(U)$ , luego por Lema 3,  $|\lambda| = 1$ , lo que es una contradicción. En consecuencia  $R(\lambda - U)$  es cerrado. Ahora como  $\lambda I - U$  es inyectiva, se sigue que  $\lambda - U$  es semi-Fredholm e  $i(\lambda - U) = 0 - \beta(\lambda - U) \leq 0$ . Observe que  $i(\lambda - U)$  no puede ser 0, ya que si lo es, entonces  $\lambda - U$  es invertible lo que es falso. Así  $i(\lambda - U) < 0$ , y esto implica que  $\lambda \in \Phi_-(U)$ . Por lo tanto  $\{\lambda \in \mathbb{C} \mid |\lambda| < 1\} \subseteq \Phi_-(U)$ . En consecuencia  $\{\lambda \in \mathbb{C} \mid |\lambda| \leq 1\} \subseteq \overline{\Phi_-(U)}$ .  $\square$

El espectro  $\sigma$  visto como función de  $B(H)$  en el conjunto de subconjuntos compactos y no vacíos de  $\mathbb{C}$  es semicontinuo superiormente en cada  $T \in B(H)$  (véase [3, Theorem 1]). Más aún, usando [2, Lema 1.8] se puede de ver que  $\sigma_{su}$  y  $\sigma_{ap}$  son siempre semicontinuas superiormente en el álgebra  $B(H)$ .

5. TEOREMA. Si  $U$  es un desplazamiento unilateral de multiplicidad finita, entonces  $\sigma_{su}$  y  $\sigma$  son continuos en  $U$ .

DEMOSTRACIÓN. Para probar que  $\overline{\sigma_{su}}$  es continuo en  $U$  es suficiente demostrar que  $\sigma_{su}$  es semicontinua inferiormente en  $U$ . Sea  $\{A_n\}_{n \in \mathbb{N}} \subseteq B(H)$  tal que  $A_n \rightarrow U$ .

Primero verifiquemos que  $\Phi_-(U) \subseteq \liminf \sigma_{su}(A_n)$ . Si existe  $\lambda \in \Phi_-(U)$  tal que  $\lambda \notin \liminf \sigma_{su}(A_n)$ , entonces podemos hallar un  $\epsilon > 0$  y una sucesión creciente de números naturales  $n_1 < n_2 < n_3 \cdots$  tales que  $B(\lambda, \epsilon) \cap \sigma_{su}(A_{n_k}) = \emptyset$  para toda  $k \in \mathbb{N}$ . Como  $\lambda \notin \sigma_{su}(A_{n_k})$ ,  $\lambda - A_{n_k}$  es un operador semi-Fredholm con índice  $i(\lambda - A_{n_k}) = \alpha(\lambda - A_{n_k}) - 0 \geq 0$ . Por otra parte  $\lambda - A_{n_k} \rightarrow \lambda - A$ , así dado que  $\lambda - A$  es semi-Fredholm y por la continuidad del índice, se sigue que  $i(\lambda - A) \geq 0$ , lo cual es falso dado que  $\lambda \in \Phi_-(A)$ . En conclusión  $\Phi_-(U) \subseteq \liminf \sigma_{su}(A_n)$ .

Ahora, se sabe que el límite inferior de cualquier sucesión de subconjuntos compactos es cerrado, de esta forma,  $\Phi_-(U) \subseteq \liminf \sigma_{su}(A_n)$ . Así por Corolario 4, se tiene que

$$\sigma_{su}(U) = \sigma(U) = \{\lambda \in \mathbb{C} \mid |\lambda| \leq 1\} \subseteq \overline{\Phi_-(U)} \subseteq \liminf \sigma_{su}(A_n).$$

Lo cual implica que  $\sigma_{su}$  es continua en  $U$ . La continuidad de  $\sigma$  en  $U$  resulta por lo siguiente:

$$\sigma(U) = \sigma_{su}(U) \subseteq \liminf \sigma_{su}(A_n) \subseteq \liminf \sigma(A_n).$$

$\square$

6. COROLARIO. Si  $U$  es un desplazamiento unilateral de multiplicidad finita, entonces  $\sigma_{ap}$  y  $\sigma$  son continuos en  $U^*$ .

DEMOSTRACIÓN. Sea  $A_n \rightarrow U^*$  y  $\lambda \in \sigma_{ap}(U^*)$ . Como  $\sigma_{ap}(U^*) = \{\bar{\lambda} \mid \lambda \in \sigma_{su}(U)\}$ , entonces existe  $\beta \in \sigma_{su}(U)$  tal que  $\bar{\beta} = \lambda$ . Dado que  $A_n^* \rightarrow U$ , por el Teorema 5, podemos tomar  $\{\beta_n\} \subseteq \mathbb{C}$  tal que  $\bar{\beta}_n \rightarrow \bar{\beta}$  y  $\beta_n \in \sigma_{su}(A_n^*)$ . Observe que  $\sigma_{ap}(A_n) = \{\bar{\lambda} \mid \lambda \in \sigma_{su}(A_n^*)\}$ , así  $\bar{\beta}_n \in \sigma_{ap}(A_n)$  para todo  $n \in \mathbb{N}$ . Esto implica que  $\lambda = \bar{\beta} \in \liminf \sigma_{ap}(A_n)$ . De esta forma  $\sigma_{ap}$  es continuo en  $U^*$ . Ahora, la continuidad de  $\sigma$  en  $U^*$  resulta por lo siguiente:

$$\sigma(U^*) = \sigma_{ap}(U^*) \subseteq \liminf \sigma_{ap}(A_n) \subseteq \liminf \sigma(A_n).$$

$\square$

Sabemos que en la práctica a menudo no es posible encontrar directamente los valores propios, o cualquier otro conjunto espectral, de un operador lineal acotado

que nos interese. En ese caso se busca una manera de "aproximar" en algún sentido el valor espectral que deseamos, y la continuidad del espectro utilizando la métrica de Hausdorff toma un rol esencial. Los desplazamientos unilaterales son operadores de los cuales ya se conoce su espectro o algunas de sus partes, sin embargo, como se sabe estos operadores, por su "sencillez", son usados para construir ejemplos y contraejemplos en la teoría de operadores, y la continuidad del espectro en esta clase de operadores nos provee, pues, de una fuente potencial de ejemplos para estudios más profundos sobre el tema.

## REFERENCIAS

- [1] Carlos S. Kubrusly, *Elements of Operator Theory*, Birkhäuser, 2001.
- [2] J.B. Conway and B.B. Morrel, *Operators that are points of spectral continuity II*, Integral Equations Operator Theory 4 (1981), 459-503.
- [3] J.D.Newburgh, *The variation of spectra*, Duke Math. J. **18** (1951), 165-176.
- [4] S. R. Caradus, W.E. Pfaffenberger, B. Yood, *Calkin Algebras and Algebras of Operators on Banach Spaces*, Marcel Dekker, 1974.

Facultad de Ciencias Físico Matemáticas, BUAP.  
Rio Verde y Av. San Claudio, San Manuel,  
Puebla, Pue. 72570.  
es21254@yahoo.com.mx



# Sesión de Divulgación

Quinta Gran Semana Nacional de la Matemática



# EL RAZONAMIENTO MATEMÁTICO

EDUARDO ARIZA VELÁZQUEZ  
PEDRO GARCÍA JUÁREZ  
ROSA GARCÍA TAMAYO  
CARLOS PALOMINO JIMÉNEZ  
FACULTAD DE CIENCIAS DE LA COMPUTACIÓN - BUAP

RESUMEN. El surgimiento de paradojas (o antinomias) en la periferia de la teoría de conjuntos provoca toda una reconstrucción de la matemática. En medio de la polémica provocada por éstas, surgen tres escuelas del pensamiento sobre los fundamentos de la Matemática: Logicista, Intuicionista y Formalista. Aquí retomamos el análisis realizado por S. C. Kleene, el cual, por un lado hemos depurado y por otro, lo hemos complementado. Al final agregamos una pequeña introducción a las teorías paraconsistentes.

## 1. INTRODUCCIÓN

En general, nuestro conocimiento sobre las paradojas es más bien en un sentido lúdico y sabemos también, porque nos lo han dicho, que no se tienen soluciones aceptables universalmente. Tal vez, alguien más nos contó que éstas representan un parteaguas en la matemática, pero no explican por qué. En todo caso, ¿quién no ha dedicado un tiempo al análisis de alguna paradoja?

El desarrollo del pensamiento matemático da un gran paso al dejar en jaque la teoría de conjuntos de Cantor con la aparición de paradojas. La trascendencia es mucho más de lo que parece pues, como sabemos, ramas importantes de la matemática están fundamentadas sobre la teoría de conjuntos, por ello los conflictos bien pueden ser heredados.

El primer paso hacia la solución del problema planteado por las paradojas es limitado, éste consiste en evitar algunos objetos de la teoría de conjuntos que, aparentemente, están en la periferia. Con esto se buscaba tener una teoría ad hoc. Sin embargo, este primer paso es significativo en el sentido de ser el detonante para observar, no sólo la teoría de conjuntos, sino otras teorías de la matemática y poder dimensionar la magnitud real del problema; que llegó a tocar fondo con el planteamiento de preguntas como ¿Cuál es la naturaleza de la verdad matemática?

El presente trabajo está basado en el análisis hecho por S. C. Kleene [1], el cual hemos actualizado y enriquecido, apoyándonos con el material citado en la bibliografía, con la intención de presentar un panorama más general sobre el razonamiento matemático actual. Por otro lado, se anexa una pequeña introducción a las teorías paraconsistentes, en las que no se busca eliminar las antinomias, sino más bien vivir con ellas.

## 2. LAS PARADOJAS

Las paradojas son afirmaciones que llevan a contradicciones, también suelen usarse las palabras contradicción o antinomia para hacer referencia a éstas. El surgimiento de las paradojas ha llevado a una reestructuración del razonamiento en toda la matemática, en particular en la teoría de conjuntos. Cantor observa la necesidad de trabajar en forma general con el concepto de conjuntos infinitos. En uno de sus trabajos propone la siguiente definición de conjunto.

2.1. DEFINICIÓN. Un conjunto  $M$  es una agrupación de objetos (llamados elementos de  $M$ ), bien definidos y distinguibles de nuestra intuición o de nuestras ideas de un todo.

Tal concepción de conjunto resultó ser punto de inspiración para el surgimiento de algunas paradojas y retomar la discusión sobre otras, ya conocidas. A continuación listamos algunas de las más conocidas, ver también [4], [2], [5].

- (A) La paradoja de Burali Forti 1897, también descubierta por Cantor en 1895, se origina en la teoría de Cantor de los ordinales transfinitos con la siguiente afirmación: *A todo conjunto bien ordenado le corresponde un ordinal.*

Observemos que tal afirmación lleva a una contradicción: Si  $M$  es el conjunto de todos los ordinales, éste es bien ordenado y por lo tanto le corresponde un ordinal. Sea  $\alpha$  tal ordinal, el cual es mayor que cualquier otro ordinal, es decir  $\alpha > \alpha + 1$ . Pero también sabemos que  $\alpha < \alpha + 1$ . De aquí que  $\alpha > \alpha + 1$  y  $\alpha < \alpha + 1$ .

- (B) La paradoja de Cantor (encontrada por él en 1899) surge al considerar lo siguiente: *El conjunto de todos los conjuntos.*

Efectivamente: Sean  $M$  el conjunto de todos los conjuntos y  $\mathcal{P}(M)$  el conjunto potencia de  $M$ . Por el teorema de Cantor,  $|\mathcal{P}(M)| > |M|$ . Por otro lado, como  $M$  es el conjunto de todos los conjuntos entonces  $\mathcal{P}(M)$  está contenido en  $M$ . De aquí que  $|\mathcal{P}(M)| \leq |M|$ , por lo que no puede ocurrir que  $|\mathcal{P}(M)| > |M|$ . Entonces, si  $M$  es el conjunto de todos los conjuntos tenemos que;  $|\mathcal{P}(M)| > |M|$  y, es falso que  $|\mathcal{P}(M)| > |M|$ .

- (C) La paradoja de Russell, descubierta también por Zermelo de forma independiente, se obtiene al considerar: *El conjunto de todos los conjuntos que no son miembros de sí mismos.*

Observemos que si  $T$  es el conjunto de todos los conjuntos que no son miembros de sí mismos entonces  $T \notin T$ . Ahora bien, si  $T$  no forma parte del mismo, entonces  $T$  es un conjunto que se tiene como elemento, es decir  $T \in T$ . Pero  $T \notin T$ , se tiene una contradicción. Otra versión más actual de esta paradoja refiere al tablero de un vehículo conformado por algunas luces que ayudan a mantener el funcionamiento del mismo. Suponga una luz auxiliar en el tablero que se prende cuando alguna de las luces se funde. ¿Es la luz auxiliar parte de las luces del tablero?

- (D) La paradoja de Richard 1905, también dada por Dixón en 1906. Trata con la noción de definibilidad finita. Ésto se refiere a un lenguaje dado (pensemos en el español) con un alfabeto, diccionario y una gramática pre-asignados. El alfabeto consiste del espacio en blanco (para la separación de palabras), las 26 letras latinas, la coma “,” y las vocales acentuadas. Por una expresión en el lenguaje entenderemos una secuencia finita de estos símbolos (cadenas) no iniciando con espacio en blanco. Consideremos la expresión: *Sea  $K$  el número entero que sigue al mayor de los números enteros que se pueden definir con menos de cincuenta palabras.*

Richard reproduce la demostración, diagonal, de la enumerabilidad del conjunto de los números reales de Cantor, aplicada al conjunto de los números reales definibles en francés, que en principio parece claramente enumerable. En eso reside la contradicción.

- (E) Algunas paradojas modernas recaen en cierta medida en la teoría de conjuntos y están relacionadas con una paradoja muy antigua, atribuida al filósofo Epiménides de Creta (siglo 6 a.c.). Ésta se origina con la siguiente declaración que él mismo hace: *Los cretenses son siempre mentirosos.*

Obsérvese que podemos distinguir dos clases de mentirosos: los mentirosos que dicen la verdad alguna vez, y mentirosos de la segunda clase, quienes dicen sólo mentiras. La declaración de Epiménides se refiere a: todos los cretenses son mentirosos de la segunda clase. Supongamos que la declaración es verdadera. Por lo que dice y por el hecho de que él sea cretense, debe ser entonces falso. Ésto es una contradicción. De aquí que por reducción al absurdo, la declaración debe ser falsa. Pero si es falsa, entonces no todos los cretenses son mentirosos y por lo tanto, podría ser verdadera.

La paradoja de Epiménides, conocida también como la paradoja del mentiroso, aparece en forma simple cuando una persona hace la siguiente afirmación: *Esta declaración que estoy diciendo ahora es una mentira.*

La siguiente sentencia se desenvuelve en paradoja, una forma de este enigma, ocurre en la obra “Don Quijote” de Cervantes (1605), II, Pag. 51. *Un viajero ha caído entre caníbales. Ellos le ofrecen la oportunidad de hacer un declaración, sujeto a las condiciones de que si su declaración es verdadera, él será hervido y si es falsa, él será rostizado. ¿Cuál es la declaración que debería hacer el viajero?*

El “dilema del cocodrilo”, semejante al dilema de los caníbales, *un cocodrilo ha robado un niño. El cocodrilo promete al papá regresar al niño, siempre que el padre suponga si el cocodrilo regresará o no al niño. ¿Qué debería hacer el cocodrilo si el papá espera que el cocodrilo no regresará al niño?*

**2.1. Los primeros pasos hacia la solución.** A pesar que no se ha encontrado una solución, universalmente aceptada, al problema de las paradojas, las ideas para resolver el problema de éstas llegaron rápidamente. La primera consideración fue pensar que el error en las paradojas de Buralli Forti, Cantor y Russell consiste en

usar conjuntos muy grandes, tales como el conjunto de todos los conjuntos o el conjunto de todos los números cardinales, o bien en permitir que conjuntos sean considerados miembros de sí mismos.

Esta propuesta, aunque bien intencionada, deja el problema de refundamentar los detalles que no están totalmente implícitos: por ejemplo, si prohibimos el conjunto de todos los números cardinales, no estamos listos para introducir el conjunto de los números naturales a menos que ya sepamos que ellos no son todos los números cardinales. Si prohibimos el conjunto de todos los conjuntos nos encontraremos en conflicto con la definición de conjunto dada por Cantor. Para tener teoría de conjuntos, debemos tener teoremas de todos los conjuntos, todos los conjuntos entonces constituyen un conjunto bajo la definición de Cantor. Si no es así, se debe decir que otra definición de conjunto se usará en su lugar o debemos suplir la definición de Cantor con algún otro criterio para determinar cuando una colección de objetos, como se describe en su definición, constituirá un conjunto (Skolem 1929-1930).

**2.2. Axiomas de la teoría de conjuntos.** Reconstrucciones de la teoría de conjuntos pueden darse poniendo restricciones, para excluir conjuntos muy grandes cuando sean requeridas. En adelante consideraremos que las nociones de teoría de conjuntos son determinadas por axiomas, ya que, como hemos observado, el uso libre en la construcción de conjuntos bajo la definición de Cantor, conlleva inconsistencias. Dichos axiomas son parecidos a aquellos que rigen “punto” y “línea” en la geometría Euclidiana [2], [10], [8].

1. **Axioma de extensión.** Dos conjuntos son iguales si y sólo si tienen los mismos elementos.
2. **Axioma de especificación.** Para cada conjunto  $A$  y para cada condición  $p(x)$  existe un conjunto  $B$  cuyos elementos son exactamente aquellos elementos  $x \in A$  para los cuales  $p(x)$  es verdadera.
3. **Axioma de apareamiento.** Para cualquier par de conjuntos existe un conjunto que los contiene a ambos.
4. **Axioma de la unión.** Para cualquier colección de conjuntos, existe un conjunto que contiene todos los elementos que pertenecen al menos a uno de los conjuntos de la colección.
5. **Axioma de la potencia.** Para cada conjunto existe una colección de conjuntos que contiene a todos los subconjuntos del conjunto dado.
6. **Axioma de infinitud.** Existe un conjunto que contiene al conjunto vacío y al sucesor de todos y cada uno de sus elementos.
7. **Axioma de elección.** Si  $A$  es una colección de conjuntos no vacíos y ajenos dos a dos, entonces existe un conjunto  $C$ , de elección, tal que  $x$  y  $C$  tienen un único elemento en común, para cualquier conjunto  $x$  de  $A$ .
8. **Axioma de sustitución.** Si  $p(a, b)$  es un enunciado tal que para cada  $a \in A$  se puede formar el conjunto  $\{b / p(a, b)\}$ , entonces existe una función  $F$  con dominio  $A$  tal que  $F(a) = \{b / p(a, b)\}$  para cada  $a \in A$ .

### 3. EL FONDO DEL PROBLEMA.

Suponiendo que las paradojas son “evitables” en la axiomatización de la teoría de conjuntos, nos queda la pregunta: ¿esto constituye una solución completa del

problema propuesto por las paradojas?

En el caso de la geometría, los matemáticos han reconocido, debido al descubrimiento de la geometría no euclidiana, la existencia de más de una clase de espacios. Los sistemas de axiomas sirven para singularizar una u otra clase de espacios, o ciertos rasgos comunes de varios espacios a estudiar por el geómetra. Una contradicción que surge en una teoría axiomática formal puede significar, simplemente, que una combinación no realizable de rasgos ha sido postulada.

Sin embargo, en el caso de la aritmética y el análisis, las teorías culminan en la teoría de conjuntos. Los matemáticos anteriores al criticismo (de Kant) suponían que trataban con sistemas de objetos, genéticamente, puestos mediante definiciones, pretendiendo establecer su estructura completa. Además, los teoremas fueron pensados para expresar verdades acerca de estos sistemas, más que como proposiciones hipotéticamente aplicadas a cualquier clase de objetos que satisfaga los axiomas. Pero, a menos que haya un defecto en la lógica o algún error en los métodos de construcción y de razonamiento, acerca de objetos matemáticos en los que hemos confiado, ¿cómo fue posible que surgieran contradicciones en estos objetos?

En un sistema axiomático informal, los axiomas son expresados en el lenguaje ordinario y no hace falta explicitar las reglas de inferencia. Mientras que en un sistema axiomático formal se requiere de la simbolización y especificación de los axiomas y reglas de inferencia, propios del sistema. Sin embargo, la totalidad de la matemática no sólo se debe entender a partir de la axiomatización formal, también debe ser entendida intuitivamente [7].

El problema inmediato, de eliminar las paradojas, emerge con el problema más amplio de los fundamentos de la matemática y la lógica, que bien puede plantearse con las siguientes preguntas.

1. *¿Cuál es la naturaleza de la verdad matemática?*
2. *¿Cuál es el significado que tienen las proposiciones matemáticas y cuál es la evidencia en la que se apoyan?*

Históricamente, estas circunstancias han conducido a estudios más intensos del problema y, por supuesto, las paradojas imponen condiciones. Los intentos en esta dirección fueron a través de las definiciones impredicativas.

**3.1. Definiciones Impredicativas.** Russell fue el primero que llevó su paradoja a un contexto alejado de la teoría de conjuntos, introduciendo las expresiones predicativas. Posteriormente, se observa que esto puede extenderse, alcanzando a otras paradojas, mediante las definiciones impredicativas.

Cuando un conjunto  $M$  y un objeto particular  $m$  son definidos de tal forma que  $m$  es miembro de  $M$  y la definición de  $m$  depende de  $M$ , decimos que el procedimiento (o definición de  $m$ , o la definición de  $M$ ) es *impredicativo*. Similarmente cuando un objeto  $m$  posee una propiedad  $P$ , cuya definición depende de  $P$  (con  $M$  el conjunto de objetos que poseen la propiedad  $P$ ). Una definición impredicativa es circular pues lo que es definido participa en su propia definición.

Cada una de las antinomias, sección 2, conlleva una definición impredicativa. Efectivamente, en la paradoja de Cantor, el conjunto  $M$  de todos los conjuntos incluye como miembro al conjunto  $\mathcal{P}(M)$  definido de  $M$ . El procedimiento impredicativo en la paradoja de Russell destaca en la construcción de  $T$ , pues dividimos el conjunto  $M$  de todos los conjuntos en dos partes; la primera comprendiendo aquellos miembros que se contienen a sí mismos, y la segunda (donde está  $T$ ) conteniendo aquellos que no se contienen. Entonces ponemos  $T$  en  $M$ , para preguntar en cuál parte de  $M$  cayó. Con respecto a la paradoja de Richard, la totalidad de expresiones del lenguaje que la constituyen, es tomada como las expresiones citadas, las cuales se refieren a esa totalidad. En la paradoja de Epiménides, la totalidad de las declaraciones es dividida en dos partes, declaraciones verdaderas y declaraciones falsas. Una declaración que se refiere a esta definición es contada como de la totalidad original, cuando preguntamos si es verdadera o falsa.

Poincaré (1905-1906, 1908) analizó la causas de las paradojas a través de las definiciones impredicativas y Russell (1906-1910) enunció la misma explicación en su principio de *círculo vicioso*:

*Ninguna totalidad puede contener miembros definibles sólo en términos de su totalidad, o miembros envolviendo o presuponiendo su totalidad.*

Así, en primera instancia, parecería que tenemos una solución y una visión adecuada del problema planteado por las paradojas, puesto que el principio del círculo vicioso las excluye. Sin embargo, aunque alentador, no resuelve el problema de fondo. Excluir las paradojas con los argumentos antes mencionados, nos obliga a excluir otros conceptos:

*Parte de la matemática que queremos retener, particularmente análisis, también contiene definiciones impredicativas.*

Como ejemplo está la definición del supremo del conjunto  $M$ ,  $u = \sup(M)$ . Se podría justificar, esta definición impredicativa, interpretándola como una descripción que singulariza el número particular  $u$  de una totalidad ya existente de números reales, y no como definir o crear el número real  $u$ . Pero, el mismo argumento puede ser usado para derivar la definición impredicativa en las paradojas.

**3.2. El intento de Weyl.** El carácter impredicativo de algunas definiciones en el análisis fue especialmente enfatizada por Weyl, en 1918, quien en su libro titulado “Das Kontinuum (El continuo)” se propuso averiguar cuántas definiciones del análisis podrían ser construidas sin que fuesen impredicativas.

Weyl fué capaz de desarrollar una buena parte del análisis, pero no logró un teorema que garantice que todo conjunto acotado superiormente y no vacío,  $M$ , de números reales, tenga una mínima cota superior.

#### 4. ESCUELAS DEL PENSAMIENTO MATEMÁTICO

En medio de la polémica provocada por las paradojas surgen tres escuelas del pensamiento enfocadas en la fundamentación de la matemática: La escuela Logicista

(Russell y Whitehead), la escuela Intuicionista (Brouwer, Dutch) y la escuela Formalista o Axiomática (Hilbert y German). Esta clasificación no incluye otros puntos de vista, debido a que no fueron tan cultivados o no comprenden un grado similar de reconstrucción de la matemática, ni una filosofía que lo soporte.

**4.1. La escuela Logicista.** Según esta corriente, la matemática es una disciplina puramente lógica y como ella sigue enteramente dentro de los límites del lenguaje, no tiene que ver con la realidad ni con la intuición pura. Por el contrario, trata exclusivamente con el uso de signos o de símbolos. Estos signos o símbolos pueden ser usados con entera libertad. El único obstáculo de esta libertad es que bajo ningún concepto pueden contravenirse las reglas establecidas. Leibniz (1666) inicialmente concibe a la lógica como una ciencia que contiene ideas y principios fundamentales de todas las otras ciencias. Dedekind (1888) y Frege (1884, 1893, 1903) estaban comprometidos en definir nociones matemáticas en términos de nociones lógicas, y Peano (1889, 1894-1898) en expresar teoremas matemáticos en un simbolismo puramente lógico.

Para ilustrar cómo las nociones matemáticas pueden ser definidas de nociones lógicas, consideremos la definición de Frege-Russell de un número cardinal, y la definición del número cardinal cero y del número cardinal  $n + 1$  para cualquier número cardinal  $n$ . Entonces, un *cardinal finito* (o número natural) puede ser definido como un número cardinal el cual posee la propiedad  $P$  tal que:

1. 0 tiene la propiedad  $P$  y;
2.  $n + 1$  tiene la propiedad  $P$ , siempre que  $n$  la tenga.

Así, un número natural es definido como un número cardinal para el cual la inducción matemática se cumple. Nótese que esta definición es impredicativa, porque la propiedad de ser un número natural pertenece a la totalidad de las propiedades de números cardinales, que es presupuesta en la definición. El punto de vista dado es muy diferente del que antepone una concepción intuitiva de la sucesión en los números naturales, y extraído del principio inductivo: siempre que una propiedad particular  $P$  de números naturales es dada, satisfaciendo (1) y (2), entonces cualquier número natural dado debe tener la propiedad  $P$ . Aquí en cambio se presupone la totalidad de las propiedades de números cardinales como existían en la lógica antes de la definición de sucesión de número natural.

Para adaptar la construcción logicista, Russell excluyó definiciones impredicativas mediante su *Teoría ramificada de Tipos* (1908-1910): los objetos o individuos primarios, que no están sujetos al análisis lógico, son asignados a un tipo, digamos tipo 0, propiedades de individuos para tipo 1, propiedades de propiedades de individuos para el tipo 2, etc. y no es admitida cualquier propiedad que no caigan en uno de estos tipos. Entonces, para excluir definiciones impredicativas con un tipo, los tipos anteriores son separados en ordenes. Así, las propiedades admitidas sin mencionar ninguna totalidad, pertenecen al orden 0, y propiedades definidas usando la totalidad de propiedades de un orden dado pertenecen al siguiente orden. Las definiciones logicistas de números naturales ahora se convierten en predicativas, cuando la propiedad  $P$  se especifica sólo para propiedades de un orden dado en cuyo caso la propiedad de ser un número natural es del siguiente orden.

Esta separación, en órdenes, hizo imposible construir el análisis familiar que, como hemos visto, contiene definiciones impredicativas. Para escapar de este inconveniente, Russell postuló su *axioma de reductibilidad*, el cual afirma que dada cualquier propiedad perteneciente a un orden anterior al más pequeño, existe una propiedad co-extensiva de orden 0. Si sólo propiedades definibles son consideradas, entonces el axioma significa que para cualquier definición impredicativa con tipo dado, existe una definición predicativa equivalente.

La deducción de la matemática como instancia de la lógica está cimentada en la obra “Principia Mathematica” de Whitehead y Russell (1910-1913). Esta deducción de la matemática fue propuesta como una axiomática intuitiva. Los axiomas fueron pensados para dar por hecho su existencia o al menos para ser aceptados como hipótesis plausibles acerca del mundo. La dificultad ahora es ¿en qué parte del axioma de reductibilidad crearemos?, si las propiedades son para ser construidas, entonces la materia debería ser establecida en la base de construcciones no axiomáticas. Como el autor admite en la introducción de su segunda edición (1925), “Este axioma tiene una justificación puramente pragmática: conduce a resultados deseados y a ningún otro”, claramente no es la clase de axioma con el cual podemos estar satisfechos.

Ramsey, en 1926 clasificó las antinomias conocidas en dos clases; la paradojas “lógicas” (Burali-Forti, Cantor y Russell) y las “epistemológicas” ó “semánticas” (Richard y Epiménides). Además observó, aparentemente, que las antinomias lógicas son superadas por la simple jerarquía de tipos y que las antinomias semánticas pueden derivarse con el lenguaje simbólico mediante la ausencia de los significados requeridos para referirse a expresiones del mismo lenguaje.

*Pero ni Ramsey, Whitehead, ni Russell tuvieron éxito en lograr la meta de tener constructivamente la matemática de la lógica.*

Weyl en 1946, afirma que: *la matemática no se fundamenta en la lógica, sino más bien en un paraíso de lógicos* y observa que todo aquel que esté listo para creer en este “mundo trascendental” también podría aceptar al sistema de la teoría de conjuntos axiomatizada que, para la deducción matemática, tiene la ventaja de ser más simple en estructura.

El logismo trata la existencia de las series de números naturales como una hipótesis acerca del mundo actual, axioma del infinito. Propuestas diferentes del infinito son dadas por los intuicionistas y los formalistas. Desde los puntos de vista intuicionista y formalista, la sucesión de números naturales es más elemental que la noción de número cardinal y sus propiedades, las cuales son utilizados en la caracterización logicista de ellos.

Sin embargo, la tesis logicista es cuestionable al presuponer ideas matemáticas en su formulación. En la perspectiva intuicionista, un núcleo matemático esencial está contenido en la idea de iteración, la cual podría ser usada. Por ejemplo, en la descripción de la jerarquía de tipos o en la noción de una deducción de premisas dadas.

**4.2. La escuela Intuicionista.** Poincaré, al defender la inducción matemática como una herramienta irreducible del razonamiento matemático intuitivo (1902, 1905-1906) es también un precursor de la escuela intuicionista moderna.

Brouwer en su publicación “The untrustworthiness of the principles of logic”, 1908, desafió las creencias de las reglas de la lógica clásica, que esencialmente nos han sido legadas por Aristóteles (384-322 A.C.). Citado de Weyl 1946, según su opinión y lectura de la historia:

*La lógica fué abstraída de la matemática de conjuntos finitos y de sus subconjuntos... Olvidadizos de este origen limitado, confundimos la lógica como algo sobre y antes de toda la matemática. Finalmente la aplicamos, sin justificación, a la matemática de conjuntos infinitos.*

Dos ejemplos inmediatos ilustrarán que los principios válidos acerca de conjuntos finitos no necesariamente se cumplen en conjuntos infinitos. Uno, es el principio de que el todo es más grande que cualquier parte propia. Otro, es que un conjunto de números naturales contiene un máximo.

Un principio de la lógica clásica, válido en el razonamiento sobre conjuntos finitos, que Brouwer no acepta para escenarios infinitos, es la ley del medio excluido. La ley en su forma general dice: para cada proposición  $A$ , se tiene o  $A$  o no  $A$ .

Consideremos la proposición  $A$  como sigue: *Existe un miembro del conjunto  $D$  que tiene la propiedad  $P$* . Luego *no  $A$*  es equivalente a: *Todo miembro de  $D$  no tiene la propiedad  $P$* , o en otras palabras: *Cada miembro de  $D$  tiene la propiedad no  $P$* . La ley, aplicada a la proposición  $A$  dice: *existe un miembro de  $D$  que tiene la propiedad  $P$  o bien, todo miembro de  $D$  tiene la propiedad no  $P$* . De manera natural, tenemos dos casos por analizar.

Primero supongamos que  $D$  es un conjunto finito y que  $P$  es una propiedad tal que, para cualquier miembro en  $D$ , podemos determinar cuando éste tiene la propiedad  $P$  o no. Entonces podríamos examinar cada elemento en  $D$  y encontrar uno que tenga la propiedad  $P$ , o verificar que todos los elementos tienen la propiedad *no  $P$* . Podría ser prácticamente difícil, por ejemplo; cuando  $D$  es un conjunto muy grande teniendo un millón de elementos, o aún para un  $D$  pequeño, el determinar si un elemento tiene o no la propiedad  $P$  podría ser tedioso. Pero la posibilidad de completar la búsqueda, en principio, existe. Esta posibilidad hace que Brouwer acepte la ley del medio excluido como un principio válido por el razonamiento en conjuntos finitos  $D$  y propiedades  $P$  de clase específica.

Ahora, supongamos que  $D$  es infinito, la situación es fundamentalmente diferente: En principio no es posible buscar en todo el conjunto  $D$ , así la búsqueda por sustitución a través de todos los miembros del conjunto infinito  $D$  es imposible. Por ello, en este caso la ley no es aceptada por Brouwer. Podríamos tener éxito en algunos casos, para algunos conjuntos  $D$  y propiedades  $P$ , encontrando un miembro de  $D$  teniendo la propiedad  $P$  y en otros casos, tener éxito buscando por razonamiento matemático que todo miembro en  $D$  tiene la propiedad *no  $P$* , por ejemplo

deduciendo una contradicción del supuesto que un miembro arbitrario de  $D$  tiene la propiedad  $P$ . Un ejemplo para la segunda clase de solución es cuando  $D$  es el conjunto de todos los pares ordenados  $(m, n)$  de enteros positivos y  $P$  es la propiedad:  $m^2 = 2n^2$ . El resultado, ya lo conocemos, Pitágoras descubre que  $\sqrt{2}$  es irracional. Pero en todo caso, no tenemos algo para afirmar la posibilidad de obtener o uno, o lo otro de esta clase de soluciones.

Como otro ejemplo tenemos la Conjetura de Goldbach: *Todo número par mayor que 2 puede representarse como la suma de dos números primos.*

Aquí  $D$  es el conjunto de todas las ternas ordenadas  $(x, y, n)$  de enteros primos positivos  $x, y$ , con  $n \geq 2$  entero par y  $P$  es la propiedad:  $x + y = n$ . A pesar de un gran esfuerzo, hasta la fecha, nadie ha tenido éxito en probar o refutar la conjetura de Goldbach. Por otra parte carecemos del conocimiento de algún método sistemático de búsqueda cuya necesidad en principio lleve a una determinación en cuanto a su verdad o falsedad.

El rechazo a la ley del medio excluido para los sistemas infinitos de Brouwer se basa, fundamentalmente, en la ausencia de un método adecuado para solucionar no sólo todos los problemas matemáticos actuales, sino resolver también los casos excepcionales y cualesquiera otros que se pudieran presentar en el futuro.

**4.2.1. Intuicionistas contra Clásicos.** Llamamos clásica a la matemática familiar, con sus métodos y lógica. Por otro lado, llamamos intuicionista a la matemática que tiene los métodos, lógica de Brouwer y su escuela. La matemática clásica incluye partes que son intuicionistas y partes no intuicionistas. La matemática intuicionista de Brouwer y la matemática no intuicionista, que culminó en las teorías de Weierstrass, Dedekind y Cantor, difieren esencialmente en su opinión del infinito.

En la matemática no intuicionista el infinito es tratado como real, completado, extendido o existencial; un conjunto infinito es considerado existente en una totalidad completa, antes de o independientemente de cualquier proceso humano de generación o construcción. En el caso intuicionista, el infinito es tratado sólo como potencial o constructivo. El reconocimiento de esta diferencia, en el caso de magnitudes infinitas nos remite a Gauss, quien en 1831 escribió, *Protesto... contra el uso de una magnitud infinita como algo completado, lo cual nunca es permitido en la matemática.*

De acuerdo con Weyl 1946, Brouwer aclaró, cuando pensó más allá de cualquier duda, que *no hay pruebas que sustenten la creencia en la calidad existencial de la totalidad de todos los números naturales.* Es decir, la secuencia de números que crece más allá de cualquier etapa alcanzada, pasando al siguiente número, tiene una multitud de posibilidades, que permanece por siempre en el estado de la creación, pero no es un reino cerrado de las cosas que existen en sí mismas. El origen de nuestras dificultades, según Brouwer, está en transformar inconscientemente lo uno en lo otro, además de las antinomias. De esta forma, Brouwer nos hace ver qué tan lejos estamos de la matemática clásica, su crítica aplicada a un conjunto infinito  $D$  (digamos los números naturales) surge debido a su concepción del infinito.

La inducción matemática es un ejemplo de un método intuicionista para probar proposiciones generales de los números naturales. Una demostración por inducción de la proposición, *para todo  $n$ :  $P(n)$* , muestra que dado cualquier  $n$  se debe tener la propiedad  $P$ , mediante el razonamiento que sólo usa los números de 1 a  $n$ . Por supuesto que en una demostración particular por inducción, para ser intuicionista, también los razonamientos usados con su base y paso inductivo deben ser intuicionistas.

Una declaración de existencia tiene su significado intuicionista como una comunicación parcial, o abstracta, de un ejemplo particular, de un número natural  $n$  que tiene la propiedad  $P$  o que al menos da un método mediante el cual en principio, uno pueda encontrar un ejemplo tal. Por lo tanto, una demostración intuicionista de la proposición: *existe  $n$  tal que  $P(n)$* , debe ser “constructiva” en el siguiente sentido; la demostración de hecho exhibe un ejemplo de un  $n$  tal que  $P(n)$ , o al menos indica un método mediante el cual es posible encontrar tal ejemplo.

En la matemática clásica ocurren demostraciones de existencia “no constructivas o indirectas” que no son aceptadas por los intuicionistas. Por ejemplo, para probar: *existe  $n$  tal que  $P(n)$* , los matemáticos clásicos pueden deducir una contradicción de suponer que: *para todo  $n$ , no  $P(n)$* . Bajo las lógicas clásica e intuicionista mediante reductio ad absurdum esto da: *no; para todo  $n$ , no  $P(n)$* . La lógica clásica le permite a este resultado ser transformado en: *existe un  $n$  tal que  $P(n)$* , pero no en la lógica intuicionista (en general). Los intuicionistas se abstienen de aceptar tal demostración de existencia, porque su conclusión *existe  $n$  tal que  $P(n)$*  puede no tener más significado que como una referencia a un ejemplo de un número  $n$  tal que  $P(n)$ , y este ejemplo no ha sido dado. El significado clásico no está disponible para ellos puesto que no conciben a los naturales como una totalidad completada.

Como otro ejemplo de una demostración de existencia no constructiva, suponga que se ha demostrado para una cierta  $P$ , mediante métodos intuicionistas, que si la conjetura de Goldbach es verdadera, entonces el número 5013 tiene la propiedad  $P$  y también, que si la conjetura de Goldbach es falsa, entonces 10 tiene la propiedad  $P$ . Pero con el problema no resuelto de la conjetura de Goldbach, Brouwer anularía tal demostración de existencia porque ningún ejemplo ha sido dado. No sabemos que 5013 es un ejemplo, o no sabemos que 10 es un ejemplo, o no sabemos de algún procedimiento que nos conduciría (aparte de las limitaciones prácticas) a un número particular del que podríamos asegurar es un ejemplo. Brouwer meramente aceptaría que ha sido dada probando la implicación (o declaración condicional): *si  $F$  o no  $F$ , entonces existe un  $n$  tal que  $P(n)$* , donde  $F$  es la declaración de la conjetura de Goldbach. El matemático clásico, mediante esta ley del medio excluido, tiene la premisa  $F$  ó no  $F$  de esta implicación, y así puede inferir su conclusión: *existe  $n$  tal que  $P(n)$* . Pero Brouwer no acepta la premisa  $F$  ó no  $F$ .

Como muestra el ejemplo anterior, los métodos intuicionistas se distinguen de los no intuicionistas tanto en definiciones como en demostraciones. En el presente estado de nuestro conocimiento Brouwer no acepta “el número  $n$  tal que es igual a 5013 si  $F$ , e igual a 10 si no  $F$ ”.

Una disyunción:  $A$  o  $B$  constituye para el intuicionista una comunicación incompleta de una declaración que nos dice que:  $A$  se cumple o que  $B$  se cumple, o al menos hay un método mediante el cual podemos elegir de entre  $A$  y  $B$  una que se cumpla. Una conjunción:  $A$  y  $B$  significa que tanto  $A$  como  $B$  se cumplen. La afirmación:  $A$  implica  $B$ , expresa que  $B$  se sigue de  $A$  mediante un razonamiento intuicionista, es decir que se posee un método el cual desde cualquier demostración de  $A$ , obtendríamos una demostración de  $B$ . Además, una negación:  $\text{no } A$  dice que una contradicción:  $B$  y  $\text{no } B$  se sigue desde  $A$  mediante un razonamiento intuicionista, ó más explícitamente que uno posee un método el cual desde cualquier demostración de  $A$ , obtendríamos la demostración de una contradicción  $B$  y  $\text{no } B$ .

Desde el punto de vista de Brouwer, *la matemática es idéntica con la parte exacta de nuestro pensamiento. . . , ninguna ciencia particular, como la filosofía o la lógica, pueden estar a pre-suposición para la matemática*. Sería circular aplicar cualquier principio filosófico o lógico como significado de demostraciones puesto que concepciones matemáticas son ya presupuestas en la formulación de tales principios. En la matemática intuicionista no se extraen inferencias de acuerdo a normas fijas que pueden ser coleccionadas en una lógica, pero cada inferencia es por sí sola inmediatamente probada en su evidencia. También, existen reglas generales mediante las que, de unos teoremas matemáticos, teoremas nuevos pueden ser formados en una manera intuitivamente clara.

¿Cuán grande es la parte intuicionista que juega en la matemática clásica? A pesar de que ocurran en la teoría clásica elemental de números, esto le permite servir como la primera zona de prueba de investigación en fundamentos que surgen del pensamiento intuicionista y formalista, los métodos no intuicionistas juegan una parte relativamente pequeña. Muchas de las demostraciones de existencia no constructivas pueden ser reemplazadas por constructivas. Sin embargo, en análisis, y todavía en ramas más trascendentales de la matemática, los métodos no intuicionistas, en definiciones y pruebas, permiten la metodología completa.

Los números reales en la representación de las cortaduras de Dedekind son conjuntos infinitos de racionales. Para tratarlos como objetos en el sentido usual, necesitamos usar el infinito completado. En particular, aplicamos la ley del medio excluso para este conjunto, en conexión con las definiciones más simples del sujeto. Por ejemplo, para mostrar que para cualesquiera dos números reales  $x, y$  se cumple que:  $x < y$  ó  $x = y$  ó  $x > y$ , lo utilizamos dos veces. Así, existe un racional  $r$  en  $y$  que no pertenece a  $x$ , o todos los racionales en  $y$  pertenecen a  $x$ , y similarmente cambiando  $x$  por  $y$ . En la definición impredicativa de  $\text{sup}(M)$ , usamos la totalidad de los números reales de la misma manera.

La siguiente pregunta es: ¿qué clase de matemática puede ser construida con las restricciones intuicionistas? Si la matemática clásica existente pudiese ser reconstruida con las restricciones intuicionistas sin grandes cambios ni grandes sacrificios, el problema de sus fundamentos parecería ser resuelto.

Los intuicionistas han creado una matemática totalmente nueva, incluyendo una

teoría del continuo y una teoría de conjuntos (Heyting 1934). Esta matemática emplea conceptos y hace distinciones no encontradas en la matemática clásica; y es muy atractiva por sí misma. Como un sustituto para la matemática clásica ésta ha probado ser menos poderosa y de muchas maneras más complicada de desarrollar; por ejemplo, en la teoría del continuo de Brouwer, no se puede afirmar que cualesquiera dos números reales sean iguales o desiguales. Nuestro conocimiento acerca de la igualdad o desigualdad de  $a$  y  $b$  puede ser más o menos especificada: Mediante “ $a$  distinto de  $b$ ”, se dice que  $a = b$  lleva una contradicción, mientras que “ $a \square b$ ” es una clase más fuerte de desigualdad la cual significa que uno puede dar un ejemplo de un número racional tal que separa  $a$  de  $b$ . Por supuesto  $a \square b$  implica  $a$  distinto de  $b$ . Pero hay parejas de números reales  $a$  y  $b$  para los que no es conocido  $a = b$  ó bien que  $a$  distinto de  $b$  ( $a \square b$ ). A pesar de esto, la posibilidad de una reconstrucción intuicionista de la matemática clásica, en una manera que contenga reinterpretaciones diferentes, no es descartado.

Una controversia que surge entre Brouwer y Hilbert en los primeros años después de que el programa de Hilbert toma forma fue el siguiente: Brouwer en 1923 declaró que *una teoría incorrecta la cual no es detenida por una contradicción no es la menos incorrecta, sólo como un policía criminal no reconocido por una corte represora no es menos criminal.*

Hilbert en 1928 replicó: *negar la ley del medio excluido de los matemáticos sería parecido a negarle al astrónomo el telescopio o al boxeador el uso de sus puños.*

Según Brouwer (1928) y Heyting (1931-1932, 1934), un acuerdo entre intuicionismo y formalismo es posible, siempre y cuando (como en von Neumann 1931-1932) el formalismo refrene de atribuir a la matemática clásica un papel significativo, justificado por las pruebas de consistencia. Tal justificación, según Brouwer, *contiene un círculo vicioso, porque su justificación depende de la correctitud de la proposición más que de la consistencia de una declaración, la correctitud de la declaración se sigue en la correctitud de la “ley del medio excluido”, la cual es parte de la matemática formalista que es justificada.* El punto delicado en la posición formalista es explicar cómo la matemática clásica no intuicionista es significativa, teniendo inicialmente un acuerdo con los intuicionistas de que sus teoremas carecen de un significado real, en términos del cual son verdaderos.

**4.3. Consistencia vía Modelos.** Antes del programa de Hilbert, el método usado en pruebas de consistencia para teorías axiomáticas, fue mediante un “modelo”. Para una teoría axiomática, un modelo es simplemente un sistema de objetos, elegidos de alguna otra teoría, que satisfacen los axiomas. Esto es, para cada objeto o noción primitiva de la teoría axiomática, un objeto o noción de la otra teoría está correlacionado, de tal forma que los axiomas se transforman en (o corresponden a) teoremas de la otra teoría. Si esta última es consistente, entonces así lo debe ser la teoría axiomática. Por supuesto que si la teoría axiomática deduce una contradicción, entonces en la otra teoría, mediante inferencias correspondientes, una contradicción sería deducible de los correspondientes teoremas.

En un famoso ejemplo, Beltrami (1868) mostró que las líneas en el plano de la geometría no euclidiana de Lobatchevsky y Bolyai (el plano de la geometría hiperbólica) puede ser representado por las geodésicas en una superficie de curvatura negativa constante en el espacio euclidiano. Así, el plano de la geometría hiperbólica es consistente, si la geometría Euclidiana es consistente.

Las pruebas de consistencia mediante un modelo son relativas; la teoría, bajo un modelo, es consistente si el modelo tomado es consistente. Sólo cuando esto último es intachable hace que el modelo nos dé una prueba absoluta de consistencia. Veblen y Bussey en 1906 lograron pruebas absolutas de consistencia para ciertas geometrías proyectivas rudimentarias mediante la puesta de modelos usando sólo una clase finita de objetos para representar los puntos. Pero, para probar la consistencia absoluta de teoría de números clásica, análisis y teoría de conjuntos (axiomatizada) el método a través de un modelo “no” ofrece esperanza, como a continuación observamos.

La imposibilidad de trazar un perceptor, o mundo físico, para un modelo es argumentado por Hilbert y Bernays en 1934. Ellos lo ilustraron considerando la primera paradoja de Zenón (Siglo V a.c.) en la que un corredor no puede recorrer una distancia específica en un tiempo finito. Antes de hacerlo, debe correr la primera mitad, después el siguiente cuarto, después el siguiente octavo y así sucesivamente. Para lograrlo, le requeriría un número infinito de pasos. La solución usual de la paradoja consiste en observar que la serie infinita de intervalos de tiempo requeridos para recorrer los segmentos sucesivos converge.

Por lo tanto, *si consistencia debe ser probada en alguna teoría, teoría de números, análisis, etc., ésta debe ser justificada por otro método.*

Esta es la contribución de Hilbert por haber concebido una nueva aproximación directa y por haber reconocido que está involucrada la axiomatización. Este método directo está implícito en el significado de consistencia (al menos como ahora lo pensamos), a saber que ninguna contradicción lógica puede derivarse en la teoría deducida de los axiomas. Así, para probar directamente la consistencia de una teoría, se debería probar una proposición acerca de la teoría misma, específicamente, acerca de todas las pruebas de los teoremas de la teoría. Entonces, la teoría matemática cuya consistencia se busca probar, se transforma en objeto de estudio matemático, a la cual Hilbert llama **Metamatemática o Teoría de Pruebas**.

**4.4. La escuela Formalista.** El método logicista de Russell y Whitehead planteaba nuevos problemas, tres de los cuales presentaban singular relevancia. El primero de ellos hace referencia a la consistencia y trata con la interrogante de si el sistema formal es o no contradictorio. Para no tropezar con un sinnúmero de problemas, la demostración de la consistencia debería realizarse con métodos finitistas.

El segundo problema hace referencia a la axiomatización de las teorías matemáticas: ¿Existe para toda teoría un conjunto finito o, al menos recursivamente, definible de axiomas a partir de los cuales son formalmente derivables todos los teoremas de la teoría en cuestión? ¿Existiría un conjunto de proposiciones aritméticas, tales que todas las sentencias válidas de la teoría de números pudiesen ser obtenidas

mediante la aplicación de unas, determinadas, reglas lógicas de inferencia?

El tercer problema hace referencia a la decidibilidad del sistema y consiste en encontrar, en caso de ser posible, un procedimiento efectivo -un algoritmo en su sentido más amplio- que permitiera decidir en un número finito de pasos si una expresión era o no formalmente derivable a partir de un conjunto de axiomas.

*La resolución de estos tres problemas constituyó el núcleo del famoso programa propuesto por David Hilbert a la Academia de Matemáticas de Hamburgo en 1927.*

Hilbert aceptó que las proposiciones de la matemática clásica, que involucran al infinito completado, van más allá de la evidencia intuitiva. Pero se negó a perder el interés en la matemática clásica, como lo hizo Brouwer. Hilbert trazó una distinción entre declaraciones *reales* y declaraciones *ideales* en la matemática clásica: las *declaraciones reales* son aquellas que están siendo utilizadas por tener un significado intuitivo, mientras que las *declaraciones ideales* son aquellas que no están siendo utilizadas de esta manera. Las declaraciones que tratan al infinito como real son ideales. La matemática clásica adjunta las declaraciones ideales a las reales para retener las reglas simples de la *lógica aristotélica* en el razonamiento acerca de conjuntos infinitos. La adición de *elementos ideales* a un sistema, para completar su estructura y simplificar la teoría, es un procedimiento común actualmente.

Por ejemplo, en el plano de la geometría Euclidiana, dos líneas distintas se intersecan en un único punto, excepto cuando las líneas son paralelas. Para quitar esta excepción, Poncelet en su geometría proyectiva (1822) introduce *un punto en el infinito* en cada una de las líneas originales, tal que líneas paralelas tiene el mismo punto en el infinito y líneas no paralelas tienen distintos puntos en el infinito. La totalidad de estos puntos en el infinito hacen una *línea en el infinito*. Como una línea a través de un punto finito en el plano proyectivo gira, su punto en el infinito traza la línea en el infinito. Mediante este procedimiento, las relaciones de incidencias entre puntos y líneas están simplificadas. Dos puntos distintos determinan una única línea y dos líneas distintas determinan un único punto (el cual está en ambas líneas). Existe el *principio de la dualidad* para el plano de la geometría proyectiva, que dice: *Para cada teorema del tema, la declaración obtenida de él mediante el intercambio de las palabras "punto" y "línea", también es un teorema.*

Como otro ejemplo de la adición de elementos a un sistema, consideremos las extensiones sucesivas de los de números, empezando con los números naturales, después adjuntamos los enteros negativos, siguen las fracciones, introducimos los irracionales y finalmente los números imaginarios. Como bien sabemos, agregar los enteros negativos, a los naturales, simplifica la teoría de adición, haciendo la operación inversa, sustracción, siempre posible, etc.

Una teoría en la matemática clásica puede ser vista como un simple y elegante esquema sistematizado, en el que una variedad de declaraciones reales (presumiblemente) verdaderas, aparentemente heterogéneas y no relacionadas, y con frecuencia previamente desconocidas, son comprendidas como consecuencias de los teoremas ideales en la teoría (Cf. von Neumann 1947, Einstein 1944).

El ejemplo de la teoría analítica de números ilustra que los teoremas de análisis, que carecen de un significado aceptable para el intuicionismo, con frecuencia vinculan teoremas de teoría de números, que sí son intuicionistamente significativos y que de ellos ninguna prueba no-analítica ha sido descubierta, o bien sólo hay pruebas mucho más complicadas.

Para que una teoría sea evaluable, las declaraciones reales comprendidas deben ser verdaderas. Anteriormente los matemáticos lo suponían para garantizar la veracidad de los teoremas que reconocemos como ideales. Ahora, esperamos garantizarlo mediante una prueba de consistencia. Con etapas fáciles de transición, la teorización puede subir a niveles muy altos. Es de interés saber si las construcciones teóricas superiores, que se agregan sucesivamente al cuerpo de proposiciones reales de la clase original, lo hacen tan bien como si realmente permitieran simplificaciones sustanciales de las pruebas de aquellas comprendidas previamente.

Es discutible qué tan justificada está una estructura teórica para sistematizar una clase de verdades reales, es decir, si el análisis clásico está justificado como una sistematización de verdades teórico-numéricas. Históricamente la teoría analítica de números fue un subproducto, así como el ímpetu para el desarrollo del análisis clásico viene de las ciencias y de sus aplicaciones. El análisis sirve como una “formación de ideas” donde las teorías pueden ser expresadas, o reducidas, mediante el método de modelos. Una prueba de la consistencia del análisis, nos aseguraría de la consistencia en la idealización hecha en dicha teoría.

En los años 1926, 1928 y 1931, Weyl observa que la física teórica no separa las declaraciones que están confrontadas con la experiencia, aunque no representan una descripción verdadera de lo que está dado, en teoría, son construcciones puramente simbólicas del mundo. Esto representa una pregunta filosófica profunda de la verdad u objetividad a la que pertenece esta construcción teórico mundial. También, está estrechamente conectado con la pregunta, *¿qué nos motiva a tomar como base un sistema axiomático particular?* Para su consistencia, es argumento necesario pero no suficiente. Cuando la matemática es tomada sólo por sí misma, se restringiría a las verdades intuitivas (Brouwer), no encontraría motivos suficientes para ir más allá. Pero cuando la matemática es unida completamente con la física en el proceso de la construcción teórico-mundial, se está de lado de Hilbert.

Un veredicto en el pensamiento formalista dependerá en parte del fruto del programa propuesto. Este programa se llama *Metamatemática*, en el que se pretende establecer, de forma particular, la consistencia de la matemática clásica. Notemos que en adelante la metamatemática estaría basada en proveer una técnica matemática rigurosa, para investigar una gran variedad de problemas fundamentales, para la matemática y la lógica, entre los que el problema de consistencia es sólo uno de ellos. Por ejemplo, en la actualidad los métodos metamatemáticos están aplicados en estudios de sistematización de matemáticas que nacen de las escuelas logicista, intuicionista y de Hilbert. Inversamente, la metamatemática debe mucho, en su inicio, a las investigaciones logicistas e intuicionistas.

4.4.1. *Formalización de una teoría.* En una teoría matemática, estudiamos un sistema de objetos matemáticos. ¿Cómo puede una teoría matemática ser, en sí misma, objeto para el estudio matemático?. El resultado de tales actividades está incluido en proposiciones, las proposiciones en afirmaciones o teoremas de la teoría matemática dada. No se puede saber en términos exactos qué está en la mente de los matemáticos, pero podemos observar el sistema de tales proposiciones.

El sistema de estas proposiciones debe ser hecho enteramente explícito. No todo de las proposiciones puede ser escrito. La disciplina y el estudiante de la teoría deben proporcionar las condiciones que determinen que proposiciones se cumplen en la teoría. Como primer paso, las proposiciones de la teoría deben ser arregladas deductivamente, algunas con las que otras son lógicamente deducibles, siendo identificadas como los axiomas, o postulados. Este paso no termina hasta que todas las propiedades, términos indefinidos o técnicos de la teoría, que son materia en la deducción de los teoremas que han sido expresados por axiomas, sean aclarados. Explícitamente, el programa de Hilbert consiste de:

- Elementos básicos (indefinidos), axiomas y reglas de inferencia.
- Generación rutinaria de las verdades (demostración de teoremas).
- Los conceptos indefinidos conllevan la existencia de modelos o interpretaciones del sistema de axiomas.
- Un sistema de axiomas será, más o menos, aplicable a un contenido.
- “Todo lo que puede ser objeto de pensamiento científico entra en la esfera del método axiomático”.
- El sistema de axiomas es independiente (el conjunto de axiomas es mínimo).
- Una formalización de toda la matemática: todas las afirmaciones matemáticas deben ser escritas en un lenguaje formal preciso y manipuladas siguiendo reglas bien definidas.
- Completitud. Una demostración de que todas las afirmaciones matemáticas verdaderas pueden ser demostradas dentro del formalismo.
- Consistencia. Una demostración de que en el formalismo de la matemática no se pueden obtener contradicciones. Esta prueba debe usar preferiblemente razonamientos “finitos” acerca de objetos matemáticos finitos.
- Conservación. Una prueba de que cualquier resultado acerca de “objetos reales” obtenido razonando acerca de “objetos ideales” (como conjuntos no numerables) se puede demostrar sin usar objetos ideales.
- Decidibilidad. Debe existir un algoritmo para decidir la verdad o falsedad de cualquier afirmación matemática.

En 1899 Hilbert presenta un sistema axiomático para la geometría. Él y su equipo estaban dedicados al proyecto, sin embargo el programa sólo pudo ser llevado a cabo parcialmente.

## 5. EL TRABAJO DE GÖDEL

En su tesis doctoral [3], Kurt Gödel demuestra la completitud del cálculo de predicados de primer orden, un sistema de axiomas y reglas que permite derivar formalmente todas las proposiciones tautológicas de un lenguaje lo suficientemente expresivo como para que en él puedan ser formalizadas una parte muy considerable de las expresiones matemáticas. El teorema de completitud, poco conocido fuera de

ámbitos especializados, entraña sorprendentes resultados de incompletitud de otros sistemas.

El primero de ellos es la incompletitud del cálculo de predicados de segundo orden, basado en un lenguaje de mayor poder expresivo que el primero, lo que permite dar en él una descripción completa de los números naturales. Puesto que todo teorema de completitud de un cálculo conduce a otro de finitud (o compactidad) y puesto que éste implica que en el lenguaje del cálculo no puede darse una descripción completa de los números naturales (ya que del teorema de finitud se sigue la existencia de conjuntos cuyos elementos satisfacen, aparte de las propiedades descritas, otras que los números no verifican), el cálculo de predicados de segundo orden, en el que los números naturales pueden ser caracterizados, no es posible que sea completo.

Otro resultado de incompletitud que, también a través del teorema de finitud, se deriva de la completitud del cálculo de predicados de primer orden, es la paradoja que ya había sido encontrada por Skolem. Ésta en esencia consiste en el resultado, sólo aparentemente contradictorio, de que existen modelos de la teoría de conjuntos que sólo poseen un conjunto numerable de elementos (conjuntos), pese a que en la referida teoría sea un teorema que existe un conjunto no numerable de conjuntos.

Una tercera consecuencia del teorema de completitud, íntimamente relacionada con las anteriores, es la existencia de un análisis “no estándar”, calificado por Gödel como el análisis del futuro, en el que existen elementos mayores que todo número real y, como consecuencia, otros elementos infinitamente pequeños. El análisis “no estándar”, desarrollado por Robinson a partir de 1961 se ha utilizado con éxito para simplificación de las pruebas de algunos teoremas clásicos e incluso ha permitido encontrar resultados nuevos. En la actualidad se investiga en su aplicación a la demostración automática de teoremas y las perspectivas parecen ser alentadoras.

Así, Gödel demostró que no se podía probar la completitud de ningún sistema formal, no contradictorio, que fuera suficientemente amplio para incluir al menos la aritmética, sólo mediante sus propios axiomas. En 1931 su teorema de la incompletitud mostró que el programa de Hilbert era imposible tal como se planteaba.

Pero, el teorema de completitud no da información respecto de la demostración de la completitud de la matemática mediante un sistema formal diferente. La necesidad de entender el trabajo de Gödel llevó al desarrollo de la teoría de la recursividad y después a la lógica matemática como disciplina autónoma en la década de 1930. De aquí nació directamente la base para la informática teórica de Alonzo Church y Alan Turing.

## 6. LA PROPUESTA PARACONSISTENTE

Sin lugar a duda las contradicciones abundan en las personas, a pesar de que las personas mantienen su razonamiento en forma sensata, esto debido a que no asocian conclusiones arbitrarias a premisas contradictorias. En términos lógicos, es rechazado el principio “Ex falso quodlibet” (de una contradicción, cualquiera puede ser su conclusión). Pero las contradicciones aparecen, no sólo por falta de cuidado

sino por el conocimiento per se, y la Matemática no escapa de tal situación. Por ello se presentan, como una posibilidad, las teorías paraconsistentes [6].

Stanislaw Jaskowski, en 1948, lanza el problema de trabajar con sistemas que incluyan contradicciones y todavía contar con deducciones razonables. Aunque el mismo Jaskowski da respuesta a su planteamiento, sólo lo hace en forma parcial. Sin embargo el trabajo de Jaskowski ha sido importante y todavía es motivo de estudio.

En 1958, Newton Caneiro Affonso da Costa destacó la importancia del estudio de lógicas contradictorias. De acuerdo con él, las teorías contradictorias no pueden ser excluidas a priori, porque la elección de los postulados de una teoría es libre y de hecho, hay teorías cuyas suposiciones iniciales implican contradicciones. En ese mismo año propone el “Principio de Tolerancia en la Matemática”:

*Toda teoría es permisible, si no es trivial.*

Las teorías inconsistentes tienen, entonces, la misma importancia que las consistentes. La particularidad de las teorías inconsistentes es que éstas deben ser basadas en sistemas diferentes a la lógica clásica.

En la propuesta paraconsistente, las antinomias parecerían no causar problemas al tratarlas como contradicciones locales. Es decir que la propuesta es vivir con algunas contradicciones sin que ello implique perder credibilidad.

Da Costa da respuesta al problema planteado por Jaskowski, en 1963 publica su trabajo conteniendo una jerarquía de lógicas de primer orden para el estudio de teorías inconsistentes y las aplicó para la construcción de teoría de conjuntos inconsistentes no triviales. Junto con sus colaboradores, da Costa ha introducido el estudio de diversas lógicas paraconsistentes, teorías de conjuntos, semánticas apropiadas y álgebras asociadas a éstos sistemas, procedimientos de decibilidad, teoría de modelos paraconsistentes, un cálculo diferencial paraconsistente etc.

Actualmente, la paraconsistencia se ha hecho un campo de estudio, con aplicaciones tanto en los fundamentos de la ciencia y del análisis filosófico como en informática y aplicaciones tecnológicas.

## REFERENCIAS

- [1] S. C. Kleene, Nort-Holland, *Introduction to Metamathematics*, 1952.
- [2] J. E. Marsden y M. Hoffman, Addison - Wesley Iberoamericana, *Análisis Clásico Elemental*, Segunda edición.
- [3] Profesor José Fernández Brida, *KURT GÖDEL “Revolucionador de la metamatemática”*, Periódico “EL PAÍS” 17 de Febrero de 1978.
- [4] Bryan H. Bunch, Reverté, *Matemática Insólita paradojas/paralogismos*
- [5] E. Mendelson, *Introduction to Mathematical Logic*. Wadsworth, Belmont, CA, third edition, 1987.
- [6] Eduardo Ariza y José Arrazola, *Topología y Sistemas Dinámicos vol. II, Capítulo 7, Pags, 181-213*, Primera edición, 2009.
- [7] Juan Antonio Valor Yebenes, *Axiomatización conjuntista de las teorías de la ciencia experimental*, Universidad Europea de MAdrid, Departamento de Filosofía, 1998.
- [8] José A. Amor, Servicios Editoriales de la Facultad de Ciencias de la UNAM, *Teoría de Conjuntos para estudiantes de Ciencias* 1997.
- [9] Dirk van Dalen, *Logic and Structure*, Springer-Verlag, 2003, 1986.

[10] J. Dugundji, *Topology*, Allyn and Bacon, Inc., Boston, 1966.

aveariza@hotmail.com

pgarciajz\_2@hotmail.com

tamayorg@hotmail.com

carlos\_cpj@hotmail.com

# EL REFRESCO DE COLA Y LA ESTADÍSTICA

HORTENSIA REYES CERVANTES  
SARA RODRÍGUEZ RODRÍGUEZ  
FCFM - BUAP

RESUMEN. La publicidad de las empresas refresqueras ha ocasionado que la mayoría de la gente guste por ingerir refresco y, en especial, el refresco de Cola, y se sabe que algunas de las marcas más vendidas en México son: Coca-Cola, Pepsi-Cola y Big-Cola. Por otra parte, el consumo excesivo de este tipo de bebidas afecta la salud del ser humano debido a los ingredientes que las constituyen. Una forma de saber qué tanto afectan estas bebidas al cuerpo humano, es por medio de un estudio del pH, que a su vez nos indica el nivel de acidez en disoluciones acuosas. En el presente trabajo se realiza un experimento para descubrir si existen factores que influyan en el nivel del pH, tales como el tipo de refresco y el tiempo en el que se encuentran expuestos a una cierta temperatura. De acuerdo a lo expuesto, se utiliza el diseño que más se acopla al experimento, como es el análisis de varianza.

## 1. INTRODUCCIÓN

El objetivo del trabajo es presentar a la estadística como una de las herramientas matemáticas que nos ayuda a analizar diversos tipos de problemas o asuntos que involucran información sobre lo que se tenga un interés en particular. Comúnmente esto se logra por medio de la recolección de datos, manejo de gráficas, el uso de diversos tipos de técnicas propicias para la información y la experiencia que se tenga sobre el tema.

México tiene el segundo lugar a nivel mundial como consumidor de refrescos de cola, con 112 litros por persona al año, ya que en los últimos 70 años se ha convertido en "parte de la dieta básica del mexicano" [1]. La empresa Coca-Cola controla el 70% de la industria refresquera en México, seguida de Pepsi-Cola con el 15% y Big-Cola con el 5% de la participación total.

El refresco de cola tiene ingredientes que son nocivos para la salud. Uno de los principales compuestos de la bebida de cola es el ácido fosfórico, empleado como aditivo y cuyo efecto en el organismo es desmineralizante de la estructura ósea, ya que el fósforo obstaculiza la absorción de calcio por parte del cuerpo, además que la acidez de esta bebida daña el organismo del ser humano, pues su presencia en el cuerpo provoca que no se puedan absorber todas las vitaminas y minerales que se necesitan para su buen funcionamiento [6]. Estos productos contienen otro elemento llamado gas carbónico, que da el sabor ácido y fresco, y atenúa el gusto dulce en algunos paladares, causando adicción psicológica y problemas físicos en los individuos que la beben [5].

Algunas personas usan estas bebidas como remedio a ciertos males, como por ejemplo, se ha escuchado que si algún individuo tiene la presión arterial baja y se toma un vaso de refresco de cola caliente, esta hace que su presión arterial se normalice. Por lo anterior, resulta interesante construir un experimento para verificar si el nivel de acidez cambia dependiendo de la marca, la temperatura o el tiempo que está destapado el refresco a un nivel de confianza fijado. Se llega a concluir al final de este trabajo, que la acidez del refresco de cola sí depende de los tres factores mencionados.

## 2. REFRESCO DE COLA

En México existen tres compañías importantes que producen este tipo de producto: Coca Cola, Pepsi Cola y Big Cola; aunque existen más, como Chivas Cola, sólo las tres primeras tienen una participación significativa en el mercado, el cual es de tipo oligárquico (ya que no existen muchos productores que satisfagan una demanda grande), donde cada vez es más desigual la competencia ya que los porcentajes de participación de unas compañías se están reduciendo y para otras están aumentando de manera extraordinaria. En el transcurso del año 2005 el porcentaje de control de la industria refresquera estuvo dominado por Coca Cola, quien llegó a la cifra del 70% del mercado nacional, mientras que el 20% se lo reparten entre las otras dos compañías, Pepsi y Big Cola [1]. Este aumento en el porcentaje dominado por Coca Cola tiende a volver a esta competencia un monopolio ya que la supremacía de esta compañía influye de manera notable en el control de precios y volúmenes de producción

En México hace unos años se notaba una preferencia significativa por Pepsi sobre Coca Cola ya que ésta era mucho más dulce y gustaba más entre la población mexicana; poco a poco esta tendencia fue disminuyendo hasta convertirse en el mercado que conocemos actualmente, donde la Coca Cola ha sobrepasado casi por completo a las demás compañías, únicamente la compañía que hace una notable competencia a Coca Cola es Big Cola debido a que el precio al consumidor es mucho más accesible y por esta razón tiene una demanda creciente. Por lo tanto, este mercado ha sufrido una gran transformación ya que antes era un mercado duopolio pues sólo existían dos productores significativos; esta situación se transformó con la entrada de Big Cola que con sus bajos precios amenaza con adquirir una presencia notable en el mercado, mientras que Pepsi ha perdido notablemente su fuerza hasta el momento.

Una bebida de cola es un refresco cuyos ingredientes principales son azúcar, aceites cítricos (naranja, lima o cáscara de limón), canela, vainilla y un saborizante ácido, los cuales en su justa medida y proporción se diluyen en agua carbonatada junto a colorantes y conservantes. Muchos fabricantes de bebidas de cola añaden otros ingredientes a éstos para crear un sabor propio de la marca. Estos incluyen nuez moscada, lavanda y una gran variedad de ingredientes, siendo la base vainilla y canela. La acidez es aportada por el ácido fosfórico, algunas veces acompañado de algún ácido cítrico o aislado.

La fórmula original del refresco de cola fue creada por el farmacéutico John Pemberton y esta bebida se vendía en farmacias. En la elaboración de las primeras

recetas de Coca-Cola se usaban hojas de coca y semillas de Cola acuminata, en una búsqueda de un remedio contra el dolor, pues contenía alcaloides psicoestimulantes. Esto le dio el nombre a Coca-Cola.

Las bebidas de cola son ácidas por contener gas (el ácido carbónico se forma cuando el dióxido de carbono se disuelve en agua), y pueden reaccionar violentamente con bases químicas como el bicarbonato de sodio. Muchas bebidas de cola también contienen ácido fosfórico y/o ácido cítrico, que incrementa altamente la acidez. Las bebidas de cola con ácido fosfórico han sido vinculadas con complicaciones crónicas de riñón, ya que beber dos o más de estas bebidas por día duplica la frecuencia de afecciones de riñón, mientras que bebidas de cola con ácido cítrico no tienen este efecto [6].

**2.1. Potencial de hidrógeno pH.** El pH es una medida de la acidez o basicidad de una solución. En otras palabras, es la concentración de iones hidronio  $[H_3O^+]$  presentes en determinadas sustancias. Esta sigla significa "potencial de hidrógeno", que en 1909, el químico danés Sorensen definió como menos el logaritmo de base 10 de la actividad de los iones hidrógeno,  $pH = -\log_{10}[a_{H_3O^+}]$ . El pH típicamente va de 0 a 14 en disolución acuosa, siendo ácidas las disoluciones con pH menores a 7 y básicas las que tienen pH mayores a 7. El  $pH = 7$  indica la neutralidad de la disolución (donde el disolvente es agua).

<i>Algunos valores comunes del pH</i>	
Sustancia/Disolución	pH
Refresco de cola	2,5
Vanagre	2,9
Cerveza	4,5
Café	5,0
Té	5,5
Saliva (pacientes con cáncer)	4,5 a 5,7
Leche	6,5
Agua pura	7,0
Saliva humana	6,5 a 7,4
Sangre	7,35 a 7,45
Orina	8,0
Agua de mar	8,0
Jabón de manos	9,0 a 10,0

FIGURA 1

El valor del pH se puede medir de forma precisa mediante un potenciómetro, también conocido como pH-metro, un instrumento que mide la diferencia de potencial entre dos electrodos: un electrodo de referencia y un electrodo de vidrio que es sensible al ión hidrógeno. También se puede medir de forma aproximada

empleando indicadores, ácidos o bases débiles que presentan diferente color según el pH. Generalmente se emplea papel indicador, que se trata de papel impregnado de una mezcla de indicadores cualitativos. El papel de litmus o papel tornasol es el indicador mejor conocido en los laboratorios.

### 3. ANÁLISIS DE VARIANZA

Para evaluar la información contenida en una muestra, se efectúa un experimento planeado en el cual los investigadores pueden imponer deliberadamente una o más condiciones experimentales sobre las unidades que se van a analizar, con el fin de determinar su efecto en la respuesta, comparando las respuestas medias  $E(y)$  para varias combinaciones distintas. Las respuestas generadas en una situación experimental controlada siempre exhiben cierta cantidad de variabilidad. El procedimiento estadístico mediante el cual se comparan las medias de la población es el análisis de varianza. Se divide la variación total de las medidas de respuesta en porciones atribuibles a distintos factores de interés para el experimentador.

Supongamos que queremos verificar la cantidad de refresco que contiene una botella y elegimos 5 personas para medirlo, seguramente habrá variación entre las mediciones. Consideremos a la cantidad media de refresco que quiere estimarse como  $\mu$ , una constante desconocida a partir de las observaciones denotadas por  $Y_1, \dots, Y_5$ , que son variables aleatorias independientes que se usan para poder inferir sobre el parámetro. Ahora, como la lectura puede coincidir o discrepar del valor de la longitud en virtud de los errores, se toman  $\epsilon_1, \dots, \epsilon_5$  los errores de medición, entonces cada variable aleatoria se puede representar como:

$$Y_i = \mu + \epsilon_i, i = 1, \dots, 5.$$

Los errores  $\epsilon_i$  deben ser variables aleatorias ya que los  $Y_i$  lo son, mientras que  $\mu$  es una constante. Cualquier inferencia estadística que se quiera hacer sobre  $\mu$  dependerá del modelo probabilístico que se suponga para las variables  $\epsilon_1, \dots, \epsilon_5$ . Se adopta que las observaciones tienen un modelo de distribución Normal con media cero y varianza  $\sigma^2$ . Así, tenemos que:

$$\begin{aligned} E(Y_i) &= E(\mu + \epsilon_i) = \mu + E(\epsilon_i) = \mu, \\ \text{Var}(Y_i) &= \text{Var}(\mu + \epsilon_i) = \text{Var}(\epsilon_i) = \sigma^2. \end{aligned}$$

Por lo cual,

$$Y_i \sim N(\mu, \sigma^2); i = 1, \dots, 5.$$

Notemos que las variables aleatorias  $\epsilon_i$  son no observables en virtud de que el valor  $\mu$  es desconocido y la base de la inferencia es por medio de las variables  $Y_i$ .

Tomando en cuenta la metodología en que se sustenta el análisis de varianza, se tiene que

$$\epsilon_i = Y_i - \mu; i = 1, \dots, n.$$

CUADRO 1. ANOVA

Fac. Var.	G.L.	(SS)	(MS)	$F_0$ .
Media ( $\mu$ )	1	$n(\bar{Y} - \mu_0)^2$	$\frac{n(\bar{Y} - \mu_0)^2}{1}$	$\frac{n(\bar{Y} - \mu_0)^2}{S^2}$
Error	n-1	$\sum_{i=1}^n (Y_i - \bar{Y})^2$	$\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n-1} = S^2$	
Total	n	$\sum_{i=1}^n (Y_i - \mu)^2$		

El error  $i$ -ésimo es la diferencia entre una observación  $i$ -ésima y el valor verdadero del parámetro. Este valor puede expresarse como la diferencia de términos de promedios muestrales, como sigue

$$Y_i - \mu = (Y_i - \bar{Y}) + (\bar{Y} - \mu).$$

De aquí se sigue que el error  $i$ -ésimo está compuesto por la desviación de una observación  $i$ -ésima con respecto a la media muestral, más la desviación entre la media muestral y la media poblacional. A partir de esta partición se realizan transformaciones distribucionales; ésta información se expresa comúnmente usando propiedades probabilísticas de la distribución normal [2], llegando a la tabla de Análisis de Varianza para el modelo  $Y_i = \mu + \epsilon_i$ . Para este caso se establece la prueba e hipótesis  $H_0 : \mu = \mu_0$  en oposición a  $H_a : \mu \neq \mu_0$ , comúnmente llamada ANOVA (véase el Cuadro 1).

Donde se rechaza  $H_0$  si  $F_0 > F_{\alpha,1,n-1}$ , que tiene la regla de decisión al nivel de significancia  $\alpha$ , donde  $n$  son los grados de libertad.

**3.1. Diseños Factoriales.** Para el siguiente caso, se dividen los experimentos tomando en cuenta las diferentes condiciones experimentales y las repeticiones de éstas, es decir, se construyen bloques con el fin de homogeneizar las condiciones experimentales (tratamientos).

Cuando se diseña un experimento se establecen las variaciones que se van a analizar. En nuestro caso, el diseño depende de tres factores  $A, B, C$  donde  $A$  representa la marca del refresco,  $B$  la temperatura que tienen y  $C$  los minutos que están destapados. Los resultados y el análisis estadístico para un experimento factorial se realiza a partir de, Factor  $A$ , que tiene tres diferentes niveles  $A_1 = Coca - Cola$ ,  $A_2 = Pepsi - Cola$  y  $A_3 = Big - Cola$ , Factor  $B$ , que involucra la temperatura del refresco medida en grados Celsius, con tres niveles  $B_1 = 2^\circ$ ,  $B_2 = 10^\circ$  y  $B_3 = 18^\circ$  y Factor  $C$  que es el tiempo que están destapados los refrescos, también con tres niveles  $C_1 = 0$ ,  $C_2 = 10\text{min}$  y  $C_3 = 20\text{min}$ .

Tal experimento factorial requiere  $3 * 3 * 3$  combinaciones o condiciones experimentales llamadas celdas; en cada una de estas celdas se tienen 2 observaciones, por lo cual, el número total de datos es  $54 = 3 * 3 * 3 * 2$ , cada observación se denota como  $y_{ijkl}$ , con  $i = 1, \dots, a = 3$ ,  $j = 1, \dots, b = 3$ ,  $k = 1, \dots, c = 3$ ,  $l = 1, n = 2$ . Así que el modelo estadístico resultante es:

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + \tau_i\beta_j + \tau_i\gamma_k + \beta_j\gamma_k + \tau_i\beta_j\gamma_k + \epsilon_{ijkl}$$

La variable aleatoria  $y_{ijkl}$  depende del valor verdadero que queremos obtener, más la influencia de los factores  $A, B$  y  $C$ , más la influencia de las combinaciones

CUADRO 2. ANOVA Diseños Factoriales

Factor.	G.L.	(SS)	(MS)	$F_0$ .
A	a-1	$\sum_i^a \frac{T_{i...}^2}{bcn} - \frac{T_{....}^2}{N}$	$\frac{SSA}{a-1}$	$F_A = \frac{MSA}{MS_\epsilon}$
B	b-1	$\sum_j^b \frac{T_{.j..}^2}{acn} - \frac{T_{....}^2}{N}$	$\frac{SSB}{b-1}$	$F_B = \frac{MSB}{MS_\epsilon}$
C	c-1	$\sum_k^c \frac{T_{..k.}^2}{abn} - \frac{T_{....}^2}{N}$	$\frac{SSC}{c-1}$	$F_C = \frac{MSC}{MS_\epsilon}$
AB	(a-1)(b-1)	$\sum_{i=1}^a \sum_j^b \frac{T_{ij..}^2}{cn} - \frac{T_{....}^2}{N}$	$\frac{SSAB}{(a-1)(b-1)}$	$F_{AB} = \frac{MSAB}{MS_\epsilon}$
AC	(a-1)(c-1)	$\sum_{i=1}^a \sum_k^c \frac{T_{i.k.}^2}{bn} - \frac{T_{....}^2}{N}$	$\frac{SSAC}{(a-1)(c-1)}$	$F_{AC} = \frac{MSAC}{MS_\epsilon}$
BC	(b-1)(c-1)	$\sum_{j=1}^b \sum_k^c \frac{T_{.jk.}^2}{an} - \frac{T_{....}^2}{N}$	$\frac{SSBC}{(b-1)(c-1)}$	$F_{BC} = \frac{MSBC}{MS_\epsilon}$
ABC	(a-1)(b-1)(c-1)	$\sum_i^a \sum_{j=1}^b \sum_k^c \frac{T_{ijk.}^2}{n} - \frac{T_{....}^2}{N}$	$\frac{SSABC}{(a-1)(b-1)(c-1)}$	$F_{ABC} = \frac{MSABC}{MS_\epsilon}$
Error	abc(n-1)	$SST - SABC$	$MS\epsilon = \frac{SS\epsilon}{abc(n-1)}$	
Total	N-1=abcn-1	$\sum_{i=1}^a \sum_j^b \sum_k^c \sum_{l=1}^n y_{ijkl}^2 - \frac{T_{....}^2}{N}$		

de los factores  $A$ ,  $B$  y  $C$  y finalmente se añade el error de medición. Las pruebas de hipótesis que se establecen son para verificar si algunos de los factores influyen en la variabilidad de las mediciones, es decir:

$$\begin{aligned}
 H_0 : \tau = 0 \text{ vs } H_a : \tau \neq 0, \\
 H_0 : \beta = 0 \text{ vs } H_a : \beta \neq 0, \\
 H_0 : \gamma = 0 \text{ vs } H_a : \gamma \neq 0, \\
 H_0 : \tau\beta = 0 \text{ vs } H_a : \tau\beta \neq 0, \\
 H_0 : \tau\gamma = 0 \text{ vs } H_a : \tau\gamma \neq 0, \\
 H_0 : \beta\gamma = 0 \text{ vs } H_a : \beta\gamma \neq 0, \\
 H_0 : \tau\beta\gamma = 0 \text{ vs } H_a : \tau\beta\gamma \neq 0.
 \end{aligned}$$

La tabla de la ANOVA para este diseño factorial se presenta en el Cuadro 2.

Las regiones de rechazo para las respectivas pruebas de hipótesis son:

$$F_A > F_{\alpha, a-1, abc(n-1)},$$

$$F_B > F_{\alpha, b-1, abc(n-1)},$$

$$F_C > F_{\alpha, c-1, abc(n-1)},$$

$$F_{AB} > F_{\alpha, (a-1)(b-1), abc(n-1)},$$

$$F_{AC} > F_{\alpha, (a-1)(c-1), abc(n-1)},$$

$$F_{BC} > F_{\alpha, (b-1)(c-1), abc(n-1)},$$

$$F_{ABC} > F_{\alpha, (a-1)(b-1)(c-1), abc(n-1)}.$$

Para el caso de comparar dos medias poblacionales,  $\mu_i$  y  $\mu_j$ , el intervalo de confianza es

$$(\bar{x}_i - \bar{x}_j) \pm t_{\alpha/2} \sqrt{s^2 \left( \frac{1}{n_i} + \frac{1}{n_j} \right)}$$

Con  $s = \sqrt{s^2} = \sqrt{MSE}$

4. METODOLOGÍA

Para realizar el experimento, se comenzó por definir qué tipo de refresco se compararía y en que presentación se tomaría, ya que éste podría ser otro factor para que el nivel de pH variara.



FIGURA 2

Se usaron refrescos en lata, y se tomaron 2 observaciones por cada tipo de tratamiento (Figura 2), así que se colocaron vasos de la misma medida y se adecuó la temperatura inmediatamente que se destapaba el refresco, seguido de colocar una de las tiras tornasol y medir el pH. La medición se hace comparando la tira tornasol con la tabla contenida en su caja. Este procedimiento se realizó para cada uno de los tres tipos de refresco (Figura 3).



FIGURA 3

Los niveles de pH obtenidos los tenemos en el cuadro 3.

NIVEL DE PH EN REFRESCOS DE COLA									
TIPO DE REFRESCO	TEMPERATURA (CENTIGRADOS)								
	2°			10°			18°		
	TIEMPO (MINUTOS)								
	0	10	18	0	10	18	0	10	18
R1	1.5	2	1.5	2	2	2	2.5	2.5	2.5
	1	1	2	2	2.5	2	2.5	3	2.5
R2	2	2.5	3	2	2.5	3	2.5	3	3
	2	2	3	2.5	3	2.5	2.5	3	3
R3	1.5	2.5	2.5	2.5	3.5	3	3	3	4
	2	2.5	3	2	2	3	3	3	3.5

FIGURA 4. Cuadro 3

## 5. RESULTADOS

Al trabajar con los datos probamos graficamente que cumplen con las hipótesis para realizar un análisis de varianza, es decir, que son independientes y que tienen una distribución normal. Los resultados que se obtienen se encuentran en el Cuadro 4, donde se muestran con un nivel de significancia de  $\alpha = 0.05$ . Tanto el tipo de refresco, como la temperatura y el tiempo que están destapados influyen en el nivel de pH de las bebidas.

FACTOR	SS	G.L.	MS	Fo		
A	4.86	2	2.43	20.95	>	$F_{\alpha,2,27} = 3.354$
B	5.86	2	2.93	25.26	>	$F_{\alpha,2,27} = 3.354$
C	2.86	2	1.43	12.32	>	$F_{\alpha,2,27} = 3.354$
AB	0.79	4	0.198	1.71	<	$F_{\alpha,4,27} = 2.728$
AC	0.78	4	0.195	1.68	<	$F_{\alpha,4,27} = 2.728$
BC	0.45	4	0.113	0.97	<	$F_{\alpha,4,27} = 2.728$
ABC	0.49	8	0.061	0.53	<	$F_{\alpha,8,27} = 2.305$
Error	3.12	27	0.116			
Total	19.21	59				

FIGURA 5. Cuadro 4

Se realizó un análisis de varianzas y luego se prosiguió a comparar las medias de los distintos factores, teniendo como resultado con un nivel de significancia de  $\alpha = 0.05$  que el nivel de pH de Coca-Cola es menor que el de Pepsi-Cola y Big-Cola, es decir, es más ácida. En cuanto a la temperatura, a los 20° el nivel de pH es mayor que cuando están a 2° y 10°. Por último, con un nivel de significancia de  $\alpha = 0.1$  entre más tiempo est'en destapados los refrescos serán menos ácidos.

**De acuerdo al estudio realizado, concluimos que el refresco de cola es dañino para la salud y aún más en la forma que la mayoría de la gente acostumbra a ingerirlo, ya que lo prefieren frío y con gas.**

En relación a la parte matemática, se nota que el modelo es muy fácil de utilizar ya que sólo involucra expresiones sencillas de sumatorias y las tablas de las distribuciones t-student y Fisher. Cabe aclarar que nuestros resultados serán los más cercanos a la realidad siempre y cuando nuestros datos cumplan las hipótesis de la teoría que se quiere aplicar.

## REFERENCIAS

- [1] *La Cronica de Hoy*(2005)*México*, mayor bebedor refrescos cola, en disputa por refresqueras. <http://www.cronica.com.mx/nota.php?idc=199175>.
- [2] Infante, G. (2008) *Métodos estadísticos*, Editorial Trillas.
- [3] Mendenhall, M. (2008) *Introducción a la probabilidad y la estadística*, Editorial Cengage Learning.
- [4] Montgomery, D. (2006) *Diseño y análisis de experimentos*, Editorial Iberoamerica.
- [5] Serrano Raúl(2009) *REFRESCOS DE COLA, ¡AGUAS!*. <http://www.saludymedicinas.com.mx/nota.asp?id=1681>.
- [6] *El Tao de la vida sana*(2009) *Los alimentos: la dieta de desintoxicación y los complementos*. Pag:146,147.

Facultad de Cs. Físico Matemáticas, BUAP.  
Rio Verde y San Claudio, Col. San Manuel,  
Puebla, Pue. C.P. 72570.  
[hreyes@fcfm.buap.mx](mailto:hreyes@fcfm.buap.mx)  
[sarus13@hotmail.com](mailto:sarus13@hotmail.com)



# LA DIVULGACIÓN DE LA MATEMÁTICA: UN ENFOQUE PERSONAL

ROBERTO TORRES HERNÁNDEZ  
UNIVERSIDAD AUTÓNOMA DE QUERÉTARO

RESUMEN. En el presente trabajo se comentan de manera general algunas consideraciones sobre el quehacer del divulgador de la matemática, a la luz de la experiencia acumulada en años de participación en eventos de este tipo. También se ilustran esas ideas con algunos ejemplos concretos.

## 1. INTRODUCCIÓN

En el presente trabajo se exponen de manera muy general, algunas ideas acerca de la divulgación de la matemática. Además, se esbozan algunos ejemplos que pretenden ilustrar estas ideas. En la sección 2, se analizan los objetivos y obstáculos a los que se puede enfrentar un divulgador. Se analiza también el público al que se enfrenta. A continuación, en las secciones siguientes, se exponen dos ejemplos de conferencias que combinan diversos aspectos de la matemática con otras áreas del conocimiento o de la misma matemática. En la sección 3, se exploran el problema isoperimétrico, la criptografía y la inducción a través de tres cuentos de reconocidos escritores de la literatura universal: Leon Tolstoi, Edgar Allan Poe y Robert Louis Stevenson, combinando de esta manera la literatura y la ciencia. En la última sección se habla del juego del tangrama, su contenido lúdico y algunos problemas interesantes desde el punto de vista matemático. Se presentan también otros rompecabezas geométricos como los poliminós.

Es menester señalar que sólo se esboza aquí el contenido de las pláticas, algunas con más detalles que otras, pero esperando que se vislumbre la manera como se lleva a cabo.

## 2. ACERCA DE LA DIVULGACIÓN

Como en casi todas las actividades humanas, es casi imposible (además de soberbio) señalar la manera como se deben de hacer las cosas. No es la intención de estas líneas pretender indicar el modo de divulgar la matemática, entre otras muchas razones por la gran diversidad de formas que hay de hacerlo. Sin embargo, a la luz de la experiencia se pueden comentar, de manera muy general, algunas ideas globales que pueden servir de guía en el diseño de nuestras actividades.

### OBJETIVOS DE LA DIVULGACIÓN

- Compartir las matemáticas, su belleza y poder, con un público amplio, demoliendo los prejuicios y barreras mucho más rígidas en matemáticas que en otras ciencias.
- Intentar cambiar la actitud hacia la matemática, pues ésta puede ser perjudicial para el desarrollo de la matemática y la cultura.

- Animar al público a ser más activo matemáticamente, para que lleve una vida más plena en el plano intelectual.
- Desarrollar la actividad matemática en libertad, no por obligación, alejándose de la odiosa sensación que niños y adultos tienen.

### OBSTÁCULOS DE LA DIVULGACIÓN

- La relación afectiva, la gran mayoría de los casos negativa, con las matemáticas a raíz de la experiencia escolar.
- La dificultad de presentar la verdadera dinámica de las Matemáticas como una ciencia viva.
- Los temas de investigación de los matemáticos, a veces de difícil vinculación con problemas importantes de la vida real.
- El lenguaje técnico propio de la matemática y el lenguaje formal y crítico de los matemáticos al hablar de sus trabajos.

### LA DIVULGACIÓN, ¿PARA QUIÉNES?

- Niños pequeños: Comunicarse con ellos adecuadamente, a través de juegos, exhibiciones, competencias.
- Para el público en general: Evitar tecnicismos, transmitir el impacto y los métodos del pensamiento matemático en algunos temas concretos.
- Para estudiantes: Hablar de la historia, evolución y aplicaciones de los temas que se les enseña. El impacto de las matemáticas en la historia de la humanidad. Tomar en cuenta el nivel escolar.
- Profesionales dentro y fuera del mundo académico: Tratar temas que aunque no usen en su trabajo, pueden arrojar luz sobre su modo de pensar y resolver problemas.
- Matemáticos investigadores y profesores universitarios: Ideas nuevas en un campo pueden ser útiles en otro. Hacer accesibles las ideas sin tecnicismos superfluos.

Parafraseando las últimas líneas del prólogo del clásico libro de Courant y Robbins *¿Qué es la matemática?* podemos decir "Tanto para el docto como para el profano, no es la especulación filosófica sino la experiencia activa en la divulgación matemática lo único que puede responder a la pregunta: ¿Qué es la divulgación de la Matemática?"

Pasemos pues a los ejemplos concretos.

## 3. LITERATURA Y MATEMÁTICAS

Existe una gran variedad de ejemplos en la literatura que tocan temas matemáticos. Pretender dar una lista completa sería poco menos que imposible. En esta sección se tocan tres, pero quedan en el tintero muchos otros clásicos, como *El libro de arena* de Borges, *Alicia en el país de las maravillas* y *Al otro lado del espejo* de Lewis Carroll, etc.

**3.1. Leon Tolstoi y el problema isoperimétrico.** El Conde Leon Nicolae-vich Tolstoi (1828-1910), nació en Iasnaia Poliana, Rusia. Hijo de una familia de la nobleza, siempre mostró simpatía por los campesinos (mujiks) y los desprotegidos. Fue un devoto cristiano y su tesis se basa en el amor a los semejantes. Él mismo cortaba su leña y hacía labores agrícolas antes de sentarse a escribir. Mantuvo

correspondencia con Lenin y Gandhi, a quienes influyó notablemente. Mostró siempre desprecio por la ciencia. Sus principales obras son *Guerra y Paz* (1864-1869), *Ana Karenina* (1873-1876) y *Resurrección* (1899). Entre 1879 y 1882 publica sus *Cuentos populares*. De estos, tomamos el cuento *Pakhom el mujik*:

La esposa del campesino Pakhom, convence a éste a que intente comprar mas tierra para obtener mejores ganancias y ascender de posición. La familia se muda a varios lugares, mejorando cada vez más.

Cuando ya tienen una situación económica muy cómoda, llega a oídos de Pakhom la existencia de una tierra lejana, muy fértil y agradable, en el pueblo de los Baschkirs, donde todo aquel que desee instalarse allí se le vende por un precio fijo (1000 rublos) toda la tierra que el comprador pueda abarcar en un día de camino, es decir, inicia a caminar al amanecer y se señala el camino por donde va, con la condición de que antes de que se oculte el sol se tiene que regresar al punto de partida.

El terreno abarcado de este modo será suyo por el costo fijado. Si no pudiera volver exactamente al punto de partida al anochecer, se perderá el dinero.

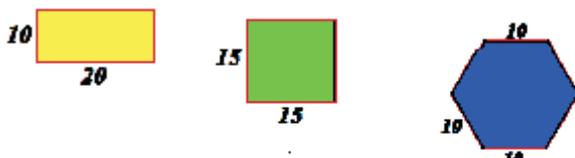
Pakhom decide probar suerte y se prepara lo mejor que puede. Planea todo con absoluta precisión: Piensa recorrer en un día alrededor de 60 versetas (medida rusa un poco mayor a un kilómetro) lo cual será el perímetro del terreno que estima en 10 000 deciatinas (medida un poco mayor a una hectárea).

Inicia el camino conforme a lo planeado, pero en el trayecto ve un bosquecillo y un arroyo que decide abarcar. Finalmente se atrasa y en el último tramo tiene que correr desesperado, lo que finalmente le provoca la muerte.

Obsevemos que como parte central de la trama, se puede plantear la pregunta: ¿Qué forma debe dar Pakhom a su terreno para abarcar la mayor cantidad de área? Esto nos lleva al Problema isoperimétrico:

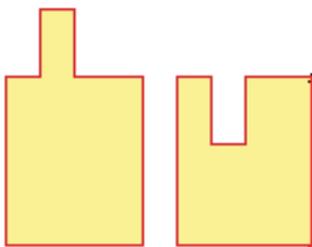
Dada una longitud fija, encontrar la figura cuyo perímetro sea esa longitud y que encierre la mayor área posible.

A continuación, se puede ilustrar el hecho de que con una longitud fija, que en este caso sería las 60 versetas recorridas en un día por Pakhom, dependiendo de la forma del terreno, se obtienen diferentes áreas:



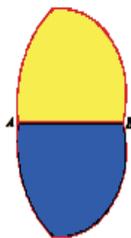
Supongamos que ya tenemos la figura. ¿Qué propiedades debe tener? Debe ser convexa, ya que si la figura no es convexa, con el mismo perímetro se puede formar una figura con mayor área:

Es importante señalar que ambas figuras no son convexas, sin embargo, el esquema puede servir para iniciar una discusión para generar una figura sin las “entradas” de una figura no convexa y con el mismo perímetro.



Una proposición que es fácil de ilustrar es la siguiente: De todos los triángulos con dos lados de longitud fija, el de mayor área es el triángulo con un ángulo recto.

Además, si A y B son dos puntos que dividen a la curva en dos arcos de igual longitud, entonces el segmento AB divide a la figura en dos partes de igual área:



Finalmente, la figura debe ser una circunferencia pues todo ángulo inscrito en una circunferencia es recto.

**3.2. Edgar Allan Poe y la criptografía.** Edgar Allan Poe (1809-1849), fue descendiente de actores, nació en Boston, aunque su infancia trascurrió en Inglaterra. Regresa a los doce años a Richmond, Virginia, donde se inicia en la literatura. De 1835 a 1847 sobrevive apenas de su trabajo literario. Es considerado como un erudito, al conocer diferentes idiomas y sus literaturas. Empieza a entregarse al alcohol. Escribe en este periodo sus *Narraciones extraordinarias* (*Los asesinatos de la calle Morgue*, *El corazón delator*, *El pozo y el péndulo*, *La caída de la casa Usher*) y su cuento en verso *El cuervo* que son cuentos de suspenso y terror. Escribe también ensayos y críticas sobre ciencia (matemáticas y criptografía), política y deporte. En un bar cercano a la estación de trenes de Baltimore muere de congestión alcohólica.

De las narraciones extraordinarias, tomamos *El escarabajo de oro*:

La historia se desarrolla en Charleston, Carolina del Sur, donde Legrand, el heredero de una rica familia de Nueva Orleáns venida a menos, se establece en la costa. En una caminata por la playa, descubre un pedazo de pergamino al intentar atrapar a un escarabajo que llama la atención por su color dorado. Al tratar de dibujar al bicho en ese pergamino aparece un texto en clave que no es fácil descifrar. Compuesto por una serie de signos en apariencia sin sentido, Legrand utiliza algunas nociones de matemáticas elementales para su comprensión.

Una vez “traducido” se observa que el mensaje contiene las instrucciones para encontrar la ubicación de un tesoro, enterrado años atrás por un famoso pirata en

una isla cercana. Después de varias peripecias por las equivocaciones de Júpiter (el sirviente negro de Legrand) e interpretaciones de las instrucciones para encontrar el tesoro, finalmente dan con el.

El texto es:

53++†305))6\*;4826)4+.)5;086\*;48†8960))85;1+(;+\*8†83(88)5\*†;46(;88\*  
76\*?;8)\*+(;485);5\*†2:\*+(;4756\*2(5\*-4)898\*;4067285);6†84++;1(+7;4808  
1;8:8+1;48†85;4)85;4)485†528806\*81(+7;48;(88;4(+?34;48)4+;161;:188;+?;

Aquí se puede platicar sobre la historia de la criptografía y los primeros encriptamientos por sustitución, es decir, mensajes en que se han cambiado las letras por signos o números. En este cuento, el protagonista de la historia procede de la siguiente manera (que el público sigue sin mucho esfuerzo):

Primero confecciona una tabla donde escribe el número de veces que aparece cada signo en el mensaje. A continuación, razona: La letra que mas se utiliza en inglés es la e (seguida de a, o, i, d, h, n, r, s, t, u, . . .) por lo que se puede sospechar que 8 en la letra e, ya que en el texto aparece 33 veces contra 26 del punto y coma y 18 del 4. Además, aparece repetida varias veces. La palabra más popular es “the”, luego, como en el texto aparecen con frecuencia los tres signos junto ;48, tenemos otras equivalencias: t es ; y h es 4. Continuando de esta manera, se descifra el mensaje completo.

Dependiendo del público, puede uno mencionar el encriptamiento utilizando matrices de codificación.

**3.3. Robert Louis Stevenson y la inducción matemática.** Robert Louis Stevenson (1850-1894) nació en Edimburgo, en el seno de una familia muy celosa en materia de religión, aunque su padre era ingeniero en faros. De salud frágil desde pequeño, inició los estudios de ingeniería civil, aunque terminó graduándose con mucho esfuerzo de abogado en 1875. Dada la severidad de la vida familiar, el joven Stevenson se rebela entregándose a una vida disipada. En 1873 se inicia como escritor y también empiezan a manifestarse los primeros síntomas de tuberculosis. Buscando climas más agradables para su salud, recorre Francia, Estados Unidos y los mares del sur. En 1883 inicia su obra *La isla del tesoro*. Otras importantes son *El extraño caso del Dr. Jekyll y Mr. Hyde* y *Cuentos de los mares del Sur*. Muere finalmente de tuberculosis.

De estos cuentos, tomamos *El diablillo de la botella*:

El hawaiano Keawe, visitando la ciudad de San Francisco conoce a una persona rica que le ofrece venderle, por una cantidad muy barata (50 dólares) una botella que concede casi cualquier deseo (excepto prolongar la vida).

En esa botella habita un diablo, que reclamará el alma del dueño de la botella que muera en la posesión de la misma. Si el propietario de la botella desea venderla, deberá de ser por un precio menor al que pagó por ella.

Keawe compra la botella, regresa a su tierra natal, donde se construye un palacio (Casa resplandeciente), vende la botella y trascurre su vida muy felizmente.

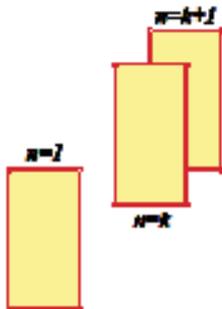
Un día se enamora de una bella y virtuosa muchacha, se conocen y planean casarse, pero él se enferma de lepra.

Desesperado, busca la botella para pedir su curación pero el precio ha bajado notablemente (5 centavos). La compra, se cura y se casa pero empieza a sentirse mal por el miedo a ir al infierno.

Su esposa que lo ama, se desespera y cree que su mal humor es porque no la quiere. Le pregunta a Keawe y él le narra toda la historia. Para salvarlo, ella contrata a un viejo, para que compre la botella en 4 centavos y ella a su vez la paga en 3. Los papeles se invierten: Keawe anda feliz y la triste es su mujer. Él se entera del sacrificio de su mujer y hace lo propio: Convence a un marino para que compre la botella en 2 centavos para luego comprarla él a 1. Pero en el último momento, el marino se arrepiente y se queda con la botella y Keawe y su esposa son felices siempre.

El método de inducción matemática puede plantearse así: Si una afirmación es cierta para el valor  $n=1$  y si cada vez que sea cierta para  $n=k$  es cierta también para  $n=k+1$ , entonces es cierta para todo número entero positivo  $n$ .

A menudo, este método se denomina, en términos coloquiales, el método del dominó, ya que si tenemos una fila infinita de dominós formados y tiramos la primera ficha, si sabemos que cada vez que se cae una, se cae la siguiente, entonces podemos estar seguros que se caerán todas.



Con esto en mente, volvamos al cuento y apliquemos nuestras matemáticas:

Afirmación: La botella no puede venderse a ningún precio  $n$ .

Nadie comprará la botella a 1 centavo. (Nadie quiere ir al infierno)

Nadie comprará la botella a 2 centavos. (Pues nadie se la compraría a 1 centavo)

Si nadie la compra a  $n=k$  centavos, entonces nadie la comprará a  $n=k+1$  centavos. (Pues no la podría vender)

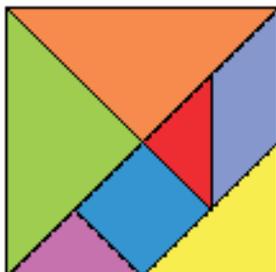
¿Qué pasó? ¿El cuento está mal planteado? Hay que tener cuidado con la matemática y el sentido común. Si alguien me ofreciera esa botella en 200 pesos, yo sí la compraría.

#### 4. EL TANGRAMA Y LAS MATEMÁTICAS

Una de las ramas más antiguas de las matemáticas recreativas es la que se ocupa de los problemas de disección. Si cortamos en varias piezas una figura plana nuestro reto consistiría en reconstruir nuevamente la figura original, o construir una nueva con las piezas dadas.

Un ejemplo de este tipo de pasatiempo son los rompecabezas chinos llamados tangramas. El tangrama clásico está formado por siete figuras geométricas llamadas

tans, los cuales se obtienen de seccionar un cuadrado de modo que se produzcan dos triángulos grandes, un triángulo mediano, dos triángulos pequeños, un cuadrado y un paralelogramo.



El juego consiste en formar figuras atendiendo las siguientes reglas:

- 1) Hay que utilizar todas las piezas.
- 2) No se permite encimar las piezas.

Los juegos con tangramas pueden agruparse en las siguientes categorías:

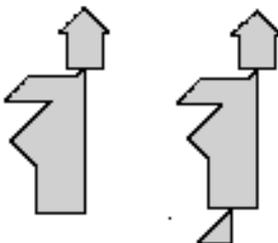
- Búsqueda de varias maneras de construir un tangrama dado, o de hallar una demostración elegante de la imposibilidad de formarlo.
- Encontrar la forma de representar artística o humorísticamente siluetas de animales, figuras humanas u otros objetos reconocibles.
- Resolver una diversidad de problemas de geometría combinatoria que los siete tans plantean.

Estos temas proporcionan una fuente muy rica de contenidos para la divulgación. Tiene además la ventaja de que pueden adecuarse al nivel de casi cualquier tipo de público.

Por ejemplo, para los más pequeños pueden ser retos atractivos formar la letra inicial de su nombre.

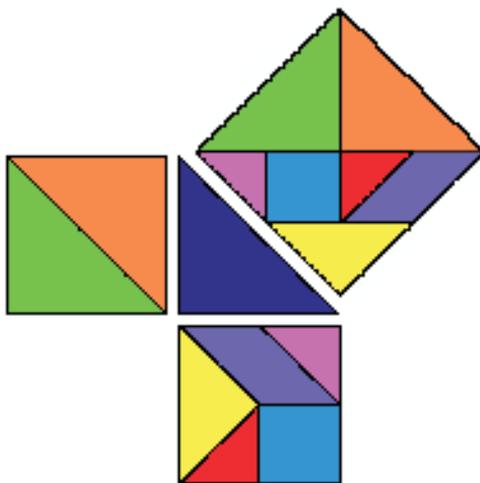
La manipulación física de las piezas nos permite formar diversas figuras, en las cuales podemos apreciar que para la misma área, las formas cambian.

Un ejemplo, se ilustra a continuación:



La figura muestra a un par de hombres, uno con un pie y el otro sin pie. Aparentemente se tiene la idea de que se ha olvidado alguna pieza al construir al primer hombre, pero conviene hacer notar que si ambas figuras han sido hechas conforme a las reglas del tangrama, tendremos entonces que los dos hombres tienen las mismas áreas.

Para el nivel secundaria, se pueden ilustrar algunos teoremas clásicos, como el Teorema de Pitágoras (al menos un caso, particular ¿por qué?):



Para el nivel preparatoria, se puede abandonar el tangrama, pero conservando la idea de rompecabezas para factorizar ecuaciones o ilustrar identidades como la ley de los cosenos, eligiendo adecuadamente las formas geométricas a manipular.

Sin embargo, aún para nivel universitario se pueden proponer problemas lo suficientemente interesantes para captar la atención de un público más selecto.

4.1. TEOREMA. Con el tangrama se pueden formar 13 figuras convexas.

La demostración consistiría en exhibir trece figuras convexas. Un buen reto para el lector es formarlas con el tangrama.

Algo mucho más difícil de demostrar es:

4.2. TEOREMA. Con el tangrama se pueden formar solamente trece figuras convexas.

Dependiendo de la preparación del público se podría intentar esbozar la prueba.

Para finalizar con estas ideas de disecciones geométricas, se presentan los poliminós (conjuntos de cuadrados unidos por sus lados) y algunos problemas relacionados a ellos:

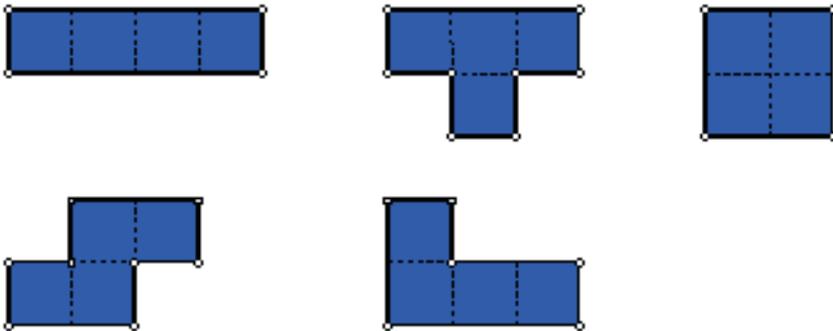
El monominó y el dominó son uno y dos cuadrados:

Y ahora los tetraminós:

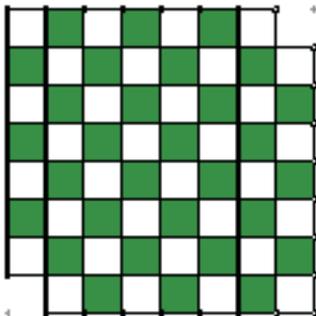
Se puede invitar al público a que proponga soluciones a los problemas planteados, inclusive se pueden obsequiar poliminós de cartón. Aquí nos contentaremos con



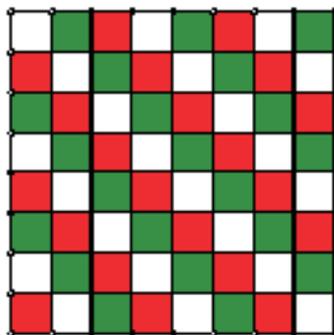
enunciarlos. Los poliminós rectos son aquellos formados por una fila de cuadrados. La generalización a pentominós y hexaminós debe ser inmediata.



4.3. PROBLEMA. ¿Se puede cubrir el siguiente tablero “recortado” con dominós que cubren dos cuadritos exactamente?



4.4. PROBLEMA. ¿Se puede cubrir el siguiente tablero con triminós rectos y un monominó?



4.5. PROBLEMA. ¿Se puede cubrir un tablero con tetraminós rectos?

4.6. PROBLEMA. Cubrir el tablero con los 12 pentaminós y un tetraminó recto. Demostrar que el tetraminó puede estar donde sea en el tablero.

4.7. PROBLEMA. Mismo problema anterior pero con el tetraminó cuadrado.

4.8. PROBLEMA. Demostrar que los hexaminós no pueden formar un rectángulo.

#### REFERENCIAS

- [1] Courant y Robbins. *¿Qué es la Matemática?* Editorial Aguilar (1979). (El Fondo de Cultura Económica ha traducido la última versión con notas y actualizaciones de Ian Stewart)
- [2] Gardner, Martin. *Viajes por el tiempo y otras perplejidades matemáticas*. RBA Editores, España, 1994.
- [3] Poe, Edgar Allan. *Narraciones Extraordinarias*. Colección “Sepan Cuantos” Núm. 210. Editorial Porrúa.

- [4] Stevenson, Robert Louis. *La isla del tesoro y cuentos de los mares del sur*. Colección "Sepan Cuantos" Núm. 110. Editorial Porrúa.
- [5] Tolstoi, Leon. *Cuentos escogidos*. Colección "Sepan Cuantos" Núm. 295. Editorial Porrúa.
- [6] Torres, Roberto y Gómez, Gonzalo. *El tangrama y las matemáticas*. Cuaderno No. 16, Taller de Apoyo Didáctico. Universidad de Guanajuato, México, 1995.

Universidad Autónoma de Querétaro  
Cerro de las campanas s/n. Col. Las Campanas  
C.P. 76010. Querétaro, Qro. México  
robert@uaq.mx



# **Sesión de Ecuaciones Diferenciales y Modelación Matemática**

Quinta Gran Semana Nacional de la Matemática



# EJEMPLOS NUMÉRICOS DE PROBLEMAS CON INESTABILIDAD NUMÉRICA Y SU REGULARIZACIÓN MEDIANTE EL MÉTODO DE TIJONOV

FÉLIX AQUINO CAMACHO  
JACOBO OLIVEROS OLIVEROS  
FCFM - BUAP

RESUMEN. El método de regularización de Tijonov (MRT) es utilizado en ecuaciones operacionales de la primera especie lineales mal planteadas para corregir el problema de la inestabilidad de la solución ante errores de los datos de entrada. El MRT consiste en penalizar el funcional de mínimos cuadrados (funcional de Tijonov) para corregir el mencionado problema de la inestabilidad. En esta penalización aparece el llamado parámetro de regularización de Tijonov el cual debe elegirse apropiadamente en función del error en los datos de entrada para garantizar que, cuando dicho error tiende a cero, la solución aproximada, converge a una solución de la ecuación para datos sin error. En este trabajo se presentan ejemplos numéricos de operadores lineales y no lineales, mostrando el MRT como solución al problema de la inestabilidad numérica.

## 1. INTRODUCCIÓN

El MRT es utilizado en ecuaciones operacionales de la primera especie mal planteadas para corregir el problema de la inestabilidad de la solución ante pequeñas variaciones de los datos de entrada, la cual puede llevar a que estas pequeñas variaciones produzcan variaciones sustanciales en la solución buscada. El MRT proporciona una herramienta con la que pueden hallarse soluciones aproximadas dependientes de un parámetro, llamado parámetro de regularización de Tijonov, el cual debe elegirse de forma conveniente, en dependencia del error cometido, para garantizar que cuando dicho error tiende a cero, la solución aproximada, converja a una solución de la ecuación para datos sin error. Para el caso en que la ecuación operacional de la primera especie sea lineal, se sabe que el MRT cumple lo anterior [7] y el parámetro puede elegirse por medio del principio de discrepancia de Morozov [7].

Para el caso en que la ecuación es no lineal, el funcional de Tijonov no es convexo, por lo que puede tener varios minimizadores locales y las rutinas clásicas de optimización pueden fallar [13]; aunque este tema esta en desarrollo, muchos investigadores han introducido métodos iterativos para la minimización del funcional de Tijonov, los cuales, reconstruyen un minimizador global del funcional [15].

La importancia de este tema en las aplicaciones es muy amplia ya que una gran cantidad de problemas inversos pueden escribirse como ecuaciones operacionales de la primera especie con operadores tanto lineales como no lineales [1], [2], [3], [5], [12], [17], [18], [19].

**1.1. Los problemas inversos.** Los problemas inversos son problemas que consisten en encontrar una propiedad desconocida de un objeto o de un medio a partir de las observaciones de una respuesta de este objeto o medio de una señal de prueba. Por lo tanto los problemas inversos proporcionan una base teórica para la detección lejana, la visualización y la evaluación no destructivas. Podemos considerar a los “problemas directos” como aquellos en los que se tiene información sobre las causas que describen un proceso en un medio y la solución del problema nos conduce a descubrir el “efecto” producido por dichas causas, mientras que en los “problemas inversos” se tiene una información parcial sobre los resultados o efectos producidos en el medio por ciertas causas que se desea descubrir a partir del análisis de dichos resultados. Así podemos hablar de los problemas directos como problemas “causa-efecto” y de los problemas inversos como problemas “efecto-causa” [12]. A continuación veremos una formulación operacional conveniente para el análisis matemático de los problemas inversos. Para ampliar el concepto de problemas inversos se sugiere la lectura de [7], [9], [10], [12] [17], [19]. En [4], [6], [8] se pueden hallar ejemplos más concretos de problemas inversos aplicados a problemas reales.

**1.2. Concepto de buen y mal planteamiento de un problema.** Hadamard clasificó a los problemas físicos que son representados por un modelo matemático, en el siguiente sentido [7]:

*Se dice que un problema es bien planteado o propiamente planteado si cumple:*

1. *Existe una solución del problema (existencia).*
2. *Existe a lo más una solución del problema (unicidad).*
3. *La solución depende continuamente de los datos de entrada (estabilidad).*

Matemáticamente esta noción de buen planteamiento se define para ecuaciones operacionales de la siguiente manera:

1.1. DEFINICIÓN. Sean  $X$  y  $Y$  espacios normados,  $A : X \longrightarrow Y$  un operador continuo (lineal o no lineal), la ecuación operacional

$$(1) \quad Ax = y,$$

es llamada bien planteada si cumple:

1. Existencia: Para cada  $y \in Y$  existe (al menos uno)  $x \in X$  tal que se satisface(1).
2. Unicidad: Para cada  $y \in Y$  existe a lo más un  $x \in X$  tal que se satisface(1).
3. Estabilidad: La solución  $x$  depende continuamente de  $y$ , es decir, para toda sucesión  $(x_n) \subset X$  con  $Ax_n \rightarrow Ax \quad (n \rightarrow \infty)$ , se sigue que  $x_n \rightarrow x \quad (n \rightarrow \infty)$ .

*Las ecuaciones que no cumplan alguna de estas propiedades son llamadas mal planteadas.*

Las propiedades 1 y 2 pueden resumirse diciendo que el operador  $A$  es biyectivo y la propiedad 3 diciendo que es bicontinuo.

Del hecho que generalmente observamos un único resultado al actuar sobre un medio, se supone que los modelos asociados a problemas directos deben satisfacer la propiedad de buen planteamiento. Los problemas mal planteados son importantes en muchas aplicaciones en las que podemos reducir un problema físico a la ecuación

(1), donde el inverso del operador  $A$  es no acotado, por ejemplo ver [1]

1.2. EJEMPLO. Consideramos la matriz  $\bar{A}$  dada por

$$\bar{A} = \begin{pmatrix} 1 & 2 \\ 1 + \sigma & 2 \end{pmatrix},$$

donde  $\sigma \neq 0$  y cuyo determinante es igual a  $-2\sigma$ , que puede ser muy pequeño si lo es  $\sigma$ .

Sea  $A : X \rightarrow Y$  definido por  $A(x) = \bar{A}x$  donde  $X = Y = \mathbf{R}^2$ . Ya que los operadores actuando sobre espacios normados de dimensión finita son continuos y que la inyectividad es equivalente a la sobreyectividad, entonces la ecuación  $Ax = y$  es bien planteada. Sin embargo si  $\sigma$  es muy pequeño dicha ecuación puede presentar problemas de inestabilidad numérica. Esto se mide con el número de condición de  $A$  [21]. A continuación se ilustra este hecho. Sea

$$y = \begin{pmatrix} 3 \\ 3 + \sigma \end{pmatrix}.$$

Entonces la solución del sistema

$$(2) \quad \bar{A}x = y,$$

está dada por  $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ .

Supóngase que en vez de conocer a  $y$ , se conoce una perturbación  $y^\delta$  donde

$$(3) \quad y^\delta = \begin{pmatrix} 3 \\ 3 + \sigma \end{pmatrix} + \begin{pmatrix} 0 \\ \delta \end{pmatrix}, \quad \|y - y^\delta\|_Y < \delta, \quad \delta > 0.$$

En este caso la solución está dada por

$$(4) \quad x^\delta = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} \frac{\delta}{\sigma} \\ -\frac{\delta}{2\sigma} \end{pmatrix}.$$

Observe que no necesariamente  $\|x - x^\delta\|_X$  es un valor pequeño; en un caso más específico, por ejemplo si  $\sigma = 0.001$  y  $\delta = 0.01$  la solución “aproximada” del sistema será

$$(5) \quad x^\delta = \begin{pmatrix} 11 \\ -4 \end{pmatrix}, \quad \|x - x^\delta\| = 11.18033.$$

Es importante aclarar que el operador  $A$  es bien planteado, lo que este ejemplo ilustra es la inestabilidad numérica ocasionada por el mal condicionamiento de la matriz.

1.3. **Método de los mínimos cuadrados.** Supongamos que la ecuación  $Ax = y$  es bien planteada, donde  $A$  es un operador lineal entre espacios separables de Hilbert  $X$  y  $Y$ , y además solo conocemos una aproximación  $y^\delta$  del dato exacto  $y$ , donde  $\delta > 0$  es conocido y

$$(6) \quad \|y - y^\delta\|_Y < \delta.$$

Se espera encontrar una solución aproximada  $x^\delta \in X$  a la solución exacta  $x \in X$ , para lo cual podemos tomar como posible solución aproximada, la solución en mínimos cuadrados:

1.3. DEFINICIÓN.  $\hat{x}$  es una solución en mínimos cuadrados de la ecuación  $Ax = y$  si cumple que

$$(7) \quad \hat{x} = \arg \min_{x \in D(A)} \|Ax - y\|^2.$$

El operador  $A$  es lineal y acotado entre espacios separables de Hilbert. Observe que es posible que  $\hat{x}$  pueda ser no única, por lo que en este caso tendremos un conjunto de soluciones en el sentido de mínimos cuadrados por lo que es necesario definir la solución mejor aproximada o también conocida como solución en mínimos cuadrados de norma mínima.

Observemos que encontrar la solución en mínimos cuadrados se reduce a encontrar el mínimo del funcional:

$$(8) \quad J(x) = \|Ax - y\|^2.$$

Supongamos que  $A : X \rightarrow Y$  lineal y continuo, con  $X, Y$  espacios reales separables de Hilbert por lo que este funcional es continuamente diferenciable y convexo, por lo tanto es suficiente que  $J'(x) = 0$  para que  $J(x) = \inf \{J(z) : z \in X\}$ . Entonces busquemos a  $J'(x)$ :

$$\begin{aligned} J(x+h) - J(x) &= \|A(x+h) - y\|^2 - \|Ax - y\|^2 \\ &= \langle A(x+h) - y, A(x+h) - y \rangle_Y - \langle Ax - y, Ax - y \rangle_Y \\ &= \langle Ax - y, Ah \rangle_Y + \langle Ah, Ax - y \rangle_Y + \langle Ax - y, Ax - y \rangle_Y \\ &\quad + \langle Ah, Ah \rangle_Y - \langle Ax - y, Ax - y \rangle_Y \\ &= \langle Ax - y, Ah \rangle_Y + \langle Ah, Ax - y \rangle_Y + \langle Ah, Ah \rangle_Y \\ &= 2 \langle Ax - y, Ah \rangle_Y + \langle Ah, Ah \rangle_Y \quad (\text{por ser un esp. de Hilbert real}) \\ &= 2 \langle A^*[Ax - y], h \rangle_X + \langle Ah, Ah \rangle_Y \\ &= \langle 2(A^*Ax - A^*y), h \rangle_X + \langle Ah, Ah \rangle_Y \end{aligned}$$

y como  $A$  es lineal y acotado existe  $\|A\|$  [12] y se cumple por la desigualdad de Cauchy-Schwarz que

$$(9) \quad \langle Ah, Ah \rangle_Y = \|Ah\|_Y^2 \leq \|A\|^2 \|h\|_X^2,$$

De aquí

$$(10) \quad \frac{\langle Ah, Ah \rangle_Y}{\|h\|_X} \leq \|A\|^2 \|h\|_X \rightarrow 0,$$

cuando  $\|h\|_X \rightarrow 0$ , se comporta como una  $0(h) \rightarrow 0$ , cuando  $\|h\|_X \rightarrow 0$ , se tiene que:

$$(11) \quad J'(x) = 2(A^*Ax - A^*y),$$

y, por lo tanto, la condición necesaria de primer orden de extremo para  $J$  es:

$$(12) \quad A^*Ax = A^*y.$$

Observemos la siguiente interpretación: resulta que resolver (12) es mejor que resolver  $Ax = y$  ya que el operador compacto  $A^*$  suaviza los datos  $y$  al aplicárselos, o sea,  $A^*y$  es una suavización a  $y$  y esto es una ventaja sobre la resolución de  $Ax = y$ .

Entonces en nuestro problema cuando tenemos una aproximación  $y^\delta$  del dato exacto  $y$ , donde  $\delta > 0$  es conocido, nuestra solución aproximada será  $x^\delta$  que satisface las **ecuaciones normales**

$$(13) \quad A^*Ax^\delta = A^*y^\delta.$$

Es importante mencionar que si  $x^\delta$  es solución de las ecuaciones normales, es precisamente la solución en mínimos cuadrados de norma mínima, lo cual está fuertemente relacionado con la idea de inversas generalizadas [22].

Ahora, ¿qué pasa si la ecuación operacional  $Ax = y$  es mal planteada? Pues resulta que las ecuaciones normales (13) son también una ecuación operacional mal planteada. En el caso de que  $A$  sea una matriz mal condicionada se sabe que el número de condición de  $A^*A$  es el cuadrado del número de condición de  $A$ . De acuerdo a las ideas del método de solución por mínimos cuadrados, Tijonov planteó que la solución pudiese ser

$$(14) \quad x^\alpha = \arg \min_{x \in D(A)} \{\|Ax - y^\delta\| + \alpha \|x\|^2\},$$

donde  $\alpha > 0$  y  $\alpha \|x\|^2$  es un funcional continuo y lineal, que “penaliza” el defecto. Se puede verificar que este  $x^\alpha$  cumple las **ecuaciones normales modificadas** [20]

$$(15) \quad A^*Ax^\alpha + \alpha Ix^\alpha = A^*y^\delta,$$

además se puede elegir  $\alpha = \alpha(\delta)$  tal que  $\alpha(\delta) \rightarrow 0$  cuando  $\delta \rightarrow 0$  y  $x^\alpha$  converja a la solución exacta cuando  $\delta \rightarrow 0$ .

## 2. REVISIÓN DE RESULTADOS DE LA TEORÍA CLÁSICA DEL MÉTODO DE TIJONOV PARA EL CASO LINEAL.

El método de regularización de Tijonov es ampliamente utilizado para resolver ecuaciones operacionales de la primera especie  $Ax = y$ , donde  $A$  es un operador lineal compacto, entre espacios de Hilbert  $X$  e  $Y$ . Este método fué desarrollado por Tijonov en dos artículos presentados en 1963 y 1965, pero fué establecido en forma más general por Arsenin y el mismo Tijonov en 1977 en su libro “Solution of ill-posed problems” [19].

**2.1. El funcional de Tijonov y sus propiedades para operadores lineales compactos.** Recordemos que de acuerdo a las ideas del método de solución por mínimos cuadrados, Tijonov planteó que la solución pudiese ser

$$(16) \quad x^\alpha = \arg \min_{x \in D(A)} \{ \|Ax - y\| + \alpha \|x\|^2 \},$$

donde  $\alpha > 0$  y  $\alpha \|x\|^2$  es un funcional que “penaliza” el defecto. Verifiquemos que este  $x^\alpha$  cumple las ecuaciones normales modificadas

$$(17) \quad A^*Ax^\alpha + \alpha Ix^\alpha = A^*y,$$

donde  $I$  es el operador identidad actuando sobre  $X$ , para esto es necesaria la siguiente:

**2.1. DEFINICIÓN.** Dado un operador lineal y acotado  $A : X \longrightarrow Y$ , entre espacios de Hilbert, se define “el funcional de Tijonov” como:

$$(18) \quad J_\alpha(x) = \|Ax - y\|^2 + \alpha \|x\|^2,$$

para  $x \in X$ , donde  $\alpha > 0$  y  $\alpha \|x\|^2$  es un funcional continuo y lineal.

Ahora es necesario probar la existencia y la unicidad del mínimo del funcional de Tijonov.

**2.2. TEOREMA.** Sea  $A : X \longrightarrow Y$  un operador lineal compacto entre espacios separables de Hilbert, con  $\alpha > 0$ . Entonces el funcional de Tijonov tiene un único mínimo  $x^\alpha \in X$ , que depende de manera continua de  $y$ , para cualquier  $y \in Y$  fijo. Este mínimo  $x^\alpha$  es la única solución de las ecuaciones normales modificadas

$$A^*Ax^\alpha + \alpha Ix^\alpha = A^*y.$$

La solución  $x^\alpha$  puede ser escrita en la forma  $x^\alpha = R_\alpha y$  con

$$(19) \quad R_\alpha := (\alpha I + A^*A)^{-1}A^* : Y \longrightarrow X.$$

La demostración de este teorema y de los siguientes se pueden hallar en [21].

**2.2. El teorema de Tijonov para operadores lineales compactos.** Ahora, supongamos que para el lado derecho exacto  $\bar{y}$ , la ecuación

$$(20) \quad Ax = y,$$

tiene solución única  $\bar{x}$ . Pero en lugar de  $\bar{y}$  conocemos el lado derecho aproximado  $y^\delta$ ,  $\delta > 0$  y  $\|y^\delta - \bar{y}\| \leq \delta$ . Utilizando  $y^\delta$  y  $\delta$  necesitamos construir la solución aproximada de (20)  $x^\delta$  que converja a  $\bar{x}$  cuando  $\delta \rightarrow 0$ .

Se puede mostrar que bajo ciertas condiciones los elementos que proporcionan un ínfimo para el funcional de Tijonov (18) pueden ser considerados como soluciones de (20) con el lado derecho dado aproximadamente. Denotemos por  $x^{\alpha(\delta)}$  el elemento que da un ínfimo para el funcional:

$$J_{\alpha(\delta)}(x) = \|Ax - y^\delta\| + \alpha(\delta) \|x\|^2,$$

donde  $\alpha(\delta) > 0$  para  $\delta > 0$ . Del teorema 2.2 se tiene que  $x^{\alpha(\delta)}$  existe y es único.

**2.3. TEOREMA (de Tijonov).** Si  $A$  es compacto, lineal e inyectivo entre dos espacios de Hilbert  $X, Y$  separables. Si la ecuación (20) tiene solución exacta  $\bar{x}$  para datos exactos  $\bar{y}$ , entonces el ínfimo  $x^{\alpha(\delta)}$  del funcional de Tijonov correspondiente a la ecuación  $Ax = y^\delta$  donde  $\|y^\delta - \bar{y}\| \leq \delta$  cumple que si se toma  $\alpha(\delta) > 0$ ,  $\delta > 0$  con  $\alpha(\delta) \rightarrow 0$ , y  $\frac{\delta^2}{\alpha(\delta)} \rightarrow 0$  cuando  $\delta \rightarrow 0$ , entonces  $\|x^{\alpha(\delta)} - \bar{x}\| \rightarrow 0$  cuando  $\delta \rightarrow 0$ .

**2.4. EJEMPLO.** Retomemos el ejemplo 1.2, al aplicar el método de regularización de Tijonov, obtenemos que la solución aproximada es

$$x_\alpha^\delta = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} \frac{-\alpha^2 - 4\alpha + 2\alpha\sigma + \alpha\delta + \alpha\delta\sigma + 4\delta\sigma}{\alpha^2 + 10\alpha + 2\alpha\sigma + \alpha\sigma^2 + 4\sigma^2} \\ \frac{-\alpha^2 + 2\alpha + 2\alpha\delta - 2\delta\sigma - \alpha\sigma^2}{\alpha^2 + 10\alpha + 2\alpha\sigma + \alpha\sigma^2 + 4\sigma^2} \end{pmatrix} = x + x_e^{\alpha,\delta}.$$

Tomando los valores  $\sigma = 10^{-3}$ ,  $\delta = 10^{-2}$ ,  $\alpha = 10^{-5}$ , encontramos la siguiente aproximación

$$x_\alpha^\delta = \begin{pmatrix} 1.00009998 \\ 1.00199855 \end{pmatrix},$$

donde se obtiene un error menor. Es decir

$$\|x - x_\alpha^\delta\| = 0.00223472.$$

Es importante mencionar que el parámetro de regularización fue escogido a priori (en este caso a prueba y error) ya que es sencillo mostrar que la regularización esta siendo adecuada, pero en la mayoría de los casos es necesario fundamentar con un proceso formal la elección de dicho parámetro. En este ejemplo también es posible observar el comportamiento gráfico de los funcionales, el de mínimos cuadrados y el de Tijonov, respecto a como existen cambios al implementar la penalización, y como esto ayuda a encontrar el elemento que minimiza (Figura 1).

### 3. EL MÉTODO DE REGULARIZACIÓN DE TIJONOV PARA EL CASO NO LINEAL.

**3.1. Introducción.** Hemos visto en la sección anterior que la teoría para el caso lineal de problemas mal-planteados está bien desarrollada y puede ser considerada como completa. Ahora queremos revisar el caso no lineal, donde la teoría no está tan desarrollada como en el caso lineal. Se pretende resolver:

$$(21) \quad F(x) = y,$$

donde  $F : D(F) \subset X \rightarrow Y$  es un operador no lineal entre los espacios de Hilbert  $X$  y  $Y$ . Nosotros entenderemos como mal planteamiento de un problema no lineal la inestabilidad, es decir, cuando la solución no depende continuamente de los datos de entrada. Durante este capítulo se asume que:

- (i)  $F$  es continuo.
- (ii)  $F$  es débilmente (secuencialmente) cerrado, es decir para cualquier sucesión  $\{x_n\} \subset D(F)$ ,  $x_n$  converge débilmente a  $x \in X$  y  $F(x_n)$  converge débilmente a

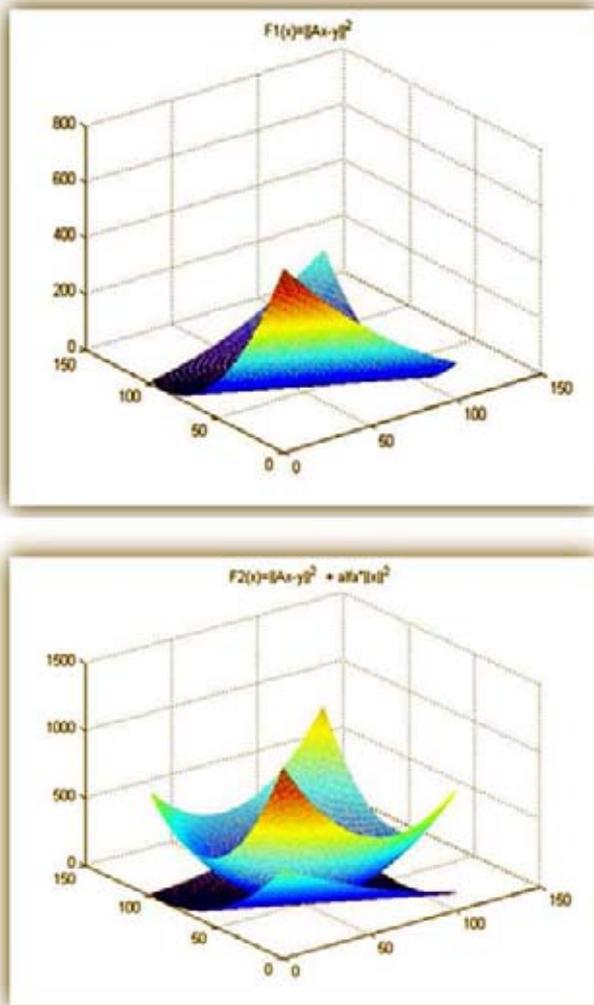


FIGURA 1. Gráfica del funcional de mínimos cuadrados, observe como gráficamente es difícil hallar un mínimo (arriba). Gráfica del funcional de Tijonov sobrepuesta a la gráfica del funcional de mínimos cuadrados, observe como mejoran las condiciones para encontrar un mínimo (abajo).

$y \in Y$  implica que  $x \in D(F)$  y  $F(x) = y$ .

Ambas suposiciones se satisfacen si  $F$  es continuo, compacto y si  $D(F)$  es débilmente cerrado (respecto a la topología débil), por ejemplo si es cerrado y convexo.

De manera más general que en el caso lineal, donde se ocupó la noción de solución de norma mínima, para el problema no lineal (21), se utiliza el concepto de una  $x^*$ -solución de norma mínima  $x^\dagger$ , es decir:

$$F(x^\dagger) = y$$

y

$$\|x^\dagger - x^*\| = \min \{\|x - x^*\| \quad : \quad F(x) = y\}.$$

La razón por el cual tratamos el caso general  $x^*$  es que para problemas no lineales,  $x^* = 0$  no juega un papel especial. Por el contrario, la elección de  $x^*$  será muy crucial, especialmente para resultados locales sobre la convergencia.

Es claro que la información a-priori disponible acerca de las soluciones de  $F(x) = y$  se debe considerar para la selección de  $x^*$ .

En el caso de soluciones múltiples,  $x^*$  juega el papel de un criterio de selección.

Como para los problemas lineales, una  $x^*$ -solución de norma mínima no necesariamente existe, y aunque exista, no necesariamente es única, ya que  $F$  es no lineal.

En lo que sigue se asumirá la existencia de una  $x^*$ -solución de norma mínima  $x^\dagger$  para el dato de entrada  $y \in Y$ , lo cual se puede conseguir, suponiendo que la ecuación (21) tiene una solución exacta y que  $F$  sea débilmente cerrado. El caso en que (21) tiene solución únicamente en el sentido de mínimos cuadrados es un problema más complicado [17].

Como lo es para los problemas lineales, se debe encontrar un criterio para decidir si un problema no lineal es mal planteado o no.

Si el operador no lineal  $F$  es compacto, se puede dar una condición suficiente para el mal planteamiento de (21), similar a la contraparte lineal compacta. La inyectividad del operador lineal es una condición suficiente para que  $X$  tenga dimensión infinita. Esta condición puede llevarse al caso no lineal, en el sentido que una inyectividad local alrededor de  $x^\dagger$  de un operador compacto no lineal  $F$ , implica el mal planteamiento del problema (21), a condición de que  $D(F)$  es de dimensión infinita alrededor de  $x^\dagger$ .

**3.1. TEOREMA.** Sea  $F$  un operador no lineal compacto y continuo, y sea  $D(F)$  débilmente cerrado. Más aún, supongamos que  $F(x^\dagger) = y$ , y existe un  $\varepsilon > 0$  tal que  $F(x) = \bar{y}$  tiene una única solución para toda  $\bar{y} \in R(F) \cap B_\varepsilon(y)$ . Si existe una sucesión  $\{x_n\} \subset D(F)$  que satisfaga

$$(22) \quad x_n \rightarrow x^\dagger \text{ pero } x_n \not\rightarrow x^\dagger,$$

entonces  $F^{-1}$  (definida sobre  $R(F) \cap B_\varepsilon(y)$ ) es no continua en  $y$ .

**3.2. Análisis de convergencia.** Como en el caso lineal, se reemplaza el problema  $F(x) = y$  por el problema de minimizar:

$$(23) \quad \|F(x) - y^\delta\|^2 + \alpha \|x - x^*\|^2 \longrightarrow \min, \quad x \in D(F),$$

donde  $\alpha > 0$ ,  $y^\delta \in Y$  es una aproximación del lado derecho exacto  $y$  del problema (21) y  $x^* \in X$ . Bajo la suposición que el problema (23) admite una solución, es posible que la solución en general no sea única ya que  $F$  es no lineal. Como en el caso lineal, cualquier solución de (23) será denotada por  $x_\alpha^\delta$ . Primero se probará que el problema de resolver (23) es estable en el sentido de la dependencia continua de la solución de los datos  $y^\delta$ .

**3.2. TEOREMA.** Sea  $\alpha > 0$  y sean  $\{y_k\}$  y  $\{x_k\}$  sucesiones donde  $y_k \rightarrow y^\delta$  y  $x_k$  es un minimizador de

$$\|F(x) - y_k\|^2 + \alpha \|x - x^*\|^2.$$

Entonces existe una subsucesión convergente de  $\{x_k\}$  y el límite de toda subsucesión convergente es un minimizador de (23).

En el siguiente teorema se muestra que bajo las mismas condiciones sobre  $\alpha(\delta)$  como en el caso lineal, las soluciones de (23) convergen hacia una  $x^*$ -solución de norma mínima de (21).

Las demostraciones de los teoremas de estas notas se pueden encontrar en [21].

**3.3. TEOREMA.** Sea  $y^\delta \in Y$  con  $\|y - y^\delta\| \leq \delta$  y sea  $\alpha(\delta)$  tal que  $\alpha(\delta) \rightarrow 0$  y  $\frac{\delta^2}{\alpha(\delta)} \rightarrow 0$  como  $\delta \rightarrow 0$ . Entonces toda sucesión  $\{x_{\alpha_k}^{\delta_k}\}$ , donde  $\delta_k \rightarrow 0$ ,  $\alpha_k := \alpha(\delta_k)$  y  $x_{\alpha_k}^{\delta_k}$  es una solución de minimizar

$$\|F(x) - y^{\delta_k}\|^2 + \alpha_k \|x - x^*\|^2; \quad x \in D(F),$$

tiene una subsucesión convergente.

El límite de toda subsucesión convergente es una  $x^*$ -solución de norma mínima. Además si la  $x^*$ -solución de norma mínima  $x^\dagger$  es única, entonces

$$\lim_{\delta \rightarrow 0} x_{\alpha(\delta)}^\delta = x^\dagger.$$

Respecto a la teoría del método de regularización de Tijonov se sugiere en especial [16]; recientemente se publicó un artículo donde se revisan las tendencias actuales de la investigación y modificaciones sobre este método [18]; También es importante revisar métodos iterativos que están basados en este método [14], así como algunas modificaciones del funcional de Tijonov como generalizaciones o como mejoras del método [13], [15].

#### 4. EJEMPLOS NUMÉRICOS PARA EL CASO NO LINEAL.

En esta sección se presenta un ejemplo numérico en el cual se muestra, por una parte la inestabilidad de la ecuación operacional y por otra parte su regularización

mediante el método de Tijonov.

Consideremos la función

$$(24) \quad F(x_1, x_2) = a(x_1^4 + x_2^4),$$

donde  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ , y  $a$  es un real positivo pequeño, por ejemplo  $a = 0.00001$ .

El problema que se pretende resolver es el de encontrar un mínimo de la función mediante el método de Newton, es decir

$$(25) \quad X_{k+1} = X_k - H(X_k)^{-1} \nabla F(X_k),$$

donde  $X_k = (x_{k1}, x_{k2}) \in \mathbb{R}^2$ ; al realizar el proceso iterativo se encuentra el mínimo sin problema (figura 2). Ahora supongamos que cometemos un pequeño error en el gradiente en (25), entonces la iteración con error es

$$(26) \quad \tilde{X}_{k+1} = X_k - H(X_k)^{-1} \tilde{\nabla} F(X_k),$$

donde al realizar el proceso iterativo ya no es posible hallar el mínimo, ya que el proceso oscila (figura 2), aquí se ilustra la iteración hasta el paso 51. El comportamiento de los valores en las entradas lleva a sospechar que el método no converge, además es claro que ahora es afectada la velocidad de convergencia; es importante comentar que al experimentar con otro tipo de errores en (26) se han obtenido resultados similares.

¿Dónde está el problema de la inestabilidad? Se observa que  $H(X_k)^{-1}$  magnifica el error en cada iteración:

$$(27) \quad H(X_k)^{-1} = \frac{1}{12ax_{k1}^2 x_{k2}^2} \begin{pmatrix} x_{k2} & 0 \\ 0 & x_{k1} \end{pmatrix},$$

$$(28) \quad \tilde{X}_{k+1} = X_k - H(X_k)^{-1} \tilde{\nabla} F(X_k).$$

Nótese que  $a$  es pequeña y estando en el denominador magnifica el error, por lo que se regulariza en cada paso de la iteración bajo el siguiente esquema.

Observe que:

$$(29) \quad \tilde{X}_{k+1} = X_k - H(X_k)^{-1} \tilde{\nabla} F(X_k),$$

$$(30) \quad \tilde{X}_{k+1} = X_k - \tilde{d}(X_k),$$

donde la dirección de descenso es

$$(31) \quad \tilde{d}(X_k) = H(X_k)^{-1} \tilde{\nabla} F(X_k),$$

de aquí

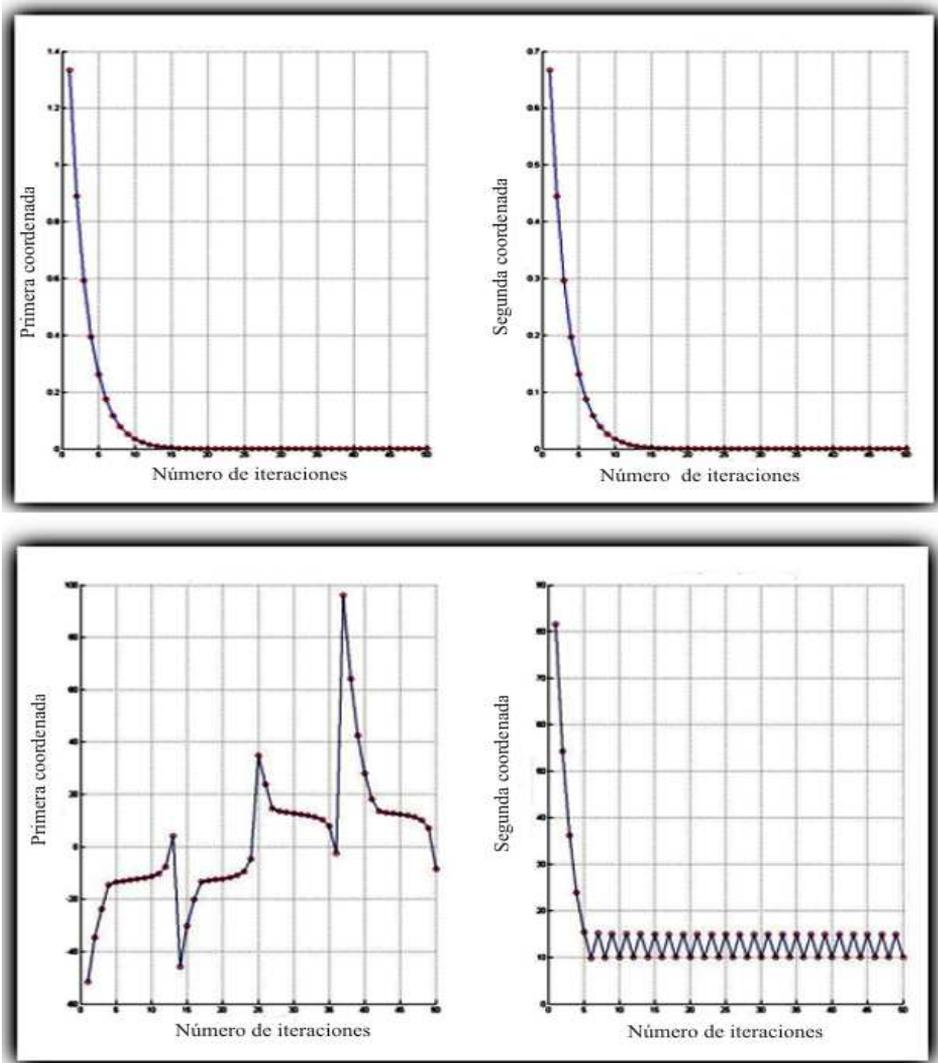


FIGURA 2. Las gráficas muestran el proceso iterativo por coordenada en la función exacta (arriba) y para el proceso con error en el gradiente (abajo).

$$(32) \quad H(X_k)\tilde{d}(X_k) = \tilde{\nabla}F(X_k).$$

Observe que (32) es en realidad una ecuación operacional la cual vamos a regularizar, teniendo así una dirección de descenso regularizada

$$(33) \quad d^{\alpha_k}(X_k) = (H^t H + \alpha_k I)^{-1}(X_k)H^t(X_k)\tilde{\nabla}F(X_k),$$

donde  $\alpha$  es obtenida mediante el principio de discrepancia de Morozov, es decir

$$(34) \quad \left\| H(d^{\alpha_k}(X_k)) - \tilde{\nabla}F(X_k) \right\|^2 = \delta_k^2.$$

Sustituyendo (33) en (34), en realidad se resuelve

$$(35) \quad \left\| H((H^t H + \alpha_k I)^{-1}(X_k)H^t(X_k)\tilde{\nabla}F(X_k)) - \tilde{\nabla}F(X_k) \right\|^2 - \delta_k^2 = 0,$$

Observe que  $\alpha_k$ , depende de quién sea  $\delta_k$  y  $X_k$ . Por lo tanto nuestra iteración regularizada es

$$(36) \quad X_{k+1}^{\alpha_k} = X_k^{\alpha_{k-1}} - (H^t H + \alpha_k I)^{-1}(X_k^{\alpha_{k-1}})H^t(X_k^{\alpha_{k-1}})\tilde{\nabla}F(X_k^{\alpha_{k-1}}).$$

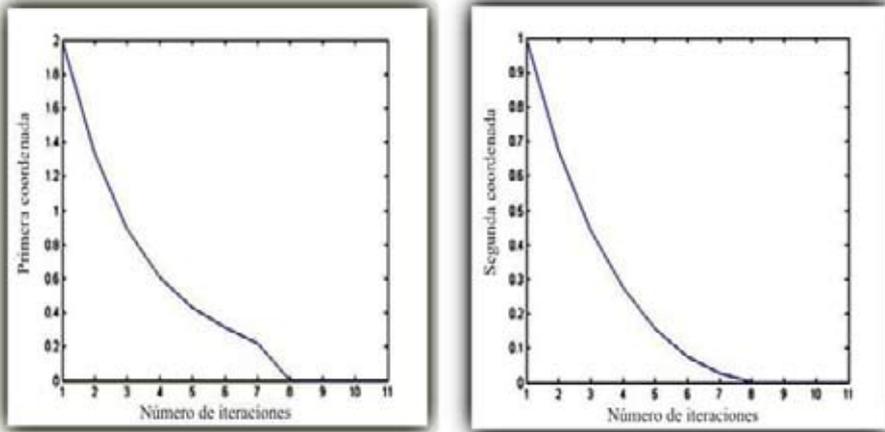


FIGURA 3. Descenso por coordenada del proceso iterativo regularizado.

La gráfica 3 muestra el descenso de las coordenadas, lo que es muestra de que se obtuvieron buenos resultados, partiendo la iteración del punto

$$X_0 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

En este ejemplo se mostro como al regularizar una ecuación mal planteada se corrige la inestabilidad creada por errores en los datos de entrada.

Es importante mencionar que en la bibliografía existen otros metodos iterativos que se construyen para que tengan propiedades regularizantes tal como el método del máximo descenso [11]. Existen otros métodos como los cuasi-Newton que son más robustos que el método de N ewton y que pueden aplicarse al problema presentado aqu ı.

## 5. CONCLUSIONES

El método de regularización de Tijonov es una herramienta importante para solucionar la inestabilidad de ecuaciones operacionales mal planteadas, sean estas lineales o no lineales.

La investigación sobre el MRT ha dado solución a problemas particulares creando modificaciones o generalizaciones del método original, por lo cual es importante revisar las ideas originales del método.

Este trabajo se realizó una revisión de los diferentes resultados que se presentan en libros especializados en métodos de regularización tanto para operadores lineales como no lineales, así como la recopilación de artículos de reciente publicación sobre la investigación en base al método de Tijonov.

Se construyeron ejemplos que dan mayor claridad a los conceptos y resultados de la teoría de Tijonov para ambos casos, es decir, para operadores lineales y no lineales, apoyados de gráficos, esquemas y programas propios que se encuentran incluidos en el trabajo.

Es importante mencionar que en este trabajo no se está realizando una implementación numérica, solo se está ilustrando el MRT con los ejemplos presentados. Tampoco se abordó la teoría sobre la velocidad de convergencia del método, lo cual se puede encontrar en la bibliografía mencionada.

## REFERENCIAS

- [1] Fraguela A., Morín M. y Oliveros J. *Inverse Electroencephalography for volumetric Sources*. Mathematics and Computers in Simulation, 78(4)(2008): 481-492.
- [2] Fraguela A., Oliveros J., Morín M. y Cervantes L. *Inverse Electro-encephalography for cortical Sources*. Applied Numerical Mathematics 55 (2005): 191-203.
- [3] Oliveros J., Cortés S., Morín M. y Fraguela A. *Problema de Cauchy para la ecuación de Laplace en una región anular bidimensional*. Memorias de la XVII Escuela Nacional de Optimización y Análisis Numérico. Gómez Palacio, Durango, del 22 al 27 de abril de 2007.
- [4] Morín M. *Análisis del problema inverso de identificación de fuentes a través de planteamientos operacionales*. Tesis de doctorado en Ciencias Matemáticas, Postgrado en Matemáticas de la Facultad de Ciencias Físico Matemáticas de la Benemérita Universidad Autónoma de Puebla (2005).
- [5] Fraguela A., Oliveros J., Cervantes L., Morín M. y Gómez S. *Un algoritmo no iterativo para la tomografía de capacitancia eléctrica*. Revista Mexicana de Física, 51 (3)(2005): 236-242.
- [6] Oliveros J. *Modelación Matemática de la actividad eléctrica de la corteza cerebral para usar la Electroencefalografía como un método de Tomografía*. Tesis de Doctorado en Ciencias Matemáticas, Postgrado en Matemáticas de la Facultad de Ciencias Físico Matemáticas de la Benemérita Universidad Autónoma de Puebla (2001).
- [7] Kirsch A. *An introduction to the mathematical theory of inverse problems*. Springer Verlag (1996).
- [8] Chang J., Yeih W. y Shieh M. *On the modified Tikhonov's regularization method for the Cauchy problem for the Laplace equation*. Journal of marine and Science Technology, 9(2)(2001): 113-121.
- [9] Engl H.W. *Inverse problems*. Sociedad Matemática Mexicana - UNAM. México (1995).
- [10] Vasin V.V. y Ageev A.L. *Ill-posed problems with a priori information*. First edition. Utrecht, The Netherlands (1995).
- [11] Gilyazov S.F. y Gol'dman N.L. *Regularization of ill-posed problems by iteration methods*. Kluwer Academic Publishers (2000).
- [12] Fraguela A., 2007. *Notas para un curso de problemas inversos*. Manuscrito en preparación. FCFM-BUAP.
- [13] Ulbrich M. *A generalized Tikhonov regularization for nonlinear inverse ill-posed problems*. Technische Universität München 1998.

- [14] Hanke M. y Groetsch C.W. *Nonstationary iterated Tikhonov regularization*. J. Optim. Theory and Applications, 98 (1998): 37-53.
- [15] Ramlau R. y Teschke G. *Tikhonov replacement functionals for iteratively solving nonlinear operator equations*. University of Bremen, Germany (2005).
- [16] Engl H.W., Hanke M. y Neubauer A. *Regularization of inverse problems*. Kluwer Academic Publishers. Holanda (1996).
- [17] Binder A., Engl H.W., Groetsch C.W., Neubauer A. y Scherzer O. *Weakly closed nonlinear operators and parameter identification in parabolic equations by Tikhonov regularization*. Appl. Anal 55(1994): 215-234.
- [18] Tikhonov A.N., Leonov A.S. y Yagola A.G. *Nonlinear ill-posed problems*. Edición en inglés. Chapman & Hall. Gran Bretaña (1998).
- [19] Vasin V.V. *Some tendencies in the Tikhonov regularization of ill-posed problems*. J. Inv. Ill-posed problems, 14(8)(2006): 813-840.
- [20] Denisov A.M. *Elements of the theory of inverse problems*. VSP, Utrecht, Holanda (1999).
- [21] Aquino C.F. *Revisión de resultados sobre el método de regularización de Tikhonov*. Tesis de Maestría en Ciencias Matemáticas, Postgrado en Matemáticas de la Facultad de Ciencias Físico Matemáticas de la Benemérita Universidad Autónoma de Puebla (2009).
- [22] Groetsch C.W. *Generalized inverses of linear operators*. Marcel Dekker. Estados Unidos de America (1977).

Facultad de Cs. Físico Matemáticas, BUAP.  
Av. San Claudio y 18 sur,  
72570 Puebla, Pue.  
fagus\_7@hotmail.com  
oliveros@fcfm.buap.mx



# PRUEBAS MAXIMIN PARA LA CALIDAD DE ESTABILIZACIÓN ROBUSTA APLICADAS A UN ROBOT MÓVIL

HONORATO AZUCENA COYOTÉCATL  
JAVIER CABRERA MARTÍNEZ  
ELADIO FLORES MENA  
JOSÉ ELIGIO MOISÉS GUTIÉRREZ ARIAS  
FACULTAD DE CIENCIAS DE LA ELECTRÓNICA - BUAP

RESUMEN. Existen situaciones donde se desea estimar la calidad de estabilización de un sistema de control en las que se desconoce como funciona éste; pero si conocemos la entrada y la salida del mismo, estas situaciones permiten considerar los sistemas de pruebas de calidad de la estabilización robusta, como lo son las pruebas maximin. Las estrategias de pruebas maximin es una nueva dirección en la teoría de control óptimo y tienen su aplicación en la práctica de pruebas, su objetivo es evaluar el desempeño de algoritmos de estabilización desconocidos, así, mediante un sistema de pruebas determinar cuan bueno es dicho algoritmo de control. En el presente trabajo se aplican las pruebas maximin a un robot móvil, determinando así, la calidad de un controlador desconocido aplicado a dicho robot.

## 1. INTRODUCCIÓN

Los sistemas automáticos de control surgen inicialmente para reemplazar a los humanos en la realización de tareas de control poco deseables, el perfeccionamiento de las técnicas empleadas permite obtener sistemas de control con una elevada precisión o rapidez. Hacia 1960 y como respuesta a las nuevas necesidades (plantas con múltiples entradas y salidas, elevada complejidad), surge la teoría moderna de control, la cual está basada en el análisis y la síntesis en el dominio del tiempo empleando variables de estado. Los desarrollos más avanzados están representados por conceptos como el control óptimo, siendo éste, un tópico especial dentro de la disciplina de las ecuaciones diferenciales. Con el desarrollo del principio del máximo por Pontryagin (1958) y el método de la programación dinámica por Bellman (1957), actualmente, la teoría de control óptimo se considera como uno de los elementos de los métodos de optimización.

La idea de control puede ser expresada como el proceso mediante el cual se ejerce una influencia sobre el comportamiento de un objeto dinámico (que varía con el tiempo) para alcanzar un propósito previamente fijado [2]; pero si ahora, se desea lograr tal propósito en un tiempo mínimo o con un mínimo uso de recursos de control, entonces, este es un problema de control óptimo, ya que se quiere minimizar un funcional que depende del estado del sistema y del control, de esta manera, podemos decir que este funcional describe la calidad de estabilización.

Cuando existen perturbaciones en el sistema dinámico es posible considerar el peor caso y suponer que la perturbación está tratando de maximizar el mismo criterio de ejecución que el control trata de minimizar, se dice entonces que se

tiene un juego dinámico. El problema fundamental es lograr la optimización de un criterio de ejecución  $\mathcal{L}(u, v)$  entre las dos partes, con el controlador  $u$  como el minimizador y la perturbación  $v$  como el maximizador. El soporte de estas pruebas consiste en la solución de un problema extremal del tipo maximin.

Un juego dinámico es cualquier situación en la cual compiten dos o más jugadores donde cada jugador toma decisiones en cada instante de tiempo. Se supone la siguiente situación: un "jugador 1" sospecha que su adversario no se comporta racionalmente, entonces tiene sentido que adopte una estrategia maximin, esto es, aquella en la que maximiza la ganancia mínima que puede obtener [1]; para descubrir la estrategia que le conviene más se analiza una matriz que indica pagos, el "jugador 1" ignora la estrategia del otro jugador, mira cual es el mínimo resultado que puede obtener con cada una de sus cartas. En la figura (1) se ha añadido una columna indicando los resultados mínimos.

		La estrategia del otro jugador			mínimos
		A	B	C	
Mi estrategia	A	9	1	2	1
	B	6	5	4	4
	C	7	8	3	3

FIGURA 1. *Matriz de pagos.*

En efecto: si elije la tarjeta A puede obtener 9, 1 o 2, luego como mínimo obtendrá un resultado de 1, si elije la tarjeta B puede obtener 6, 5 o 4, luego como mínimo obtendrá 4, si elije la tarjeta C puede obtener 7, 8 o 3, luego como mínimo obtendrá 3. De todos esos posibles resultados mínimos, el que le conviene es 4 ya que es el máximo de los mínimos. La estrategia maximin consiste en elegir la tarjeta B ya que esa estrategia garantiza que como mínimo obtendrá 4.

## 2. PLANTEAMIENTO DE LAS PRUEBAS MAXIMIN

Consideremos el siguiente sistema de ecuaciones en desviaciones del movimiento programado:

$$(1) \quad \begin{cases} \dot{x} = Ax + Bu + Cv, & x(t_0) \in X_0, \\ u(\cdot) \in U_0 = \{u(\cdot) \in KC : |u_i(t)| \leq \nu_i, i = 1, \dots, r\}, \\ v(\cdot) \in V_0 = \{v(\cdot) \in KC : |v_j(t)| \leq \mu_j, j = 1, \dots, s\}, \end{cases}$$

donde  $x(t)$  es el  $n$ -vector de desviaciones de una trayectoria dada,  $v(\cdot)$  es el vector  $s$ -dimensional de perturbaciones permanentes,  $u(\cdot)$  es el vector  $r$ -dimensional de controles estabilizadores,  $X_0$  es el conjunto de desviaciones iniciales que también consideraremos como perturbaciones,  $\nu_i$  y  $\mu_i$  son parámetros dados que describen los recursos de control y la perturbación permanente;  $U_0$  y  $V_0$  son conjuntos de funciones continuas por tramos ( $KC$ ) en los cuales el constructor trabaja prácticamente.

Se supone que el sistema (1) es estabilizable y completamente controlable ( $x(t) \rightarrow 0$  cuando  $t \rightarrow \infty$ ) para cualquier  $v \equiv const \in V_0$ , si en el correspondiente algoritmo de estabilización  $u \in U$  es usado. Supongamos que la calidad de la estabilización es

descrita por (2) el cual expresa que tan rápido un estado inicial es llevado al estado deseado

$$(2) \quad \mathcal{L}(u, v) = \int_{t_0}^{t_1} (x^T G x + u^T N u) dt + x^T(t_1) S x(t_1),$$

donde  $t_1$  es el instante de tiempo con el cual la variedad de las desviaciones dada en el espacio se alcanza por primera vez (es posible considerar el caso de un  $t_1$  fijo,  $0 < t_1 \leq \infty$ ). La expresión (2) también es conocida como funcional objetivo y es una medida cuantitativa del comportamiento del sistema en el tiempo bajo la ley de control. El primer sumando del funcional es una integral que depende de los valores que va tomando  $x(t)$  y  $u(t)$  a lo largo del horizonte temporal, por lo tanto, valora el comportamiento del sistema a través del tiempo. El segundo sumando valora el estado en que queda el sistema al final del intervalo de tiempo que constituye el horizonte temporal del problema. Las matrices  $G$ ,  $N$  y  $S$  son constantes y satisfacen que  $G \geq 0$ ,  $N > 0$  y  $S \geq 0$ .

Denotemos a  $w = (x(0), v) \in W_0$  donde  $W_0 = X_0 \times V_0$  y se formulan los siguientes problemas de síntesis y análisis de sistemas de control [5]:

a) El problema de estabilización minimax.

$$(3) \quad \inf_{u(\cdot) \in U_0} \sup_{w(\cdot) \in W_0} \mathcal{L}(u, w),$$

el cual podemos interpretar como: dadas las peores condiciones iniciales y las peores perturbaciones se quieren minimizar las desviaciones del sistema (1) en el tiempo.

b) El problema de las pruebas maximin.

$$(4) \quad \sup_{w(\cdot) \in W_0} \inf_{u(\cdot) \in U_0} \mathcal{L}(u, w),$$

el cual se plantea bajo las mismas condiciones que el problema anterior, es decir: dado un control el cual provoque las mínimas desviaciones del sistema (1) se aplican las peores perturbaciones y condiciones iniciales con el fin de obtener un parámetro de comparación para cualquier otro controlador, este problema es considerado como el soporte del procedimiento de pruebas.

Es importante mencionar, que conocemos el sistema (1) para realizar la evaluación, el criterio (2) y los conjuntos  $U_0$  y  $W_0$ . Introducimos un conjunto  $U$ , con la finalidad de resolver el problema (4) de manera más simple y asumimos que en  $W_0$  se alcanza el máximo.

Supongamos que existe un algoritmo de control desconocido del que únicamente conocemos la salida  $\tilde{u} = \tilde{u}(t, x)$  y queremos organizar un sistema de pruebas para saber que tan bueno es este algoritmo, es decir, queremos otorgarle una calificación a dicho algoritmo.

Para establecer el planteamiento general de las pruebas maximin, es sabido que la siguiente cadena de desigualdades se cumple [6]:

$$(5) \quad \begin{aligned} \mathcal{L}_0^0 = \mathcal{L}(u^0, w^0) &= \max_{w \in W_0} \min_{u \in U_0} \mathcal{L}(u, w) \leq \min_{u \in U_0} \max_{w \in W_0} \mathcal{L}(u, w) \leq \\ &\leq \min_{u \in U_0} \max_{w \in W_0} \mathcal{L}(u, w) \leq \mathcal{L}(\tilde{u}, w^0) = \tilde{\mathcal{L}}. \end{aligned}$$

Entonces podemos considerar como nota o calificación excelente para el algoritmo de estabilización:

$$\mathcal{L}_0^0 = \max_{w(\cdot) \in W_0} \min_{u(\cdot) \in U_0} \mathcal{L}(u, w).$$

Algunas veces el algoritmo podrá alcanzar esta calificación excelente y otras veces no [3], dependiendo de si en el caso particular que se está analizando existe un punto de silla o no en la expresión (5), es decir un punto  $\{u^0, w^0\}$  para el cual

$$(6) \quad \max_{w(\cdot) \in W_0} \min_{u(\cdot) \in U_0} \mathcal{L}(u, w) = \min_{u(\cdot) \in U_0} \max_{w(\cdot) \in W_0} \mathcal{L}(u, w).$$

El esquema de la figura (2) muestra como se realizan las pruebas maximin: el sistema de control ubicado en el bloque "Algoritmo de control desconocido" actúa como una caja negra, ya que sólo conocemos información de su entrada:  $x(t)$ , así como de la salida, que será la expresión de control:  $\tilde{u}$ .

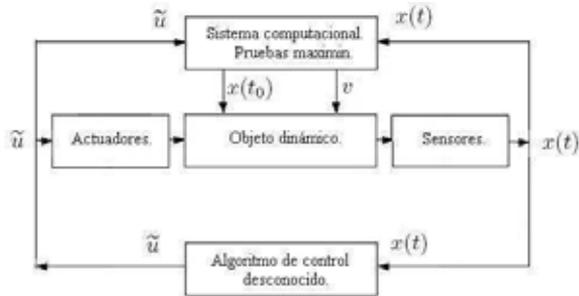


FIGURA 2. Esquema de pruebas maximin.

Al bloque "Sistema computacional. Pruebas Maximin" entran: el control  $\tilde{u}$  y las desviaciones del sistema  $x(t)$ , este bloque ejecuta las pruebas maximin, de esta manera posibilita la obtención de una evaluación del algoritmo de estabilización y la obtención de información sobre las condiciones iniciales  $x(t_0)$  y las perturbaciones  $v$ .

El procedimiento de las pruebas maximin está constituido por tres etapas [7]:

En la **primera etapa**, la solución del problema maximin, la estrategia óptima  $u^0(x, t)$ , la estrategia de las pruebas  $w^0(x, u, t)$  (si existe) y el valor del juego dinámico (4) ( $\mathcal{L}_0^0$  es la estimación inferior, es decir, la calificación excelente), son determinados; y de la primera desigualdad en (5) convenimos en tomar como parámetro de evaluación o nota excelente a la solución del problema extremal (4).

La **segunda etapa** se lleva a cabo dentro del marco de la simulación dinámica, cuando generada una contraestrategia  $w^0(x, u, t)$ , en respuesta a la acción de salida  $\tilde{u}(x, t)$  obtenemos la estimación de estabilización real  $\tilde{\mathcal{L}}$ . El algoritmo de estabilización es considerado para nosotros como una caja negra, dado que la forma en que se construyó nos es desconocida, únicamente tenemos conocimiento de la señal de salida de dicho algoritmo de estabilización. Sin embargo, en algunas ocasiones

se considera que la estructura del algoritmo a evaluar es lineal,  $\tilde{u} = K(t) x$ , donde los parámetros  $K(t)$ , son desconocidos para nosotros.

En la **tercera etapa**, el resultado de la prueba  $\tilde{\mathcal{L}}$ , se compara con la estimación inferior  $\mathcal{L}_0^0$ :

$$(7) \quad k_0 = \frac{\mathcal{L}_0^0}{\tilde{\mathcal{L}}(w^0(x, u, t), \tilde{u}(x, t))} \cdot (100),$$

la relación (7) es una estimación para la calidad de la estabilización, mientras más cercana a la nota excelente se encuentre la estimación mejor será el algoritmo de control que se está evaluando.

### 3. PRUEBAS MAXIMIN PARA UN ROBOT MÓVIL

Ahora se aplican las pruebas maximin a un ejemplo práctico, se desarrolla la estrategia de estas pruebas para evaluar un algoritmo de estabilización aplicado a un robot móvil.

Consideramos la clase de robots móviles autónomos que consisten de tres ruedas, dos activas y una pasiva, con restricciones no-holonómicas que aparecen como consecuencia de la hipótesis de no deslizamiento. El robot consta de un arreglo de sensores ubicado al frente; el centro de éste es el punto  $P$ , el cual será utilizado para determinar la posición del robot dentro del marco de referencia inercial formado por los ejes  $\xi$  y  $\eta$ , el ángulo  $\alpha$  nos sirve para determinar el giro del robot,  $m$  es el centro de masa del robot. Las velocidades del centro de las ruedas son denotadas como  $v_r$  y  $v_l$ ,  $F_r$  y  $F_l$  son las fuerzas de empuje de las llantas,  $R_r$  y  $R_l$  son las fuerzas reactivas perpendiculares al plano de rotación de las llantas, ver figura (3).

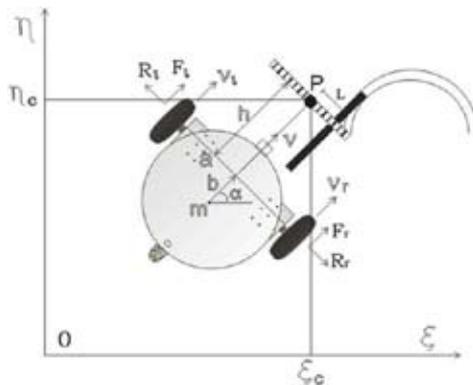


FIGURA 3. Sistema de coordenadas inercial  $(\xi, \eta)$ .

La posición del robot (centro del arreglo de sensores:  $P$ ) con respecto al marco de referencia inercial  $(\xi, \eta)$ , está dada por las siguientes relaciones [4]:

$$(8) \quad \begin{cases} \dot{\alpha} = \omega, \\ \dot{\xi} = v \cos \alpha - h \sin \alpha, \\ \dot{\eta} = v \sin \alpha + h \cos \alpha, \end{cases}$$

el robot es un sistema no-holonómico y las velocidades de las ruedas activas deben tener la dirección del eje  $X$  (el sistema de referencia  $(X, Y)$ , tiene situado el origen en la intersección del eje de simetría del robot y el centro del eje de las ruedas activas).

Para obtener las ecuaciones dinámicas del robot móvil, se utilizarán las ecuaciones de movimiento que plantea el "Principio de D'Alembert" para el equilibrio dinámico de un cuerpo rígido; dichas ecuaciones son:

$$(9) \quad \sum F = m\bar{a},$$

$$(10) \quad \sum M = \dot{L},$$

en donde la ecuación (9) relaciona la resultante de las fuerzas externas y la aceleración del centro de masa del sistema y la ecuación (10) relaciona el momento resultante de las fuerzas externas y la razón de cambio de la cantidad de movimiento angular o momento angular del sistema. Entonces, utilizando este principio se obtienen las siguientes relaciones dinámicas para el sistema de referencia inercial  $(\xi, \eta)$ :

$$(11) \quad \begin{aligned} m(\dot{v} + b\omega^2) &= F_r + F_l, \\ \tilde{J}\dot{\omega} + mb\omega v &= (F_r - F_l)a. \end{aligned}$$

Asociamos las fuerzas  $F_r$  y  $F_l$  a los motores del robot móvil; para esto, se sabe que el torque o momento de un motor está dado por:

$$(12) \quad M = F\rho,$$

donde  $M$  es el torque del motor,  $F$  la fuerza aplicada por el mismo y  $\rho$  es el radio de la rueda acoplada al motor. El modelo más simple conocido para un motor DC que describe su movimiento y relaciona el torque del motor con la velocidad angular del mismo es:

$$(13) \quad M = \chi u - \sigma \dot{\varphi},$$

donde  $M$  es el torque del motor y el miembro derecho es la diferencia de la fricción viscosa y la fuerza contraelectromotriz. Entonces para la rueda derecha se tienen las relaciones (14)

$$(14) \quad F_r = \frac{\chi u_r - \sigma \dot{\varphi}_r}{\rho} \quad ; \quad \dot{\varphi}_r = \frac{v_r}{\rho} = \frac{v + \omega a}{\rho}.$$

Análogamente se tienen las relaciones para la rueda izquierda:

$$(15) \quad F_l = \frac{\chi u_l - \sigma \dot{\varphi}_l}{\rho} \quad ; \quad \dot{\varphi}_l = \frac{v_l}{\rho} = \frac{v - \omega a}{\rho}.$$

Sustituyendo las expresiones (14) y (15) en las ecuaciones de movimiento (11) y realizando operaciones finalmente se obtienen:

$$(16) \quad \begin{aligned} m\dot{v} + mb\omega^2 + \frac{2\sigma}{\rho^2}v &= \frac{\chi}{\rho}(u_r + u_l), \\ \tilde{J}\dot{\omega} + mb\omega v + \frac{2\sigma a}{\rho^2}\omega &= \frac{\chi}{\rho}(u_r - u_l). \end{aligned}$$

Por lo tanto las ecuaciones (8) y (16) describen la dinámica del robot móvil:

$$(17) \quad \begin{cases} \dot{\xi} = v \cos \alpha - h \sin \alpha, \\ \dot{\eta} = v \sin \alpha + h \cos \alpha, \\ \dot{\alpha} = \omega, \\ m\dot{v} = -mb\omega^2 - \frac{2\sigma}{\rho^2}v + \frac{\chi}{\rho}(u_r + u_l), \\ \tilde{J}\dot{\omega} = mb\omega v - \frac{2a^2\omega\sigma}{\rho^2} + \frac{\chi a}{\rho}(u_r - u_l). \end{cases}$$

Cualquier trayectoria deseada se puede conseguir realizando combinaciones de trayectorias rectas y semicírculos como se muestra en la figura (4), se observa que el robot puede realizar movimientos en líneas rectas horizontales o verticales en cualquier sentido, así como círculos también en ambos sentidos. La tabla (1) muestra el conjunto de trayectorias deseadas, cuando el movimiento se realiza a lo largo de un segmento de línea paralela al eje  $\xi$  o al eje  $\eta$ .

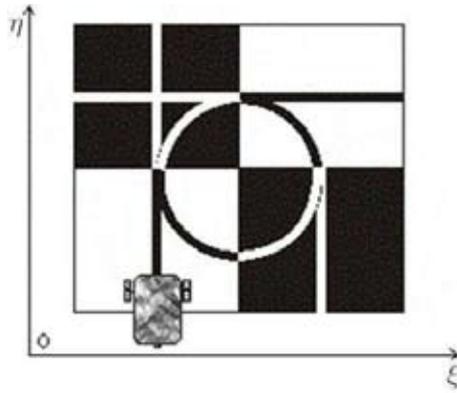


FIGURA 4. Conjunto de posibles trayectorias.

Trayectoria	$\xi^d$	$\eta^d$	$\alpha^d$	$v^d$	$\omega^d$
1.Línea horizontal sentido negativo	$v_0 t + \xi_0$	0	0	$v_0$	0
2.Línea horizontal sentido positivo	$v_0 t + \xi_0$	0	$\pi$	$v_0$	0
3.Línea vertical sentido negativo	0	$v_0 t + \eta_0$	$\frac{\pi}{2}$	$v_0$	0
4.Línea vertical sentido positivo	0	$v_0 t + \eta_0$	$-\frac{\pi}{2}$	$v_0$	0

TABLA 1. Trayectorias programadas de las líneas horizontal y vertical.

Considerando que los movimientos circulares se pueden realizar en ambos sentidos de las manecillas del reloj, entonces para este tipo de movimientos se tienen 8 posibles trayectorias programadas como se muestra en la tabla (2).

Trayectoria	$\xi^d$	$\eta^d$	$\varphi \in$	$\alpha^d$	$v^d$	$\omega^d$
5. Cuadrante I dirección reloj	$R\cos\varphi$	$R\sin\varphi$	$[\frac{\pi}{2}, 0]$	$-\omega_0 t$	$\omega_0 R$	$\omega_0$
6. Cuadrante I contra reloj	$R\cos\varphi$	$R\sin\varphi$	$[0, \frac{\pi}{2}]$	$\omega_0 t + \frac{\pi}{2}$	$\omega_0 R$	$\omega_0$
7. Cuadrante II dirección reloj	$R\cos\varphi$	$R\sin\varphi$	$[\pi, \frac{\pi}{2}]$	$-\omega_0 t + \frac{\pi}{2}$	$\omega_0 R$	$\omega_0$
8. Cuadrante II contra reloj	$R\cos\varphi$	$R\sin\varphi$	$[\frac{\pi}{2}, \pi]$	$\omega_0 t + \pi$	$\omega_0 R$	$\omega_0$
9. Cuadrante III dirección reloj	$R\cos\varphi$	$R\sin\varphi$	$[-\frac{\pi}{2}, -\pi]$	$\pi - \omega_0 t$	$\omega_0 R$	$\omega_0$
10. Cuadrante III contra reloj	$R\cos\varphi$	$R\sin\varphi$	$[-\pi, -\frac{\pi}{2}]$	$\omega_0 t - \frac{\pi}{2}$	$\omega_0 R$	$\omega_0$
11. Cuadrante IV dirección reloj	$R\cos\varphi$	$R\sin\varphi$	$[0, -\frac{\pi}{2}]$	$-\omega_0 t - \frac{\pi}{2}$	$\omega_0 R$	$\omega_0$
12. Cuadrante IV contra reloj	$R\cos\varphi$	$R\sin\varphi$	$[-\frac{\pi}{2}, 0]$	$\omega_0 t$	$\omega_0 R$	$\omega_0$

TABLA 2. Trayectorias programadas de los segmentos de círculo.

Si  $u^d$  es una entrada nominal al sistema descrito por las ecuaciones (17) y  $y^d$  es una trayectoria nominal deseada de dicho sistema, utilizamos el proceso de linealización con las trayectorias programadas obteniendo así los respectivos sistemas de ecuaciones lineales en desviaciones, para obtener la linealización del sistema aplicamos el Jacobiano o matriz Jacobiana a las ecuaciones del robot móvil dadas en (17).

Para una trayectoria en forma de línea vertical en sentido positivo (trayectoria 4, tabla (1)) el sistema linealizado es:

$$(18) \quad \begin{cases} \dot{\xi} = v_0 \alpha + h \omega, \\ \dot{\eta} = -v, \\ \dot{\alpha} = \omega, \\ \dot{v} = \frac{2\sigma}{m\rho^2} v + \frac{\chi}{m\rho} (u_r + u_l), \\ \dot{\omega} = \frac{mb\omega v_0}{J} - \frac{2a^2 \omega \sigma}{J\rho^2} + \frac{\chi a}{J\rho} (u_r - u_l). \end{cases}$$

Si ahora, la trayectoria deseada es en forma circular, que esté ubicada en el primer cuadrante del plano formado por los ejes  $\xi$  y  $\eta$  y con un sentido contrario a las manecillas del reloj (trayectoria 6, tabla (2)), se obtiene el siguiente sistema lineal:

$$(19) \quad \begin{cases} \dot{\xi} = -(\omega_0 R \sin \alpha^d + h \omega_0 \cos \alpha^d) \alpha + v \cos \alpha^d - h \omega \sin \alpha^d, \\ \dot{\eta} = (\omega_0 R \cos \alpha^d + h \omega_0 \sin \alpha^d) \alpha + v \sin \alpha^d + h \omega \cos \alpha^d, \\ \dot{\alpha} = \omega, \\ \dot{v} = -\frac{2\sigma}{\rho^2} v - 2b\omega_0\omega + \frac{\chi}{m\rho} (u_r + u_l), \\ \dot{\omega} = \frac{mb\omega_0}{J} v - \frac{mb\omega_0 R}{J} \omega - \frac{2a^2\sigma}{J\rho^2} \omega + \frac{\chi a}{J\rho} (u_r - u_l). \end{cases}$$

Para las restantes trayectorias circulares los sistemas linealizados son análogos y la diferencia entre ellos dependerá del valor de  $\alpha^d$  respectivamente.

#### 4. ETAPA 1 DE LAS PRUEBAS MAXIMIN: DETERMINACIÓN DE LA ESTIMACIÓN INFERIOR.

En esta etapa de las pruebas es importante recordar el planteamiento: se desea evaluar el desempeño de un algoritmo de control desconocido del que únicamente conocemos la salida  $\tilde{u} = \tilde{u}(t, x)$ , es decir, se desea otorgarle una calificación, para hacer ésto, es necesario contar con una calificación excelente que sirva como parámetro para comparar la calificación obtenida por el control que se evaluará.

Tomando en cuenta el argumento anterior y sabiendo que la cadena de desigualdades (5) siempre se cumple, podemos considerar como calificación excelente la solución del problema extremal  $\max_{w(\cdot) \in W_0} \min_{u(\cdot) \in U_0} \mathcal{L}(u, w)$  a la cual se le ha denominado: determinación de la estimación inferior ( $\mathcal{L}_0^0$ )

$$\mathcal{L}_0^0 = \max_{w(\cdot) \in W_0} \min_{u(\cdot) \in U_0} \mathcal{L}(u, w),$$

Para solucionar el problema interno  $\min_{u(\cdot)} \mathcal{L}(u)$  del problema extremal anterior hacemos uso de la programación dinámica de Bellman. Consideramos que el criterio de desempeño (2) es minimizado dada la dinámica del proceso controlable (1), definimos la función de costo mínima (20) llamada función de Bellman,

$$(20) \quad \omega(x, t) = \min_{u(t)} \left\{ \int_{t_0}^{t_1} (x^T G x + u^T N u) dt + x^T(t_1) S x(t_1) \right\},$$

con la cual determinamos los controles que minimizen (2). Asumiendo que las derivadas parciales de  $\omega$  existen y son acotadas, se escribe la ecuación de Hamilton-Jacobi-Bellman, donde  $\omega(x, t)$  debe satisfacer la ecuación en derivadas parciales.

$$(21) \quad 0 = \frac{\partial \omega(x, t)}{\partial t} + \min_{u(t)} \left\{ x^T G x + u^T N u + \frac{\partial \omega(x, t)}{\partial x} [Ax + Bu] \right\},$$

donde  $\omega(x(t_1), t_1) = x^T(t_1) S x(t_1)$ .

Asumimos que la solución de la ecuación (21) es de la forma  $\omega(x, t) = x^T P(t) x$  con la condición de frontera  $w(x, t_1) = x^T(t_1) S x(t_1) = 0$ .

Para encontrar  $u(t)$  que da valor mínimo a  $\left\{ x^T G x + u^T N u + \frac{\partial \omega(x, t)}{\partial x} [Ax + Bu] \right\}$  derivamos con respecto a  $u$  e igualamos con cero, obteniendo ( $N > 0$ ):

$$(22) \quad u = -N^{-1} B^T P(t) x.$$

Como la segunda derivada de  $\left\{x^T G x + u^T N u + \frac{\partial \omega(x,t)}{\partial x} [Ax + Bu]\right\}$  con respecto a  $u$  es la matriz simétrica y definida positiva  $2N$ , se tiene que (22) da la función vectorial que minimiza a  $\left\{x^T G x + u^T N u + \frac{\partial \omega(x,t)}{\partial x} [Ax + Bu]\right\}$ .

Sustituyendo el control (22) en la ecuación (21) se obtiene:

$$(23) \quad 0 = x^T \left\{ \frac{\partial P(t)}{\partial t} + G + P(t)A + A^T P(t) - P(t)BN^{-1}B^T P(t) \right\} x,$$

donde:

$$(24) \quad \frac{\partial P(t)}{\partial t} + G - P(t)BN^{-1}B^T P(t) + P(t)A + A^T P(t) = 0,$$

es la llamada ecuación diferencial matricial de Riccati con condiciones finales  $P(t_1) = 0$  la cual debe resolverse hacia atrás en el tiempo para encontrar  $P(t)$ .

Así pues obtenemos:

$$(25) \quad u^0 = -N^{-1}B^T P(t)x,$$

donde  $P(t) \forall t \leq t_1$  es solución de (24). Además  $\omega(x_0, t_0) = x^T(t_0)P(t_0)x(t_0)$  y por (20) la solución del problema interno de (4) esta dada por:

$$(26) \quad \min_{u(\cdot)} \mathcal{L}(u) = x^T(t_0)P(t_0)x(t_0).$$

Para la determinación de la estimación inferior consideramos cuatro movimientos programados; en este caso las trayectorias 2 y 4 de la tabla (1), así como las trayectorias 6 y 7 de la tabla (2) como se muestra en la figura (5).

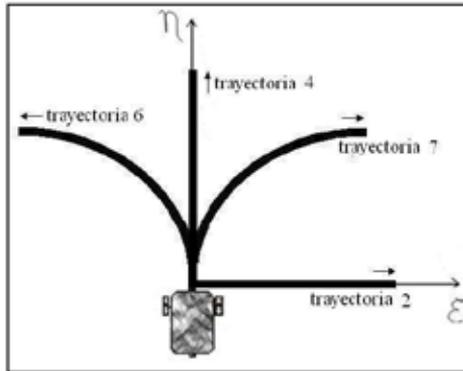


FIGURA 5. Trayectorias programadas.

Como consecuencia el conjunto  $W_0$  se considera un conjunto discreto, por lo que mediante la programación dinámica resolvemos el problema interno  $\min_{u(\cdot)} \mathcal{L}(u)$  cuatro veces, ya que se resuelve para cada uno de los sistemas de ecuaciones en desviaciones generados por las trayectorias programadas, posteriormente tomamos el máximo valor dando solución así al problema externo

$$\max_{w \in W_0 = \{1,2,3,4\}} \left( \min_{u(\cdot)} \mathcal{L}(u, w) \right),$$

determinando con ello la estimación inferior  $L_0^0$ .

Para el desarrollo de las pruebas es necesario considerar los valores numéricos de los parámetros del robot móvil [8], éstos se muestran en la tabla (3) .

Variable	Valor	Descripción
$v_0$	1.5	Velocidad inicial [m/s]
$a$	0.4	Distancia entre las ruedas [m]
$b$	0.4	Distancia del centro de masa al eje de las ruedas [m]
$h$	0.1	Distancia del eje de las ruedas al arreglo de sensores infrarrojos [m]
$m$	4.5	Masa del robot [Kg]
$\rho$	0.08	Radio de las ruedas [m]
$R$	0.35	Radio de inercia del carro [m]
$\chi$	0.01	Fricción viscosa
$\sigma$	0.009	Fuerza contraelectromotriz del motor

TABLA 3. *Parámetros del robot móvil.*

Así, el sistema lineal para la primera trayectoria deseada: línea circular en el cuadrante I con sentido contrario a las manecillas del reloj (trayectoria 6, tabla (2)) está dado en (27)

$$(27) \quad \begin{cases} \dot{\xi} = -(0.175 \sin \alpha^d + 0.05 \cos \alpha^d) \alpha + v \cos \alpha^d - 0.1 \omega \sin \alpha^d, \\ \dot{\eta} = (0.175 \sin \alpha^d - 0.05 \cos \alpha^d) \alpha + v \sin \alpha^d + 0.1 \omega \cos \alpha^d, \\ \dot{\alpha} = \omega, \\ \dot{v} = -0.625v + 0.4\omega + 0.0277(u_r + u_l), \\ \dot{\omega} = 3.1374v - 0.4706\omega - 0.1743(u_r - u_l). \end{cases}$$

El sistema (27) se puede expresar de la forma  $\dot{x} = Ax + Bu$  como se observa en (28):

$$(28) \quad \begin{bmatrix} \dot{\xi} \\ \dot{\eta} \\ \dot{\alpha} \\ \dot{v} \\ \dot{\omega} \end{bmatrix} = \begin{bmatrix} 0 & 0 & -(0.175\sin\alpha^d + 0.05\cos\alpha^d) & \cos\alpha^d & -0.1\sin\alpha^d \\ 0 & 0 & (0.175\sin\alpha^d - 0.05\cos\alpha^d) & \sin\alpha^d & 0.1\cos\alpha^d \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -0.625 & 0.4 \\ 0 & 0 & 0 & 3.1374 & -0.4706 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \\ \alpha \\ v \\ \omega \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0.0277 & 0.0277 \\ -0.1743 & 0.1743 \end{bmatrix} \begin{bmatrix} u_l \\ u_r \end{bmatrix}.$$

Consideramos el siguiente funcional:

$$(29) \quad \mathcal{L} = \int_0^{t_1} (x^T G x + u^T N u) dt,$$

con

$$(30) \quad G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad y \quad N = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

las cuales son matrices que determinan cuanto peso le corresponde a cada uno de los componentes del estado y del control.

Hagamos uso de la programación dinámica para dar solución al problema interno, para esto resolvemos la ecuación diferencial matricial de Riccati obtenida en la ecuación (24), utilizando las matrices  $A$  y  $B$  del sistema (28) así como las matrices  $G$  y  $N$  dadas en (30)

$$\frac{\partial P(t)}{\partial t} + G - P(t)BN^{-1}B^T P(t) + P(t)A + A^T P(t) = 0,$$

con la condición de contorno  $\omega(x, t_1) = 0$ .

De esta manera la solución de la ecuación diferencial de Riccati (24) que se obtiene es:

$$(31) \quad P(t_0) = \begin{bmatrix} 7.3459 & -0.0548 & 2.3223 & 7.9619 & 0.8372 \\ -0.0548 & 7.4982 & -0.1796 & -3.6565 & 2.0393 \\ 2.3223 & -0.1796 & 7.0224 & 13.5814 & 3.0014 \\ 7.9619 & -3.6565 & 13.5814 & 35.9616 & 5.3591 \\ 0.8372 & 2.0393 & 3.0014 & 5.3591 & 2.7074 \end{bmatrix},$$

donde el tiempo final fue  $t_1 = 8 \text{ seg}$ .

Las peores condiciones iniciales para este sistema son [8]:

$$(32) \quad x(t_0) = \begin{bmatrix} 0.714 \\ 0.33 \\ 0.1 \\ 0.11 \\ 0.27 \end{bmatrix}.$$

Sabemos que la solución del problema interno está dado por:

$$\min_{u(\cdot)} \mathcal{L}(u) = x^T(t_0)P(t_0)x(t_0),$$

entonces, utilizando (31) y (32) encontramos el valor del problema interno para el sistema (27), el cual es: 8.0087. Análogamente resolvemos el problema interno para los 3 sistemas restantes, encontramos los sistemas linealizados alrededor de las trayectorias programadas mostradas en la figura (5).

Así, el sistema lineal para la segunda trayectoria en línea vertical con sentido positivo (trayectoria 4, tabla (1)) es:

$$(33) \quad \begin{cases} \dot{\xi} = 1.5\alpha + 0.1\omega, \\ \dot{\eta} = -v, \\ \dot{\alpha} = \omega, \\ \dot{v} = -0.625v + 0.0277(u_r + u_l), \\ \dot{\omega} = 7.8437\omega - 0.1743(u_r - u_l). \end{cases}$$

De la misma forma obtenemos el sistema lineal para una tercera trayectoria en forma circular en el cuadrante II con sentido a las manecillas del reloj (trayectoria 7, tabla (2))

$$(34) \quad \begin{cases} \dot{\xi} = -(0.175\sin\alpha^d + 0.05\cos\alpha^d)\alpha + v\cos\alpha^d - 0.1\omega\sin\alpha^d, \\ \dot{\eta} = (0.175\sin\alpha^d - 0.05\cos\alpha^d)\alpha + v\sin\alpha^d + 0.1\omega\cos\alpha^d, \\ \dot{\alpha} = \omega, \\ \dot{v} = -0.625v + 0.4\omega + 0.0277(u_r + u_l), \\ \dot{\omega} = 3.1374v - 0.4706\omega - 0.1743(u_r - u_l). \end{cases}$$

Finalmente obtenemos el sistema linealizado para una cuarta trayectoria en línea horizontal con sentido positivo (trayectoria 2, tabla (1)):

$$(35) \quad \begin{cases} \dot{\xi} = -v, \\ \dot{\eta} = -1.5\alpha - 0.1\omega, \\ \dot{\alpha} = \omega, \\ \dot{v} = -0.625v + 0.0277(u_r + u_l), \\ \dot{\omega} = 7.8437\omega - 0.1743(u_r - u_l). \end{cases}$$

Con la programación dinámica resolvemos los problemas internos para cada uno de los sistemas (33), (34) y (35), utilizamos sus correspondientes matrices  $A$  y  $B$ , así como el funcional (29) y las matrices (30) para todos los sistemas. Una vez que se soluciona la ecuación diferencial matricial de Riccati ( $P(t_0)$ ) correspondiente a cada uno de los sistemas y utilizando las condiciones iniciales (32) obtenemos los siguientes resultados.

- (1) El valor del problema interno para el sistema (27) es: 8.0087.

- (2) El valor del problema interno para el sistema (33) es : 32.0577.  
 (3) El valor del problema interno para el sistema (34) es : 7.8776.  
 (4) El valor del problema interno para el sistema (35) es : 23.5298.

Al resolver el problema externo  $\max_{u \in W_0 = \{1,2,3,4\}} \left( \min_{u(\cdot)} \mathcal{L}(u, w) \right)$  encontramos la estimación inferior para las pruebas maximin, la cual está dada por:

$$(36) \quad \mathcal{L}_0^0 = 32.0577.$$

El valor óptimo del funcional alcanzado (36) será una cota inferior no mejorable para el algoritmo de estabilización, es decir, el algoritmo de estabilización que se quiera evaluar mientras más cerca se encuentre del valor  $\mathcal{L}_0^0$  mejor será la calidad de dicho algoritmo, con este resultado se concluye la primera etapa de estas pruebas.

### 5. ETAPAS 2 Y 3 DE LAS PRUEBAS MAXIMIN: EVALUACIÓN DE UN SISTEMA DE CONTROL DESCONOCIDO

Ahora analizamos la calidad de estabilización de un algoritmo de control desconocido. Se propone un control lineal de la forma  $\tilde{u} = K(t)x$  donde:

$$\tilde{u} = \begin{bmatrix} u_l \\ u_r \end{bmatrix} ; \quad K = \begin{bmatrix} k_{11} & k_{12} & k_{13} & k_{14} & k_{15} \\ k_{21} & k_{22} & k_{23} & k_{24} & k_{25} \end{bmatrix} ; \quad x = \begin{bmatrix} \xi \\ \eta \\ \alpha \\ v \\ \omega \end{bmatrix} .$$

Los parámetros de  $K(t)$  están dados por:

$$k_{11} = -(0.119411 \times 10^{-6})t^8 - (0.108389 \times 10^{-4})t^7 + (0.898817 \times 10^{-3})t^6 - 0.0227861t^5 + 0.0271979t^4 - 1.63701t^3 + 4.69189t^2 - 5.19183t + 123.582$$

$$k_{14} = -(0.354728 \times 10^{-5})t^8 + (0.25612 \times 10^{-3})t^7 - (0.751453 \times 10^{-2})t^6 + 0.114973t^5 - 0.978158t^4 + 4.55636t^3 - 10.6111t^2 + 9.8337t + 95.42$$

$$k_{22} = -(0.380492 \times 10^{-6})t^8 - (0.437618 \times 10^{-4})t^7 + (0.176443 \times 10^{-2})t^6 + 0.0343873t^5 + 0.35581t^4 - 1.95405t^3 + 5.24845t^2 - 5.52425t + 101.122$$

$$k_{23} = -(0.559137 \times 10^{-5})t^8 + (0.404527 \times 10^{-3})t^7 - 0.011901t^6 + 0.182708t^5 - 1.5608t^4 + 7.30545t^3 - 17.109t^2 + 15.9622t + 157.549$$

$$k_{25} = -(0.980086 \times 10^{-5})t^8 + (0.730996 \times 10^{-3})t^7 - 0.0221376t^6 + 0.349339t^5 - 3.06312t^4 + 14.6952t^3 - 35.228t^2 + 33.6019t + 101.944$$

$$k_{12} = k_{13} = k_{15} = k_{21} = k_{24} = 0$$

Así el control que será evaluado es (37):

$$(37) \quad \begin{cases} u_r = k_{11}\xi + k_{14}v, \\ u_l = k_{22}\eta + k_{23}\alpha + k_{25}\omega. \end{cases}$$

Utilizando el control (37) y sustituyendolo en los sistemas (27), (33), (34) y (35) junto con el índice de desempeño dado en (29) obtuvimos la contraestrategia de las

pruebas, que en este caso es el sistema (33), además determinamos la estimación real:

$$(38) \quad \tilde{\mathcal{L}} = 10003.$$

La última etapa de las pruebas consiste en realizar la comparación de la estimación real obtenida en (38) con la estimación inferior (36), para esto utilizamos la relación que nos determina la estimación para la calidad de la estabilización dada en (7)

$$(39) \quad k^0 = \frac{32.0577}{10003} \cdot 100 = 0.3205.$$

De acuerdo al resultado obtenido, podemos observar que el algoritmo de control evaluado presenta un bajo desempeño, ésto debido a que en el controlador no se involucran todos los estados.

## 6. CONCLUSIONES

Este trabajo se basa principalmente en el establecimiento y análisis de las pruebas maximin para la calidad de estabilización robusta y como su nombre lo dice, es una estrategia de pruebas que nos determinarán la calidad de un algoritmo de control, estas pruebas pueden ser aplicadas a cualquier sistema dinámico, para eso, solo es necesario conocer el modelo dinámico de dicho sistema.

Se aplicaron las estrategias de pruebas maximin para la calidad de la estabilización, al problema concreto de un algoritmo de estabilización de un robot móvil, utilizamos un indicador de desempeño en el cual las matrices  $G$  y  $N$  son propuestas por el diseñador, el método de la programación dinámica nos permitió encontrar la solución del problema interno, para eso resolvimos la ecuación diferencial matricial de Riccati y el valor de este resultado es solución del problema interno  $\min_u \mathcal{L}(u)$ , consideramos las condiciones iniciales pues éstas influyen en el comportamiento que presenta nuestro sistema dinámico, una vez resuelto el problema interno se resolvió el problema externo  $\max_w \left( \min_u \mathcal{L}(u, w) \right)$  para obtener así un parámetro con quien comparar cualquier controlador que se desee, esto en la primera etapa de las pruebas.

Finalmente se evaluó un control desconocido que tiene forma lineal, no se sabe como fue su planteamiento, sólo se pudo observar que los valores de sus parámetros están dados por una aproximación de polinomios en función del tiempo, una vez obtenida la estimación real (es decir la segunda etapa de las pruebas) se concluyó el trabajo comparando éste resultado con el de la estimación inferior para poder así, otorgarle una calificación.

## REFERENCIAS

- [1] BRAVO Raspeño Juan, "Teoría de Juegos, Historia de las Matemáticas.", pp. 7-9.
- [2] FERREYRA Guillermo y Pascal Jesús, "Control Óptimo Determinista vía Programación Dinámica," [en línea]; Divulgaciones Matemáticas Vol. 7 No. 2, 1999, pp. 167-185, [consulta: 19 de febrero de 2009] Disponible: <http://www.emis.de/journals/DM/v72/art6.pdf>
- [3] GUERRA Marrero Lileana, *Estabilización Minimax Pruebas Maximin de Sistemas Dinámicos Controlables*, Tesis de doctorado, México 2000, pp. 2-24.
- [4] GUTIÉRREZ Arias J. E. Moisés, *Estrategias de las Pruebas Maximin para la Calidad de la Estabilización*, Tesis de doctorado, México 2003, pp. 1-34.

- [5] V. V. Alexandrov y Gutiérrez Arias J. E. Moisés, *About the Maximin Testing Strategy*, Sociedad Matemática Mexicana, Aportaciones Matemáticas Memoria 35, México 2004, pp. 3-11.
- [6] V. V. Alexandrov y Gutiérrez Arias J. E. Moisés, *Pruebas Maximin y Juegos Geométricos*, Sociedad Matemática Mexicana, Aportaciones Matemáticas Memoria 35, México 2005, pp. 3-12.
- [7] V. A. Sadovnichii y otros, *Maximin Testing of Satellite Stabilization*, Mathematical Modeling of Complex Information Processing Systems, pp. 61-70.
- [8] V. V. Alexandrov y otros, *Optimization and Computer-Aided Testing of Stabilization Precision*, Mathematical Modeling of Complex Information Processing Systems, pp. 49-60.

Facultad de Ciencias de la Electrónica, BUAP  
Av. San Claudio y Blvd. 18 Sur, Col. Jardines de San Manuel  
C.P.72570, Puebla, Pue. México  
hazucena@ece.buap.mx  
javier-cm4@ece.buap.mx  
eflores@ece.buap.mx  
jmgutierrez@ece.buap.mx

# DISEÑO DE UN CONTROL ÓPTIMO CONTINUO PARA LA ESTABILIZACIÓN DE TRAYECTORIAS EN UN ROBOT MÓVIL

JAVIER CABRERA MARTÍNEZ  
JOSÉ ELIGIO MOISÉS GUTIERREZ ARIAS  
LUCIO HERNÁNDEZ ANGULO  
MARÍA MONSERRAT MORÍN CASTILLO  
FACULTAD DE CIENCIAS DE LA ELECTRÓNICA - BUAP

RESUMEN. Se plantea el problema de encontrar un control óptimo lineal para la estabilización de trayectorias en un robot móvil, mediante el uso de la programación dinámica como herramienta de síntesis. Se obtienen las ecuaciones dinámicas de movimiento del robot y las ecuaciones lineales de este alrededor de trayectorias deseadas. El resultado de la síntesis del control óptimo es una ecuación matricial diferencial de Riccati que es el elemento principal del control; en la literatura se trata a esta ecuación diferencial como una ecuación algebraica para un tiempo infinito. Además se presenta una modificación en la estructura del control óptimo que permite que la elección a prueba y error de las matrices de peso sea innecesaria y hace que los valores característicos del sistema sean colocados en una zona específica en el semiplano izquierdo del plano complejo. Se presentan gráficas del comportamiento obtenido para cada uno de los estados del sistema linealizado.

## 1. INTRODUCCIÓN

Un robot móvil es una máquina autónoma capaz de moverse en un ambiente dado, utilizando como herramienta alguna configuración mecánica que permita algún tipo de desplazamiento, combinado con algunos componentes computacionales. De acuerdo a su forma de locomoción se clasifican en: robots móviles con locomoción mediante orugas, mediante patas y mediante ruedas. Aunque la mayoría de los robots móviles actualmente son experimentales, ya están empezando a desarrollarse para fines industriales; las cuales, entre otras características, pueden ser inhóspitas, peligrosas o inaccesibles para seres humanos.

El caso más simple de robots móviles son los robots con ruedas; estos robots constan de una o más ruedas de tracción, tienen ruedas pasivas opcionales, y posiblemente ruedas de dirección. Considerando físicamente a un robot móvil como un conjunto de varios cuerpos sólidos con diferentes grados de movilidad, un robot móvil puede constar de 5 grados de libertad como en este caso; sin embargo, los robots móviles propulsados por ruedas a su vez se clasifican de acuerdo al número y al tipo de grados de libertad. Los grados de libertad considerados en robótica móvil son de dos tipos; de movilidad y de direccionabilidad. El grado de movilidad es un grado de libertad asociado a un variable de velocidad del robot y el grado de direccionabilidad es un grado de libertad asociado a una variable de dirección del robot. La configuración del robot móvil que se utilizará en el desarrollo de este trabajo, es el más comúnmente usado y es llamado “control diferencial”; esta configuración const!

a de dos ruedas para tracción y dirección, y una rueda pasiva para estabilidad de la plataforma del robot. Con lo anterior, se puede decir que se trata de un robot móvil de tipo (2,0). El diseño del control diferencial tiene dos motores montados en posiciones fijas al lado izquierdo y derecho del robot, independientemente tienen acoplada una rueda cada uno [1].

A causa de que se adiciona la complejidad del control para la coordinación del movimiento del robot, a la complejidad de una posible trayectoria en un ambiente posiblemente hostil, es importante la elección de un control que pueda minimizar los tiempos de estabilización a los posibles cambios de trayectoria que puede presentar el robot al recorrer el entorno dado.

En la literatura, el estudio y diseño de los sistemas de control para los robots móviles es muy amplio, así como también la cantidad de algoritmos de control con los que se puede encontrar una aplicación interesante. Aunque el robot móvil es un sistema con movimientos físicamente simples, matemáticamente no lo es, debido a que su modelo dinámico es no lineal. La aplicación de algoritmos de control para un robot móvil es basta, pasando desde algoritmos clásicos de control como en [2] donde se plantea un control PID y de linealización retroalimentada para el control de la parte dinámica (que involucra a la velocidad de los motores) y la parte cinemática (que involucra a la trayectoria deseada) respectivamente, hasta controles de tipo Difuso como en [3] donde el control se realiza en base a conjuntos de decisión y de respuesta, cabe mencionar que para este caso solo se utiliza el modelo cinemático del robot móvil. Trabajos más recientes como [4] presenta un algoritmo

de control para un robot móvil, basado en información visual para la estabilización del sistema. En [5] se presenta el modelo matemático del robot móvil en un sistema de referencia polar de 2 dimensiones, planteando el problema de un regulador o algoritmo de control basado en los teoremas de estabilidad de Lyapunov. En materia de control óptimo para robots móviles, en [6] se plantea el problema de encontrar esta ley de control considerando el modelo matemático y la información del error cometido en la orientación y la posición para una trayectoria deseada obteniendo una ecuación lineal de cuatro variables de estado, además haciendo una transformación sobre los valores característicos se obtiene un segundo sistema de dos variables de estado las cuales representa los modos lentos del sistema, el control óptimo es calculado solo para este segundo sistema.

## 2. ESTABLECIMIENTO DEL PROBLEMA

Considerando el siguiente proceso controlable que representa las ecuaciones dinámicas del robot móvil

$$(1) \quad \dot{y} = f(y, u) ,$$

$$(2) \quad u(\cdot) \in U = \{u : u(t) \in \Omega \subseteq \mathbb{R}^r\} ;$$

donde  $y \in \mathbb{R}^n$  que contiene los estados del sistema,  $u \in \mathbb{R}^m$  representa los controles de entrada. Las ecuaciones diferenciales lineales que gobiernan las desviaciones  $x(t) = y(t) - y^d(t)$  para algún movimiento deseado  $y(t) = y^d(t)$  y un control deseado  $u(t) = u^d(t)$ ,  $\in [t_0, t_1)$  pueden escribirse como:

$$(3) \quad \dot{x} = A(t)x + B(t)u ;$$

donde

$$A(t) = \frac{\partial f[y^d(t), u^d(t)]}{\partial y}, \quad B(t) = \frac{\partial f[y^d(t), u^d(t)]}{\partial u};$$

y considerando también el criterio de desempeño

$$(4) \quad \int_{t_0}^{t_1} [x^T(t)Q(t)x(t) + u^T(t)G(t)u(t)] dt + x^T(t_1)Sx(t_1),$$

donde  $S$  y  $Q(t)$  son matrices simétricas definidas no negativas ( $S = S^T \geq 0$ ;  $Q(t) = Q(t)^T \geq 0$ ); y  $G(t)$  es una matriz simétrica definida positiva ( $G(t) = G(t)^T > 0$ ) para  $t_0 \leq t \leq t_1$ . Entonces el problema de determinar un vector de entrada admisible  $u^0(t)$ ,  $t_0 \leq t \leq t_1$ , para el cual es criterio es **mínimo** es llamado **problema del control óptimo lineal** [7].

### 3. ECUACIONES DE MOVIMIENTO DEL ROBOT MÓVIL

El tipo de diseño del robot móvil que se desarrollará en este trabajo será el de *control diferencial*, el cual como ya se mencionó, consta de dos ruedas activas para tracción y dirección, y una rueda pasiva (ver figura 1).

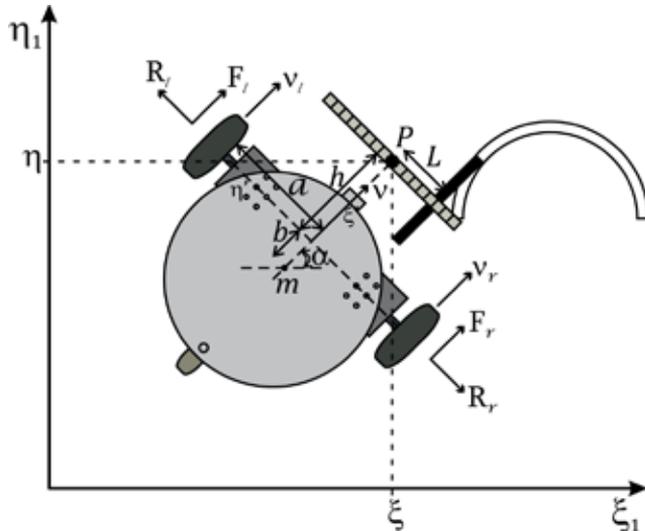


FIGURA 1. Diagrama físico del robot móvil.

Considerando que el robot móvil sólo presenta movimientos planeares translationales y rotacionales, y que no presenta ningún tipo de deslizamientos, se dice entonces que el sistema presenta restricciones no holonómicas (es decir, las velocidades de las ruedas activas tienen sólo la dirección  $P\xi$ ). Utilizando los principios básicos de la mecánica se pueden obtener las siguientes ecuaciones que describen la cinemática del robot móvil dentro del marco de referencia fijo  $\xi_1, \eta_1$ :

$$(5) \quad \dot{\xi} = v \cos(\alpha) - hw \sin(\alpha),$$

$$(6) \quad \dot{\eta} = v \sin(\alpha) + hw \cos(\alpha),$$

$$(7) \quad \dot{\alpha} = w.$$

Utilizando el principio del equilibrio dinámico de D'Alembert [8], y no considerando las fuerzas reactivas  $\vec{R}_r$  y  $\vec{R}_l$ , se obtienen las ecuaciones dinámicas que involucran las fuerzas de empuje  $\vec{F}_r$  y  $\vec{F}_l$ :

$$(8) \quad m(\dot{v} + bw^2) = F_l + F_r ,$$

$$(9) \quad \tilde{J}\dot{w} - mbvw = (F_r - F_l)a .$$

Es necesario además, asociar las fuerzas  $F_r$  y  $F_l$  a las coordenadas  $\varphi_r$  y  $\varphi_l$  que determinan el giro de cada uno de los motores del robot móvil; para esto, el torque de un motor está dado por

$$(10) \quad \tau = F\rho;$$

donde  $\tau$  y  $F$  son el torque y la fuerza del motor; y  $\rho$  es el radio de la rueda acoplada al motor. Es necesario también, un modelo matemático que describa el movimiento de un motor de DC, para esto el modelo más simple conocido es:

$$(11) \quad \tau = \chi u - \sigma\dot{\varphi};$$

donde  $\tau$  es el torque del motor (Nm);  $u$  es el voltaje aplicado a las terminales del motor (Volts);  $\dot{\varphi}$  es la velocidad angular (rad/s);  $\chi$  es la fricción viscosa; y  $\sigma$  es la fuerza contra-electromotriz del motor [9]. Considerando las expresiones para la velocidad angular de cada una de las ruedas

$$(12) \quad \dot{\varphi}_r = \frac{v_r}{\rho} = \frac{v + aw}{\rho},$$

$$(13) \quad \dot{\varphi}_l = \frac{v_l}{\rho} = \frac{v - aw}{\rho};$$

y considerando las ecuaciones (10) y (11), se pueden obtener una expresión de fuerza asociada a las ruedas

$$(14) \quad F_r = \frac{\chi u_r \rho - \sigma(v + aw)}{\rho^2},$$

$$(15) \quad F_l = \frac{\chi u_l \rho - \sigma(v - aw)}{\rho^2};$$

sustituyendo las expresiones (14) y (15) en las ecuaciones de movimiento (8) y (9), se tiene

$$(16) \quad m\dot{v} + mbw^2 = \frac{\chi}{\rho}(u_r + u_l) - \frac{2\sigma v}{\rho^2},$$

$$(17) \quad \tilde{J}\dot{w} - mbvw = \frac{a\chi}{\rho}(u_r - u_l) - \frac{2\sigma a^2 w}{\rho^2}.$$

Por lo tanto, las ecuaciones (5), (6), (7), (16) y (17) describen el movimiento del robot móvil:

$$(18) \quad \begin{aligned} \dot{\alpha} &= w , \\ \dot{\xi} &= v \cos(\alpha) - hw \sin(\alpha) , \\ \dot{\eta} &= v \sin(\alpha) + hw \cos(\alpha) , \\ \dot{v} &= -bw^2 - \frac{2\sigma v}{m\rho^2} + \frac{\chi}{m\rho}(u_r + u_l) , \\ \dot{w} &= \frac{mbvw}{\tilde{J}} - \frac{2\sigma a^2 w}{\tilde{J}\rho^2} + \frac{a\chi}{\tilde{J}\rho}(u_r - u_l) . \end{aligned}$$

## 4. TRAYECTORIAS DESEADAS Y ECUACIONES LINEALES

Ya que cualquier trayectoria irregular puede ser aproximada mediante pequeños segmentos rectos y circulares, se pueden obtener las trayectorias básicas que el robot móvil puede realizar para recrear alguna trayectoria deseada. Considerando que el robot presenta movimientos horizontales y verticales paralelos a los ejes  $0, \eta$  y  $0, \xi$  tenemos en la tabla 1 las siguientes trayectorias rectas (ver figura 2).

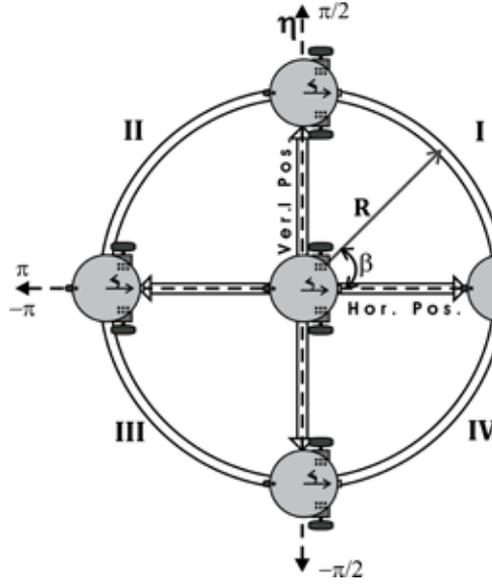


FIGURA 2. Trayectorias básicas deseadas del robot móvil.

Trayectoria	$\xi^d$	$\eta^d$	$\alpha^d$	$\mathbf{v}^d$	$\omega^d$
1. Línea Horizontal sentido Positivo	$\xi_0 + v_0 t$	0	0	$v_0$	0
2. Línea Horizontal sentido Negativo	$\xi_0 + v_0 t$	0	$\pi$	$v_0$	0
3. Línea Vertical sentido Positivo	0	$\eta_0 + v_0 t$	$\frac{\pi}{2}$	$v_0$	0
4. Línea Vertical sentido Negativo	0	$\eta_0 + v_0 t$	$-\frac{\pi}{2}$	$v_0$	0

TABLA 1. Trayectorias rectas: horizontales y verticales.

Para obtener los desplazamientos circulares se supondrá que el estado inicial o posición de casa del robot es paralelo al eje  $\xi$  y que el movimiento se describe sobre un círculo de radio  $R$ , el cual está seccionado en cuatro cuadrantes (ver figura 2) y tomando en cuenta el sentido en que se recorre; las trayectorias se muestran en el tabla 2.

Trayectoria	$\xi^d$	$\eta^d$	$\beta \in$	$\alpha^d$	$\mathbf{v}^d$	$\omega^d$
5. Cuadrante I (dirección reloj)	$R \cos \varphi$	$R \sin \varphi$	$[\frac{\pi}{2}, 0]$	$-\omega_0 t$	$\omega_0 R$	$\omega_0$
6. Cuadrante I (contra reloj)	$R \cos \varphi$	$R \sin \varphi$	$[0, \frac{\pi}{2}]$	$\omega_0 t + \frac{\pi}{2}$	$\omega_0 R$	$\omega_0$
7. Cuadrante II (dirección reloj)	$R \cos \varphi$	$R \sin \varphi$	$[\pi, \frac{\pi}{2}]$	$-\omega_0 t + \frac{\pi}{2}$	$\omega_0 R$	$\omega_0$
8. Cuadrante II (contra reloj)	$R \cos \varphi$	$R \sin \varphi$	$[\frac{\pi}{2}, \pi]$	$\omega_0 t + \pi$	$\omega_0 R$	$\omega_0$
9. Cuadrante III (dirección reloj)	$R \cos \varphi$	$R \sin \varphi$	$[-\frac{\pi}{2}, -\pi]$	$\pi - \omega_0 t$	$\omega_0 R$	$\omega_0$
10. Cuadrante III (contra reloj)	$R \cos \varphi$	$R \sin \varphi$	$[-\pi, -\frac{\pi}{2}]$	$\omega_0 t - \frac{\pi}{2}$	$\omega_0 R$	$\omega_0$
11. Cuadrante IV (dirección reloj)	$R \cos \varphi$	$R \sin \varphi$	$[0, -\frac{\pi}{2}]$	$-\omega_0 t - \frac{\pi}{2}$	$\omega_0 R$	$\omega_0$
12. Cuadrante IV (contra reloj)	$R \cos \varphi$	$R \sin \varphi$	$[-\frac{\pi}{2}, 0]$	$\omega_0 t$	$\omega_0 R$	$\omega_0$

TABLA 2. Trayectorias circulares.

Siendo las ecuaciones de movimiento del robot móvil el conjunto de ecuaciones (18), se pueden aproximar linealmente a vecindades cercanas a las trayectorias deseadas utilizando la expresión (3). Siendo la trayectoria deseada  $y^d = (\xi^d \eta^d \alpha^d v^d w^d)^T$  un movimiento en línea horizontal en sentido positivo (tabla 1, número 1), se tienen las ecuaciones lineales (19)

$$\begin{aligned}
 \dot{\xi} &= \tilde{v}, \\
 \dot{\eta} &= -v_0 \tilde{\alpha} + h \tilde{\omega}, \\
 \dot{\alpha} &= \tilde{\omega}, \\
 \dot{v} &= -\frac{2\sigma}{\rho^2 m} \tilde{v} + \frac{\chi}{\rho} (u_l + u_r), \\
 \dot{\omega} &= \frac{mbv_o}{\tilde{J}} \tilde{\omega} - \frac{2\sigma a^2}{\tilde{J} \rho^2} \tilde{\omega} + \frac{a\chi}{\tilde{J} \rho} (u_l - u_r).
 \end{aligned}
 \tag{19}$$

si la trayectoria deseada  $y^d$  pertenece a alguno de los movimientos circulares de la tabla 2, se pueden obtener las siguientes ecuaciones lineales que dependen de  $\alpha_d$

$$\begin{aligned}
 \dot{\xi} &= -(\omega_0 R \sin(\omega_0 t) + h \omega_0 \cos(\omega_0 t)) \tilde{\alpha} + \cos(\omega_0 t) \tilde{v} - h \sin(\omega_0 t) \tilde{\omega}, \\
 \dot{\eta} &= (\omega_0 R \cos(\omega_0 t) - h \omega_0 \sin(\omega_0 t)) \tilde{\alpha} + \sin(\omega_0 t) \tilde{v} + h \cos(\omega_0 t) \tilde{\omega}, \\
 \dot{\alpha} &= \tilde{\omega}, \\
 \dot{v} &= -\frac{2\sigma}{m \rho^2} \tilde{v} - 2b \omega_0 \tilde{\omega} + \frac{\chi}{m \rho} (u_l + u_r), \\
 \dot{\omega} &= \frac{mb \omega_0}{\tilde{J}} \tilde{v} + \frac{mb \omega_0 R}{\tilde{J}} \tilde{\omega} - \frac{2a^2 \sigma}{\rho^2 \tilde{J}} \tilde{\omega} + \frac{a\chi}{\rho \tilde{J}} (u_l - u_r).
 \end{aligned}
 \tag{20}$$

## 5. DEDUCCIÓN DEL CONTROL ÓPTIMO EN TIEMPO CONTINUO

Consideremos el problema básico de control óptimo en la forma más general:

$$\begin{aligned} \min_{u(t)} J &= \int_{t_0}^{t_1} F(x, u, t) dt + S[x(t_1)] , \\ \text{sujeto a} &: \dot{x}(t) = f(x, u, t), \\ \text{con} &: x(t_0) = x_0, \\ &u(t) \in \Omega(t); \end{aligned}$$

se define la *función valor*  $W^*(x, t)$ , que proporciona el mínimo valor del funcional objetivo del problema de control

$$\begin{aligned} W^*(x, t) &= \min_{u(\tau)} \int_t^{t_1} F(x, u, \tau) dt + S[x(t_1)] , \\ \text{sujeto a} &: \dot{x}(t) = f(x, u, \tau), \text{ para } t \leq \tau \leq t_1, \\ \text{con} &: x(t) = x, \\ &u(\tau) \in \Omega(\tau). \end{aligned}$$

Suponiendo que  $W(x, t)$  es una solución de la ecuación de Hamilton-Jacobi-Bellman, se verifica que:

$$(21) \quad 0 = \min_{u \in \Omega} \{F(x, u, t) + \nabla_x W(x, t) f(x, u, t)\} + W_t(x, t) , \quad \forall x, t,$$

$$(22) \quad \text{con} : W(x, t_1) = S(x), \text{ para todo } x.$$

Suponiendo también que  $u(x, t)$  alcanza el mínimo en la ecuación (21), para todo  $x$  y  $t$ , entonces la trayectoria de control  $u(t)$  para  $t_0 \leq t \leq t_1$ , es óptima.

Para solucionar este problema de control óptimo, se escribe la ecuación de Hamilton-Jacobi-Bellman de acuerdo al indicador de desempeño (4) sujeto al sistema (3) y se establece la suposición de que la solución es de la forma:

$$(23) \quad W(x, t) = x^T P(t)x,$$

donde  $P(t)$  es una matriz simétrica  $n$ -dimensional. Por lo que se puede obtener

$$(24) \quad 0 = \min_{u(t)} \{x^T Q x + u^T G u + 2x^T P(t)[Ax + Bu]\} + x^T \frac{\partial P(t)}{\partial t} x ;$$

derivando parcialmente con respecto a  $u$  la expresión entre llaves e igualando a cero se obtiene:

$$(25) \quad u = -G^{-1} B^T P(t)x.$$

Sustituyendo (25) en la ecuación (24), se deduce:

$$(26) \quad \begin{aligned} -\frac{\partial P(t)}{\partial t} &= Q - P(t) B G^{-1} B^T P(t) + P(t) A + A^T P(t) , \\ \text{con} &: P(t_1) = 0, \end{aligned}$$

que es una *ecuación matricial de Riccati*, donde  $P(t)$  se encuentra resolviendo esta ecuación.

Por lo tanto, el *control óptimo* del problema es:

$$(27) \quad u(t) = -Kx(t), \quad K = G^{-1} B^T P(t);$$

la solución  $P(t) \geq 0$  de la ecuación de Riccati (26) es llamada *solución de estabilización*; y la matriz  $K$  definida en (27) es la *matriz de ganancia del control óptimo*

o la *matriz de estabilización óptima*. Con una apropiada selección de las matrices de peso  $Q$  y  $G$  es posible tener un comportamiento apropiado del sistema en lazo cerrado, ya que las propiedades dinámicas y magnitudes de los estados son de mucha influencia. Se puede verificar el comportamiento con la ecuación lineal de control en lazo cerrado

$$(28) \quad \dot{x}(t) = (A - BK)x(t) .$$

## 6. ASIGNACIÓN DE LOS VALORES CARACTERÍSTICOS EN UN ÁREA ESPECIFICA

Considerando la solución  $P(t, t_1)$  de la ecuación diferencial matricial de Riccati (26) con condición final o de frontera  $P(t_1) = 0$ ; para todo  $t \geq t_0$ , existe el límite

$$\bar{P}(t) = \lim_{t_1 \rightarrow \infty} P(t, t_1),$$

y  $\bar{P}(t)$  es una solución de (26). Además,  $x^T(t)\bar{P}(t)x(t)$  es el índice de comportamiento óptimo, cuando el tiempo inicial es  $t$  y el estado inicial es  $x(t)$ . El control óptimo en un tiempo  $t$  (para un tiempo inicial arbitrario) es definido únicamente por (27) con  $P(t)$  reemplazada por  $\bar{P}(t)$ ; la solución  $\bar{P}(t)$  es llamada una *solución estacionaria* de la ecuación de Riccati (26) con condición  $\bar{P}(\infty) = 0$ .

El resultado puede ser fácilmente aplicado al caso invariante en el tiempo (o constante), cuando  $A, B, Q$  y  $G$  son matrices constantes, además para un horizonte de tiempo infinito la matriz  $S$  del índice de desempeño (4) es usualmente cero. Por lo tanto, el límite  $\bar{P}$  es constante (y también igual al límite de  $P(t_0, t_1)$  cuando  $t$  tiende a  $-\infty$ ). Además, la ecuación diferencial matricial de Riccati se convierte en una ecuación algebraica matricial, ya que el lado izquierdo es idéntico a cero [11]. Por lo que el control óptimo para un sistema lineal invariante en el tiempo es

$$(29) \quad u(t) = -Kx(t), \quad K = G^{-1}B^T\bar{P};$$

siendo  $\bar{P}$  la solución de la ecuación algebraica de Riccati

$$(30) \quad 0 = Q - \bar{P}BG^{-1}B^T\bar{P} + \bar{P}A + A^T\bar{P},$$

sujeto al sistema lineal invariante en el tiempo

$$(31) \quad \dot{x}(t) = Ax(t) + Bu(t).$$

En el diseño de sistemas de control óptimo, la matriz  $K$  se calcula con una apropiada selección de las matrices de peso  $Q$  y  $G$ , y usualmente se determinan a prueba y error hasta lograr el comportamiento deseado. Cuando se trabaja con sistemas grandes este proceso puede no ser muy fácil, por esta razón es común elegir a estas matrices como diagonales.

Los valores característicos del sistema de control el lazo cerrado

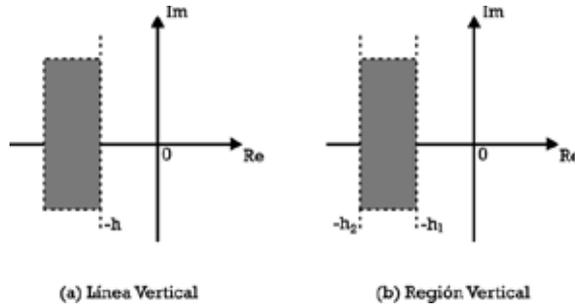
$$(32) \quad \dot{x}(t) = (A - BK)x(t) ,$$

están denotados por

$$(33) \quad \det(I\lambda - (A - BK)) = 0 ,$$

ya que el objetivo de control es mejorar el comportamiento del sistema original, los valores característicos  $\lambda_1, \dots, \lambda_n$  deben ser llevados a una región deseada del semiplano izquierdo del plano complejo (figura 3).

Siendo el  $(A, B)$  las matrices del sistema en lazo abierto (3), y  $h_1 \geq 0$  que representa el grado de estabilidad prescrito correspondiente a una región en el

FIGURA 3. Plano complejo  $S$ .

semiplano izquierdo del plano complejo. Entonces la matriz en lazo cerrado  $A_c = A - BG^{-1}B^T\tilde{P}$  tiene todos sus valores característicos ubicadas al lado izquierdo de la línea vertical  $-h_1$ , donde la matriz  $\tilde{P}$  es la solución de la ecuación de Riccati

$$(34) \quad 0 = Q - \tilde{P}BG^{-1}B^T\tilde{P} + \tilde{P}(A + hI_n) + (A + hI_n)^T\tilde{P}.$$

Con  $Q = 0_n$  los valores característicos inestables de  $A + h_1I_n$  son colocados en la posición reflejada con respecto a la línea vertical  $h_1$ , los cuales son los valores característicos de  $A_c$ .

Asumiendo que  $h_1$  y  $h_2$  son dos valores positivos que determinan la región vertical  $[-h_2, -h_1]$  sobre el eje real negativo, y que dada la matriz de  $n \times n$   $\tilde{A} = A + h_1I_n$ ; la ley de control cambia a

$$(35) \quad u(t) = \rho_g \tilde{K}x(t),$$

$$(36) \quad \tilde{F} = G^{-1}B^T\tilde{F}.$$

La matriz  $\tilde{P}$  es la solución de la ecuación de Riccati modificada

$$(37) \quad 0 = \tilde{P}BG^{-1}B^T\tilde{P} + \tilde{P}\tilde{A} + \tilde{A}^T\tilde{P},$$

donde la matriz  $G$  es seleccionada como matriz diagonal unitaria. La ganancia  $\rho_g$  es calculada con

$$(38) \quad \rho_g = \frac{1}{2} + \frac{h_2 - h_1}{2 \operatorname{tr}(\tilde{A}^*)} = \frac{1}{2} + \frac{h_2 - h_1}{\operatorname{tr}(B\tilde{K})};$$

donde  $\operatorname{tr}(\tilde{A}^*) = \sum_{i=1}^n \lambda_i^+$  y  $\lambda_i^+$  ( $i = 1, 2, \dots, n^+$ ) son los valores característicos de  $\tilde{A}$  el semiplano derecho del plano complejo. El sistema de control óptimo de lazo cerrado, es entonces

$$(39) \quad \dot{x}(t) = (A - \rho_g B\tilde{K})x(t).$$

## 7. SOLUCIÓN Y RESULTADOS

El valor de cada uno de los parámetros del sistema necesarios para obtener los resultados se presenta en el cuadro 3.

El algoritmo de solución para la obtención del control óptimo es el siguiente:

- (1) Obtener el sistema lineal (3) alrededor de la trayectoria deseada.
- (2) Verificar si el par  $(A, B)$  del sistema lineal (3) es completamente controlable, si no, no es posible obtener el control óptimo.

Variable	Valor	Descripción
$v_0$	1.5	Velocidad deseada del robot móvil, [m/s]
$a$	0.40	Distancia entre las ruedas, [m]
$b$	0.40	Distancia del centro de masa al eje de las ruedas, [m]
$h$	0.10	Distancia del eje de las ruedas al arreglo de sensores, [m]
$m$	4.5	Masa del robot, [kg]
$\rho$	0.08	Radio de las ruedas activas, [m]
$R$	0.35	Radio de inercia del carro, [m]
$\chi$	0.01	Fricción viscosa
$\sigma$	0.009	Fuerza contra-electromotriz del motor
$\tilde{J}$	0.28686	Momento de inercia

TABLA 3. Variables del Sistema.

- (3) Desarrollo y solución de la ecuación algebraica matricial de Riccati modificada (37).
- (4) Obtener el control óptimo (35).
- (5) Con la ecuación del sistema de control en lazo cerrado (39) se obtiene el comportamiento del sistema bajo la señal de control óptima.

Siendo el par de matrices  $(A, B)$  del sistema lineal invariante en el tiempo (19) correspondiente a una trayectoria recta horizontal en sentido positivo realizada por el robot móvil

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1,5 & 0 & 0,1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -0,6250 & 0 \\ 0 & 0 & 0 & 0 & 7,8435 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0,0278 & 0,0278 \\ -0,1743 & 0,1743 \end{pmatrix},$$

se puede verificar completa controlabilidad de este sistema lineal con el siguiente teorema

7.1. TEOREMA. Para que el sistema diferencial

$$(40) \quad \dot{x}(t) = Ax(t) + Bu(t),$$

sea completamente controlable, es necesario y suficiente que la **matriz de controlabilidad** de  $n \times nr$  tenga rango  $n$ ; es decir, sea de rango completo, o bien

$$(41) \quad \text{rang} ( B \quad AB \quad A^2B \quad \dots \quad A^{n-1}B ) = n.$$

Además, los valores característicos del sistema en lazo abierto son

$$\det(I\lambda - A) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -0.6250 \\ 7.8435 \end{pmatrix}.$$

Como se mencionó anteriormente, el objetivo de este método de solución es encontrar un algoritmo de control óptimo que haga que los valores característicos que hacen inestable a este sistema, sean asignados a una región vertical en el semiplano

izquierdo del plano complejo. Para este caso, se elegirá la región vertical  $[-20, -5]$ , con la cual se tiene la matriz

$$\tilde{A} = \begin{pmatrix} 5 & 0 & 0 & 1 & 0 \\ 0 & 5 & 1,5 & 0 & 0,1 \\ 0 & 0 & 5 & 0 & 1 \\ 0 & 0 & 0 & 4,3750 & 0 \\ 0 & 0 & 0 & 0 & 12,8435 \end{pmatrix},$$

la cual posee los siguientes valores característicos

$$\det(I\lambda - \tilde{A}) = \begin{pmatrix} 5 \\ 5 \\ 5 \\ 4.3750 \\ 12.8435 \end{pmatrix},$$

y con estos, se obtiene la ganancia  $\rho_g = 0.7328$ . Además, la solución de estabilización de la ecuación algebraica de Riccati modificada (37) es

$$\bar{P} = 10^6 \begin{pmatrix} 0,5695 & 0 & 0 & 0,0607 & 0 \\ 0 & 2,3289 & 0,3122 & 0 & 0,0196 \\ 0 & 0,3122 & 0,0943 & 0 & 0,0056 \\ 0,0607 & 0 & 0 & 0,0121 & 0 \\ 0 & 0,0196 & 0,0056 & 0 & 0,0008 \end{pmatrix}.$$

Así entonces, el control óptimo para que el robot móvil realice la trayectoria deseada esta dado por

$$u(t) = \rho_g \tilde{K} x(t),$$

$$\begin{pmatrix} u_l \\ u_r \end{pmatrix} = \begin{pmatrix} 1.236,6 & -2.500,6 & -710,3 & 247,3 & -96 \\ 1.236,6 & 2.500,6 & 710,3 & 247,3 & 96 \end{pmatrix} \begin{pmatrix} \tilde{\xi} \\ \tilde{\eta} \\ \tilde{\alpha} \\ \tilde{v} \\ \tilde{\omega} \end{pmatrix},$$

por lo que la ecuación lineal de control en lazo cerrado es

$$\dot{x}(t) = (A - \rho_g B \tilde{K}) x(t),$$

$$\begin{pmatrix} \dot{\xi} \\ \dot{\eta} \\ \dot{\alpha} \\ \dot{v} \\ \dot{\omega} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1,5 & 0 & 0,1 \\ 0 & 0 & 0 & 0 & 1 \\ -68,6986 & 0 & 0 & -14,3647 & 0 \\ 0 & -871,6991 & -247,6183 & 0 & -25,6353 \end{pmatrix} \begin{pmatrix} \tilde{\xi} \\ \tilde{\eta} \\ \tilde{\alpha} \\ \tilde{v} \\ \tilde{\omega} \end{pmatrix},$$

la cual tiene los siguientes valores característicos

$$\det(I\lambda - (A - \rho_g B \tilde{K})) = \begin{pmatrix} -7,1824 + 4,1367i \\ -7,1824 - 4,1367i \\ -9,7965 + 10,9741i \\ -9,7965 - 10,9741i \\ -6,0422 \end{pmatrix},$$

con los cuales se puede asegurar que el sistema de control en lazo cerrado es **estable asintóticamente**. Los resultados de las pruebas se muestran en la figura 4, en la

cual puede observarse la comparación del método empleado en este trabajo con el método general, el cual consiste en la solución de la ecuación matricial diferencial de Riccati (26) para un tiempo finito; esta solución se obtiene por métodos numéricos. También se muestra en el método general la importancia de la elección de las matrices de peso, obteniendo un control óptimo lineal con una matriz de peso  $Q$  igual a la identidad y otro control con una matriz de peso elegida a prueba y error hasta encontrar el comportamiento deseado.

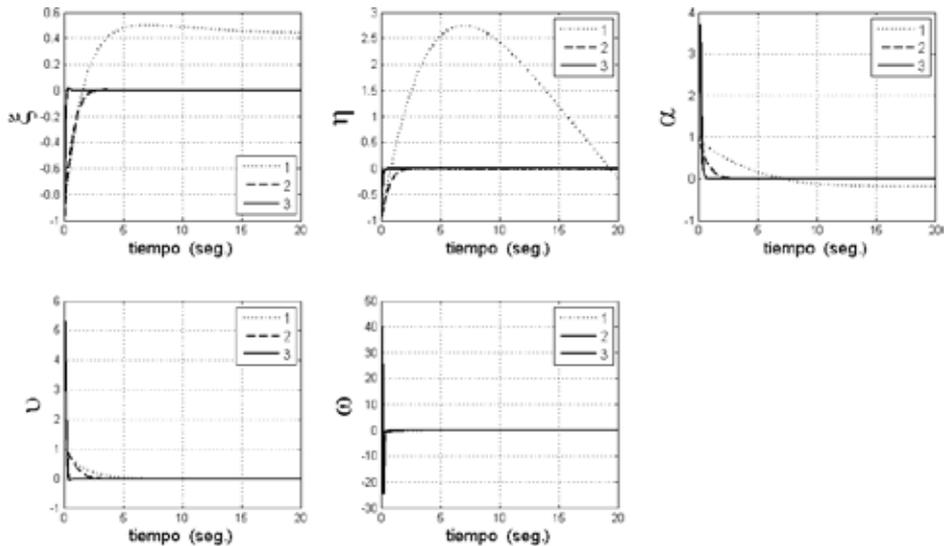


FIGURA 4. 1 Solución en lazo cerrado con matrices de peso  $Q$  unitaria. 2 Solución en lazo cerrado con matriz  $G$  obtenida a prueba y error. 3 Solución en lazo cerrado con método de asignación de valores característicos a la región  $[-5, -20]$ .

## 8. CONCLUSIONES

Las matrices de peso son un elemento de diseño primordial en el desarrollo de un control óptimo; ya que estas determinan que tan estable será el sistema dinámico. El método iterativo de prueba y error puede llevar al sistema a una estabilidad aceptable, como el caso de este trabajo, pero resultó muy difícil llegar a esto; incluso en el caso de elegir otra trayectoria se tendría que repetir este complicado proceso. Por otra parte el método de la asignación de los valores característicos hace que el diseño de un control óptimo sea mucho más sencillo, siendo el único parámetro de diseño la designación del área del semiplano izquierdo en la cual se desean colocar los valores característicos del sistema en cuestión; con este método se obtuvo una mejor respuesta, estabilizando el sistema en un menor tiempo. Cabe mencionar que cualquier región que se elija asegura la estabilidad asintótica del sistema lineal.

## REFERENCIAS

- [1] C. Canudas, B. Siciliano, G. Bastin, B. Brogliato, G. Campion, B. D'Andrea-Novet, A. De Luca, W. Khalil, R. Lozano, R. Ortega, C. Samson y P. Tomei, *Theory of Robot Control*, London, 1996.
- [2] E. Aranda, T. Salgado, M. Velasco, *Control no Lineal Discontinuo de un Robot Móvil*, Computación y Sistemas, Número Especial, 2002, págs. 042-049.
- [3] R. Carrasco, A. Cipriano, *Sistema de Guiado para un Robot Móvil basado en Lógica Difusa*, XV Congreso de la Asociación Chilena de Control Automático, 2002.
- [4] D. P. Tsakiris, C. Samson, P. Rives, *Vision-Based Time-Varying Mobile Robot Control*, Dans Final European Robotics Network (ERNET) Workshop, Germany, 1996.
- [5] T. Y. Wang, C. C. Tsai, J. L. Pang, *Nonlinear Regulation and Path Tracking of a Wheeled Mobile Robot in Polar Coordinates*, Journal of the Chinese Institute of Engineers, Vol. 28, No. 6, 2005, págs. 925-933.
- [6] A. Hemami, M. G. Mehrabi, R. M. H. Cheng, *Synthesis of an Optimal Control Law for Path Tracking in Mobile Robots*, Automática, Vol. 28, No. 2, 1992, págs. 383-387.
- [7] H. Kwakernaak, R. Sivan, *Linear Optimal Control Systems*, Wiley-Interscience, 1972, págs. 24-65.
- [8] F. Beer, R. Johnston, W. Clausen, *Mecánica Vectorial para Ingenieros: Dinámica*, McGraw Hill, 8° Ed., 2007.
- [9] S. Cetinkunt, *Mechatronics*, John Wiley & SONS INC., 2007.
- [10] D. E. Kirk, *Optimal Control Theory*, Prentice Hall, 1970.
- [11] V. Sima, *Algorithms for Linear-Quadratic Optimization*, Taylor & Francis Group, 1996.

Facultad de Cs. de la Electrónica, BUAP.  
Av. San Claudio y 18 Sur, Col. San Manuel  
C.P. 72570, Puebla, Pue., México.  
javier-cm4@ece.buap.mx  
jmgutierrez@ece.buap.mx  
lucio\_y6@hotmail.com mmorin@ece.buap.mx



# RESOLUCIÓN AUTOMÁTICA DE ECUACIONES DIFERENCIALES USANDO EL SISTEMA CHEBOP

MARIO ALBERTO CARBALLO FLORES  
FRANCISCO SERGIO SALEM SILVA  
FACULTAD DE MATEMÁTICAS - UNIVERSIDAD VERACRUZANA

RESUMEN. El problema al que nos enfrentamos en los cálculos simbólicos es que estos crecen exponencialmente, requieren gran espacio para almacenar dicho cálculo y tiempo para realizarlo. La aritmética de punto flotante propone redondear en cada paso y no al final. Se describen los sistemas Chebfun y Chebop, creados por el Grupo de Análisis Numérico de Oxford, que tienen como objetivo extender estas ideas a las funciones y combinar la sensación del simbolismo con la velocidad del cálculo numérico. La idea es representar funciones de manera muy eficiente por interpolación en puntos de Chebyshev cuyo grado se determina automáticamente, adecuándolo para mantener una exactitud cercana a la precisión de la máquina. El sistema Chebop maneja funciones suaves a trozos en intervalos arbitrarios y operadores lineales, incluye los operadores diferencial e integral. En este trabajo se hace una comparación de Chebfun con *Mathematica* y se muestra la eficiencia del sistema Chebop para solucionar ecuaciones diferenciales de forma automática con gran exactitud y rapidez.

## 1. INTRODUCCIÓN

Chebfun es un proyecto desarrollado por el Grupo de Análisis Numérico de Oxford, que forma parte del Laboratorio Computacional de la Universidad de Oxford. Inició como un proyecto doctoral de Zachary Battles, dirigido por Nick Trefethen [1]. En una segunda etapa del proyecto se unen a esta investigación los estudiantes Ricardo Pachón, quien extendió el sistema a las funciones continuas a trozos en intervalos arbitrarios. Rodrigo Platte, colabora junto con Pachón en la detección automática de la subdivisión y de borde [6]. Toby Driscoll de la universidad de Delaware también se involucró en el desarrollo del proyecto. El sistema Chebop se continúa desarrollando para prolongar el alcance a las ecuaciones diferenciales. Uno de los principales colaboradores del Chebop ha sido Folkmar Bornemann de la Universidad Técnica de Múnich [4].

Durante mucho tiempo ha existido dos clases de cálculo matemático: simbólico y numérico. Es bien sabido que los cálculos simbólicos con funciones de variable real crecen exponencialmente, necesitan espacio en la memoria para almacenar dicho cálculo y si no se hace nada para corregir este efecto, los cálculos paran debido a demandas excesivas en tiempo. La aritmética de punto flotante propone la idea alternativa de redondear el cálculo en cada paso, y no sólo al final [8]. El sistema Chebfun, basado en MATLAB, extiende estas ideas a las funciones y combina la sensación del simbolismo con la velocidad del cálculo numérico. La idea es representar funciones de manera muy eficiente por interpolación en puntos de Chebyshev y la formula baricéntrica [1]. El grado del polinomio de aproximación se determina automáticamente, adecuándolo para mantener una exactitud cercano al  $\epsilon$  de la máquina, esto es, cercano a  $10^{-16}$ . Este sistema consiste de aproximadamente ochenta archivos, el software es libre y puede ser transferido

de [www.comlab.ox.ac.uk/oucl/work/nick.trefethen](http://www.comlab.ox.ac.uk/oucl/work/nick.trefethen), se recomienda instalarlo en el lugar de trabajo de MATLAB. En las secciones dos y tres se ofrecen más detalles, ejemplos y comparaciones de Chebfun con el sistema *Mathematica*.

El sistema Chebop también se basa en la idea de aproximar funciones por interpolación en puntos de Chebyshev. Maneja funciones suaves a trozos en intervalos arbitrarios y operadores lineales, incluye los operadores diferencial e integral [4]. La analogía aquí, está entre los operadores lineales y las matrices. Los comandos `diff` y `sum` se usan para definir a los operadores diferencial e integral, respectivamente. Mientras que `*` y `\` para aplicarse a operadores en *modo de avanzar* (multiplicación) y *modo inverso* (división). Los siguientes comandos, por ejemplo, se pueden utilizar para diferenciar a  $f(x) = x^5 + x^3 + x$  en  $[-2, 2]$  usando la notación del Chebop.

```
[d, x]=domain(-2,2); D=diff(d); df=D*(x.^5+x.^3+x);
```

Una de ventajas del sistema Chebop es lo simple de la sintaxis para solucionar ecuaciones diferenciales. Para solucionar un problema de valor en la frontera, solamente definimos al operador que define nuestra ecuación diferencial, condiciones de frontera apropiadas y mecanografiar `\` [4]. La sección cuatro muestra más ejemplos de su eficaz uso en la solución de ecuaciones diferenciales. Los cálculos son realizados en una computadora portátil (memoria RAM de 1024 MB, disco duro de 160 GB y procesador AMD Turion 64X2) con MATLAB R2007a.

## 2. EL SISTEMA CHEBFUN

Chebfun es un sistema basado en MATLAB orientado a objetos y extiende muchas operaciones en vectores y matrices a las funciones y a los operadores. Un vector  $v$  en MATLAB se convierte en una función  $v(x)$  sobre  $[a, b]$  [1]. Los objetos fundamentales son llamados chebfuns, que son manipulados por comandos extendidos de las funciones usuales de MATLAB para los vectores. Un chebfun es una función de una variable definida en un intervalo  $[a, b]$ . Los datos que definen a un chebfun consiste en un sistema de números  $f_0, \dots, f_N$  para algún  $N \geq 0$ , y cada operación se define vía la interpolación polinómica de los valores  $f_j$  en los puntos de Chebyshev  $x_j$  definidos por

$$x_j = \cos(\pi j/N), 0 \leq j \leq N.$$

La interpolación es realizada numéricamente por la fórmula baricéntrica, desarrollada por Salzer [2, 5]

$$p(x) = \sum_{j=0}^N \frac{w_j}{x-x_j} f_j / \sum_{j=0}^N \frac{w_j}{x-x_j}$$

con

$$w_j = (-1)^j \delta_j, \delta_j = \begin{cases} 1/2, & j=0 \text{ o } j=N; \\ 1, & \text{en otro caso.} \end{cases}$$

Este método es conocido por ser numéricamente estable (esto evita que suceda el fenómeno de Runge), esto es, para los puntos de Chebyshev y la fórmula baricéntrica, los polinomios de interpolación tienen características casi ideales, por lo menos para aproximar las funciones que son suaves y que convergen geoméricamente a las funciones analíticas.

Una manera equivalente de representar las funciones es usar una extensión en los polinomios de Chebyshev [7] de la forma

$$p(x) = \sum_{j=0}^N \lambda_j T_j(x),$$

donde

$$T_j(x) = \cos(j \arccos(x)).$$

Al construir un objeto chebfun, el sistema calcula los coeficientes  $\lambda_j$  de la interpolación de la función objetivo  $f$  en los  $N + 1$  puntos de Chebyshev. El proceso de construcción comienza muestreando la función objetivo en  $2n + 1$  puntos, con  $n = 3, 4, \dots$ , entonces el grado óptimo  $N$  se determina tal que  $|\lambda_j|$  está cercano al épsilon de la máquina ( $\text{eps} = 2.2204e-016$ ).

La sintaxis del sistema Chebfun es casi la misma que para MATLAB, los comandos familiares para los vectores son extendidos de manera natural a este sistema. La operación más fundamental que se puede realizar es crear un chebfun llamando el programa constructor `chebfun.m`. Los siguientes comandos construyen un chebfun  $f$  que corresponde a  $f(x) = x^7 + x^5 + x^3 + x$  en el intervalo  $[0, 2]$

```
f=chebfun('x.^7+x.^5+x.^3+x', [0, 2])
f =
chebfun column (1 smooth piece)
interval length values at Chebyshev points
(0, 2)      8      7.1e-015    0.1    0.44 ... 1.7e+002
```

Chebfun necesita 8 puntos de Chebyshev para representar a  $f(x)$  de tal forma que sea exacta, cercana a la precisión de la máquina. Si no se especifica un intervalo, entonces el intervalo por defecto es  $[-1, 1]$ . La grafica del chebfun  $f$  se representa en la Figura 1 y se obtiene mecanografiando `plot(f, '-')`.

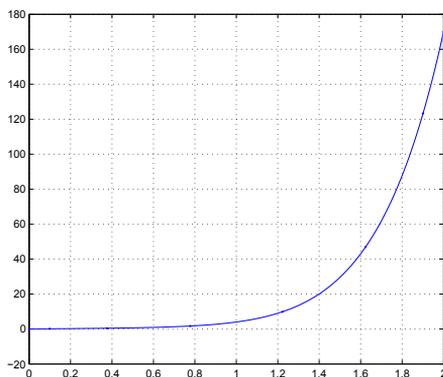


FIGURA 1. chebfun  $f(x) = x^7 + x^5 + x^3 + x$

La función usada para generar un chebfun no necesita ser un polinomio, y el programa constructor chebfun no tiene ningún conocimiento de qué forma puede tener la función. Él evalúa simplemente la función en varios puntos y determina un parámetro  $N$  que sea suficientemente grande [1]. Por ejemplo

```
f=chebfun('cos(x.^2)+sin(5.*x)', [-7, 7])
f =
chebfun column (1 smooth piece)
interval length values at Chebyshev points
(-7, 7)    111    0.73    0.68    0.52 ...    -0.13
```

La gráfica de este chebfun se muestra en la Figura 2. Evidentemente 111 puntos de Chebyshev son suficientes para representar a  $f(x)$  cerca de la precisión de la máquina.

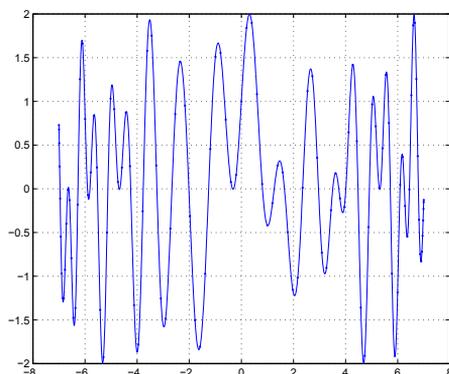


FIGURA 2. chebfun  $f(x) = \cos(x^2) + \sin(5x)$  sobre  $[-7, 7]$

El proceso termina cuando los coeficientes caen a una magnitud cercana a  $10^{-16}$ . La siguiente secuencia de comandos muestra el decaimiento de los coeficientes y gráficamente se muestra en la Figura 3.

```
c=chebpoly(f); % chebpoly regresa los coeficientes
                a_j del polinomio de aproximación
                tal que F_1=a_nT_n(x)+...+a_1T_1(x)
                +a_0T_0(x), donde T_n(x) denota el
                n-ésimo polinomio de Chebyshev y
                F_1 denota el chebfun f.

n=length(f);
semilogy(0:n-1, abs(c(n:-1:1)))
```

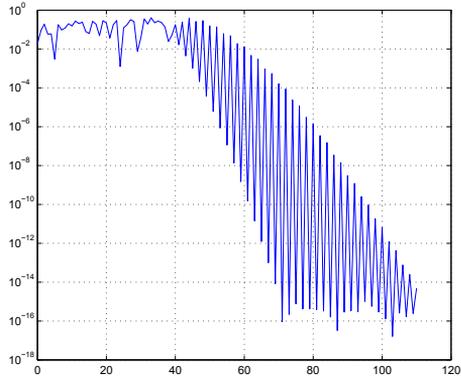


FIGURA 3. Decaimiento de los coeficientes de Chebyshev.

El `chebfun` que acabamos de representar tiene un grado prácticamente pequeño, pero `Chebfun` puede representar funciones con grados más grandes que el anterior, por ejemplo:

```
x=chebfun('x', [-14, 14]);
g=x.^2.*cos(20.*x)+x.^2.*sin(50.*pi.*x.^2);
n=length(g)
n = 31239.
```

Si  $f$  era un vector columna en MATLAB, podríamos evaluarlo en varios índices por comandos como `f(1)` ó `f([1 2 3])`. Para un `chebfun`, la analogía apropiada es que  $f$  se debe evaluar en los puntos correspondientes, no los índices. Por ejemplo, si

```
f=chebfun('x.^3');
format long
f(5)
ans = 1.2499999999999978e+002.
```

Otra forma es

```
f(0:0.5:2)
ans =
0
0.12500000000000000
1.00000000000000000
3.37500000000000000
7.9999999999999989
```

La evaluación del polinomio de interpolación baricéntrica ha calculado los resultados correctamente.

También podemos calcular el valor máximo de la función  $f(x) = \cos(x^2) + \sin(5x)$  en el intervalo  $[-7, 7]$  :

```
max(f)
ans =
    1.995208401814247
```

Este y otros comandos se pueden aplicar a una chebfun, existen más de 100. Para obtener una lista completa se escribe

```
methods chebfun
```

Para descubrir lo que hace un comando, se debe utilizar el comando `help`

```
help chebfun/max
```

Casi todos los comandos de MATLAB se han extendido al sistema Chebfun excepto `domain`, `restrict`, `chebpoly`, `define`, y el comando constructor `chebfun`.

### 3. COMPARACIÓN DE CHEBFUN CON MATHEMATICA

En MATLAB iniciamos con un vector  $v$  y comandos como `sum`, `norm` y `diff` se usan para sumar las entradas, calcular la raíz cuadrada de la suma de cuadrados y diferencias finitas, respectivamente. En el sistema Chebfun tales comandos son extendidos para calcular la integral definida, norma y derivada de una función  $f$ . Para los siguientes cálculos usamos  $f(x) = \cos(x^2) + \sin(5x)$  sobre  $[-7, 7]$ . La integral definida de esta función es

```
sum(f)
ans =
    1.116668663721936
```

Este número se calculó mediante el polinomio de integración (cuadratura de Clenshaw-Curtis), y es interesante compararlo con *Mathematica*

```
In[1]:= NIntegrate [Cos[x^2]+Sin[5*x], {x,-7,7}]
Out[1]:= 1.11667
```

Podemos encontrar la solución a una aproximación numérica de 5 dígitos. *Mathematica* encuentra ceros de una función utilizando el Método de Newton (o método de la tangente), en este caso utilizamos una variante del método de la secante tomando como puntos iniciales  $x = x_0$  y  $x = x_1$ . Observamos que la función  $f(x)$  presenta un cero en el intervalo  $[-7, -6.8]$ , ver Figura 4. Por lo tanto buscamos la raíz en ese intervalo con los siguientes comandos:

```
In[2]:= FindRoot [Cos[x^2]+Sin[5*x], {x,-7,-6.8}]
Out[2]:= {x -> -6.96056}
```

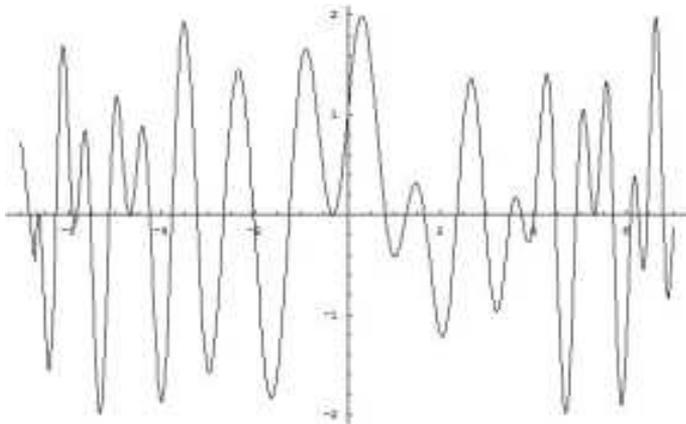


FIGURA 4. Grafica obtenida en Mathematica.

La función presenta dos ceros en el intervalo  $[-6.8, -6.5]$ , para encontrar dichos ceros hacemos lo siguiente: aumentamos el zoom en la grafica y observamos que corta al eje  $x$  en los subintervalos  $[-6.8, -6.64]$  y  $[-6.63, -6.5]$ , ahora ocupamos el comando `FindRoot` para encontrar las raíces.

```
In[3]:=FindRoot[Cos[x^2]+Sin[5*x],{x,-6.8,-6.64}]
Out[3]:={x→-6.65278}
In[4]:=FindRoot[Cos[x^2]+Sin[5*x],{x,-6.63,-6.4}]
Out[4]:={x→-6.62245}
```

Seguimos buscando todos los ceros de  $f(x)$  usando este procedimiento, como se observa es algo tedioso. Cada una de las raíces se obtienen de forma inmediata pero no lo hace automáticamente, para lograr esto se necesitaría hacer un programa específico. En cambio el sistema Chebfun automáticamente obtiene todos los ceros de la función con gran exactitud y rapidez usando el comando extendido `roots`:

```
roots(f)
ans = -6.960560518284815
      -6.652779104113806
      -6.622445944643903
      -6.271307468431964
      -5.905513036013378
      -5.810949860747620
      -5.523058275459005
      -5.121435480552959
      -4.697436590693389
      -4.663146706352830
      -4.246844311963365
      -3.763922322530252
      -3.240517185368686
```

```

-2.664334637524341
-2.015215049269980
-1.255526811776809
-0.336853293647180
-0.296568670137540
0.810949860747619
1.259995556614687
1.652779104113807
2.350645276119854
2.960031584333872
3.507922287205714
3.740004443385308
4.009862941436060
4.475779585355475
4.912468288676978
5.296568670137546
5.324824051556774
5.716511227096384
6.090357504328944
6.255526811776807
6.448599184194149
6.793041088176659

```

Los cálculos se realizan en 0.392401 segundos en la computadora mencionada en la introducción y basados en MATLAB R2007a.

```

tic, roots (f); toc
ans =
Elapsed time is 0.392401 seconds.

```

#### 4. EL SISTEMA CHEBOP

A continuación se describe la sintaxis del sistema Chebop y como resolver ecuaciones diferenciales. En esta parte no se hace una comparación con Mathematica u otro sistema sin embargo es bien sabido que Mathematica tiene problemas para solucionar ecuaciones diferenciales no lineales por ejemplo  $y' = y^3 + t^2$ , en cambio Chebop si lo soluciona [4]. Un objeto `chebop` es un operador que puede ser aplicado a los `chebfuns` [4]. Los operadores `*` y `+` se han extendido para aplicarse a los `chebops`, produciendo otros `chebops`. Iniciamos especificando un dominio

```

[d, x]=domain(0,1);
d = domain
    interval [0,1]
x =
chebfun column (1 smooth piece)
interval length values at Chebyshev points
(0, 1)      2  0      1

```

Al mismo tiempo se genera un `chebfun x` que corresponde a la variable lineal en ese dominio. Un objeto `domain` es un intervalo  $[a, b]$  en qué `chebfuns` y `chebops` pueden ser definidos. En el sistema `Chebfun`, se puede construir una función en este dominio y realizar operaciones como la diferenciación:

```
f=x.^2+x;      % no es necesario llamar al programa
                % constructor chebfun
f2=diff(f,2);  % segunda derivada.
```

En el sistema `Chebop`, podemos construir los operadores diferencial e integral que encapsulen tales cálculos. Por ejemplo, construimos un `chebop` que corresponde al operador  $L : u \mapsto u'' - 2u$  sobre  $[0, 1]$ :

```
domain(0, 1);
D2=diff(d,2);  I=eye(d);
L=D2 - 2*I;
```

Cuando el primer argumento de `diff` es un dominio, la salida es un `chebop` definido para ese dominio, en este caso un `chebop` que corresponde al operador segunda derivada.

También se ha extendido el comando `eye`, cuando su argumento es un dominio, la salida es un `chebop` que corresponde al operador identidad para ese dominio (recordemos que en `MATLAB`, `eye(n)` forma una matriz identidad de tamaño  $n \times n$ ).

Para solucionar ecuaciones diferenciales, debemos imponer condiciones de frontera y después invertir el operador. Por ejemplo, para las condiciones  $u(0) = u(1) = 0$  sólo mecanografiamos `L.bc='dirichlet'`. Pero podemos ocupar diferentes condiciones de frontera, por ejemplo

```
L.lbc=pi/4; L.rbc=sqrt(3); % Dirichlet no homogéneas
L.bc='neumann'; % Neumann homogéneas
L.bc='periodic'; % periódicas.
```

Una forma equivalente para `.bc` es usar el operador `&` [3].

```
L=D2-2*I & 'dirichlet';
```

En `MATLAB`, si queremos resolver un sistema de ecuaciones de la forma  $Ax = b$ , solo tenemos que usar el comando `\` (barra invertida). De forma inteligente, analiza que forma tiene la matriz  $A$  y elige el mejor método para resolver dicho sistema. En el sistema `Chebop`, este comando se ha extendido y se explota para solucionar ecuaciones diferenciales.

Así la solución a la ecuación diferencial  $u'' - 2u = x^2 + x$  sobre  $[0, 1]$  (cuya notación operacional es  $Lu = f$ ) y condiciones de frontera de `Dirichlet` es:

```
u=L\f
```

```

u =
chebfun column (1 smooth piece)
interval length values at Chebyshev points
(0, 1) 14 -2.2e-016 -0.0029 -0.011 ... -6.9e-018

```

Gráficamente la solución se muestra en la Figura 5, y la solución se alcanza con alta exactitud y rapidez.

```

length(u); norm(L*u-f)
ans =
4.250379051367354e-014

```

```

tic, u; toc
Elapsed time is 0.012442 seconds.

```

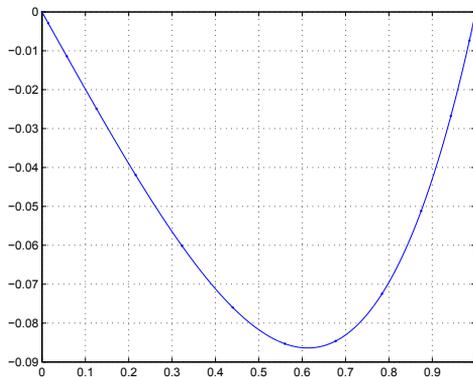


FIGURA 5. Solución de la ecuación diferencial  $u'' - 2u = x^2 + x$  y condiciones de frontera  $u(0) = u(1) = 0$ .

Con el comando `poly` conocemos los coeficientes  $a_j$  del polinomio solución  $p(x) = a_{N-1}x^{N-1} + \dots + a_1x + a_0$  de la ecuación diferencial (en este caso  $N = 14$ ), y los presenta de mayor a menor grado.

```

poly(u)
ans =
0.5000000000000577
0.5000000000013066
0.500000000180329
0.500000003415540
0.500000039668903
0.500000614830069
0.500005712326315
0.500068860953229

```

```

0.500479835408483
0.504131657196605
0.519193416339647
0.536659772718297
0.480320996075747
0.459139090873194.

```

Otra manera de construir un chebop es aplicando el comando extendido `diag` a un `chebfun`. El chebop resultante corresponde al operador multiplicación  $L : u(x) \mapsto g(x)u(x)$ . La siguiente serie de comandos construyen un chebop que corresponde al operador  $L : u \mapsto u'' - x^3u' + u$  y hacen uso de este comando:

```

d=domain(0, 1)
f=x.^3;
F=diag(f); D=diff(d); I=eye(d)
L=D^2-F*D+I;

```

En Matlab, `expm(A)` calcula la exponencial de una matriz  $A$ , y este comando también se ha extendido para calcular la exponencial de un chebop [4]. En conclusión, podemos construir un chebop haciendo uso de las siguientes operaciones:

```

eye(domain)      % operador identidad
zeros(domain)    % op. cero
diff(domain)     % op. diferenciación
cumsum(domain)   % op. integral indefinida
diag(chebfun)    % op. multiplicación
expm(chebop)     % exponencial de un operador
scalar op chebop % op es +, -, *
chebop op scalar % op es +, -, *, /
chebop op chebop % op es +, -, *
chebop^posint    % potencia entera no negativa

```

Cuando se trata de aplicar un chebop hay dos posibilidades, las cuales producen `chebfuns`:

*modo de avanzar:* `chebfun = chebop * chebfun`, por ejemplo `f=L*u`,  
*modo inverso:* `chebfun = chebop \ chebfun`, por ejemplo `u=L\f`.

Para definir a un chebop  $L$  existen diez campos, pero conceptualmente `L.oparray` y `L.varmat` son los más importantes [4]. `L.oparray` es un operador matriz, una función manejable o un conjunto de funciones manejables que determina qué sucede cuando  $L$  se aplica en *modo de avanzar*. Por ejemplo, si mecanografiamos `D=diff(d)`, se establece una función anónima creada con el campo `oparray`, que posteriormente se puede aplicar a un `chebfun` `u:D.oparray = @(u) diff(u)`. Siempre que los chebops se combinen para hacer nuevos chebops, `L.oparray` se restablece para operaciones de este tipo. El sistema Chebop no necesita guardar la pista de los detalles; las dimensiones están determinadas por las funciones anónimas de MATLAB.

`L.varmat` es un varmat, determina qué sucede cuando `L` se aplica en *modo inverso*. Un varmat es una matriz de dimensión indeterminada, o más bien una receta de cómo construir una matriz de dimensión arbitraria. Podemos combinar varmats con las operaciones habituales del álgebra, y las dimensiones están determinadas por las funciones anónimas de MATLAB. Por ejemplo:

```
I=varmat (@eye);
J=3*I+I; J(4)
ans =
    4     0     0     0
    0     4     0     0
    0     0     4     0
    0     0     0     4
```

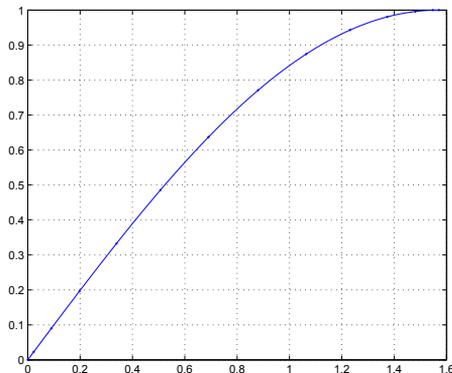
Esto es lo que el sistema Chebop realiza cuando se ejecuta un comando de la forma  $u=L \setminus f$ , donde `L` es un chebop y `f` es un chebfun:

1. El constructor `chebfun` es llamado para construir un chebfun `u` según el procedimiento habitual que implica el muestreo en  $2n+1$  puntos, conforme a la sección 1.
2. Dado un valor de  $N = 2n + 1$ , el varmat asociado a `L` es inmediatamente una matriz  $L_N$  y `f` es evaluada en puntos de Chebyshev para obtener un vector  $f_N$ .
3. En Matlab el comando `\` es entonces llamado para calcular un vector  $u_N = L_N \setminus f_N$ .
4. El proceso continúa sucesivamente sobre rejillas cada vez más finas (como de costumbre) hasta que los coeficientes de Chebyshev obtengan una precisión cercana al  $\epsilon$  de la máquina.

Presentamos un ejemplo más. Solucionar la ecuación diferencial  $u'' + u \cos^2 x = -\sin^3 x$ ,  $x \in [0, \pi/2]$  y condiciones de frontera  $y(0) = 0$ ,  $y(\pi/2) = 1$ .

```
[d,x]=domain(0, pi/2)
f=- (sin(x)).^3; g=(cos(x)).^2;
G=diag(g);
D=diff(d);
L=D^2+G;
L.lbc=0; L.rbc=1;
u=L \ f
u =
chebfun column (1 smooth piece)
interval length values at Chebyshev points
(0, 1.6) 14 1.2e-014 0.023 0.09 ... 1
```

Con `poly` obtenemos los coeficientes del polinomio solución. La solución exacta  $u(x) = \sin(x)$  se muestra en la Figura 6, y se obtiene en 0.000151 segundos con alta

FIGURA 6. Solución  $u(x) = \sin(x)$ .

exactitud.

```
tic, u; toc
    Elapsed time is 0.000151 seconds.

exact=chebfun('sin(x)', d);
error=norm(u-exact, inf)
error =
    1.210143096841421e-014.
```

## 5. CONCLUSIONES

Logramos apreciar que Chebop es fácil de usar y obtenemos soluciones exactas que se expresan generalmente de forma sencilla y fácil de manejar para otros cálculos. Es una tecnología interactiva de gran alcance para solucionar todo tipo de ecuaciones diferenciales. Chebop es capaz de solucionar ecuaciones diferenciales no lineales, parciales y sistemas de ecuaciones diferenciales [4] aunque en este trabajo no se aborda ningún ejemplo, nuestro propósito es ocuparnos de este tema en un trabajo posterior. Esta parte de Chebop que es muy grande y expandible, continúa desarrollándose para explorar múltiples áreas del campo científico, que involucran estos tópicos. El software es libre y se encuentra, junto con ejemplos, en [www.comlab.ox.ac.uk/chebfun/](http://www.comlab.ox.ac.uk/chebfun/).

## REFERENCIAS

- [1] Z. Battles and L. N. Trefethen. *An extension of MATLAB to continuous functions and operators*. SIAM J. Sci. Comput., 25(5):1743-1770, 2004.
- [2] J. P. Berrut and L. N. Trefethen. *Barycentric Lagrange interpolation*. SIAM Rev., 46(3):501-517, 2004.
- [3] T. A. Driscoll. *Chebfun guide 7: linear differential equations with chebops*. 2008 [Octubre 2009]. Disponible en [www.comlab.ox.ac.uk/projects/chebfun/](http://www.comlab.ox.ac.uk/projects/chebfun/).

- [4] T. A. Driscoll, F. Bornemann, and L. N. Trefethen. *The chebop system for automatic solution of differential equations*. BIT Numer. Math., 2008.
- [5] N. J. Higham. *The numerical stability of barycentric Lagrange interpolation*. IMA J. Numer. Anal., 24(4):547-556, 2004.
- [6] R. Pachón, R. B. Platte, and L. N. Trefethen. *Piecewise smooth chebfuns*. IMA J. Numer. Anal., submitted.
- [7] R. B. Platte and L. N. Trefethen. *Chebfun: A New Kind of Numerical Computing*. Oxford University Computing Laboratory, NAG, 2008.
- [8] L. N. Trefethen. *Computing numerically with functions instead of numbers*. Math. in Comp. Sci., 1(2007), pp. 9(199).
- [9] L. N. Trefethen. *Spectral Methods in MATLAB*. SIAM, Philadelphia, PA, 2000.

Universidad Veracruzana, Facultad de Matemáticas.  
Circuito Gonzalo Aguirre Beltrán S/N, Zona Universitaria  
CP. 91090.

malboy15@hotmail.com

fsergios@gmail.com

# IDENTIFICACIÓN DE FUENTES BIOELÉCTRICAS EN EL CEREBRO CONSIDERANDO UNA GEOMETRÍA SIMPLE DE LA CABEZA

JULIO CONDE MONES  
ANDRÉS FRAGUELA COLLAR  
MARÍA MONSERRAT MORÍN CASTILLO  
JACOBO OLIVEROS OLIVEROS  
FCFM - BUAP

RESUMEN. El Problema Inverso Electroencefalográfico (PIE) consiste en hallar fuentes bioeléctricas en el cerebro, a partir de mediciones sobre el cuero cabelludo y es un problema mal planteado, ya que dada una medición sobre el cuero cabelludo, hay diferentes fuentes bioeléctricas que producen esa medición y además porque pequeñas variaciones en las mediciones pueden producir grandes variaciones en la localización de la fuente. El PIE se estudia por medio de un modelo de medio conductor que describe a la cabeza en dos regiones ajenas: la primera representa al cerebro y la segunda al resto de la cabeza, cada una de ellas con conductividad constante y diferente y además el modelo no considera cavidades correspondientes a ojos, nariz y cuello. En este trabajo se valida, por medio de ejemplos sintéticos, un algoritmo estable para el PIE cuando la cabeza se modela por esferas concéntricas.

## 1. INTRODUCCIÓN

El Problema Inverso Electroencefalográfico (PIE) consiste en hallar fuentes bioeléctricas definidas sobre la región que representa al cerebro por medio de mediciones sobre el cuero cabelludo del potencial producido por dichas fuentes (EEG). Las fuentes bioeléctricas pueden estar ubicadas en el volumen o en la corteza cerebral. En este trabajo sólo se consideran las primeras despreciando la actividad de fuentes corticales. Al PIE para el caso de fuentes volumétricas lo llamaremos Problema Inverso Electroencefalográfico Volumétrico y lo denotaremos por PIEV.

El análisis para el PIE para el caso de fuentes volumétricas ha sido estudiado en [1], [4], [6], [7], [9], [13] y para el caso de fuentes corticales en [3].

El PIEV es un problema mal planteado en el sentido de Hadamard, porque dada una medición sobre el cuero cabelludo, podemos encontrar diferentes fuentes que producen dicha medición (no hay unicidad en la solución) y además es sensible a los errores de la medición, es decir, esos errores se amplifican al buscar la solución del problema de identificación. El problema de la unicidad ha sido estudiado en [4] donde se prueba que existe una única fuente en el espacio de las funciones armónicas ortogonales a las constantes (en el producto escalar de las funciones de cuadrado integrables) y se da un algoritmo para encontrar estas funciones armónicas en el caso de dos círculos concéntricos.

En este trabajo se analiza una estrategia de regularización para encontrar la fuente armónica cuando la cabeza se representa por dos esferas concéntricas y el

EEG está dado sobre toda la superficie exterior. Es importante mencionar que las mediciones están dadas sobre un número finito de puntos sobre el cuero cabelludo y pueden interpolarse usando diferentes métodos para ello. Desde este punto de vista el PIEV es un problema mal planteado ya que por cada interpolación recuperamos una única fuente.

El trabajo se divide en la siguiente forma: en la sección 2, se establecen la relación entre las fuentes volumétricas y el EEG sobre el cuero cabelludo, por medio de un problema de contorno y los resultados de existencia unicidad del PIEV. En la sección 3 se da una estrategia de regularización para el PIEV en el caso de mediciones con error considerando a la cabeza compuesta por capas conductoras esféricas. En la sección 4 se dan algunos ejemplos numéricos para validar la estrategia de regularización y las conclusiones en la sección 5.

## 2. PROBLEMA DE CONTORNO ELECTROENCEFALOGRÁFICO

El modelo que se presenta en esta sección ha sido propuesto para estudiar el PIE [1], [4], [6], [7], [9], [13] y en él se considera a la cabeza dividida en dos capas conductoras, a saber,  $\Omega_1$  que representa al cerebro y  $\Omega_2$  que representa a las restantes capas que componen la cabeza. Se supone que la región  $\Omega_i$  tiene una conductividad constante  $\sigma_i$ ,  $i = 1, 2$ , y  $\sigma_1 \neq \sigma_2$ . Mediante  $S_i$  se denotan las superficies que componen las fronteras de las regiones  $\Omega_i$ ,  $i = 1, 2$ :  $\partial\Omega_1 = S_1$ ,  $\partial\Omega_2 = S_1 \cup S_2$  donde  $S_1$  denota la superficie de la corteza cerebral y  $S_2$  la superficie del cuero cabelludo (ver figura 1).

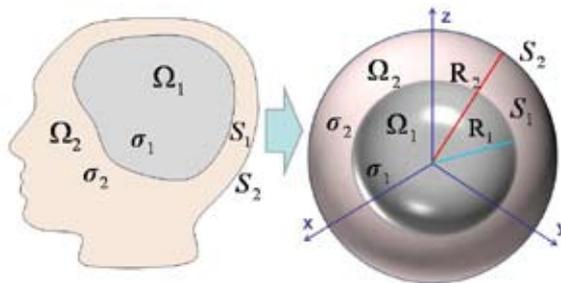


FIGURA 1. Representación esquemática de la cabeza en dos capas conductoras.

Nótese que en este modelo la corteza cerebral y el cuero cabelludo se representan por las superficies cerradas  $S_1$  y  $S_2$ , respectivamente. Este modelo es comúnmente usado como una aproximación de la cabeza [1], [6], [7], [9], [13].

Por otra parte, se sabe que la densidad de corriente total en el cerebro  $\Omega_1$  es:  $\vec{J} = \vec{J}^p + \sigma_1 \vec{E}_1$  [11],[12] donde  $\vec{E}$  es el campo eléctrico generado, y de la ley de Ohm,  $\sigma_1 \vec{E}$  denota la densidad de corriente ohmica. La corriente primaria  $\vec{J}^p$  es debida a las corrientes de difusión a través de la membrana neuronal; la corriente ohmica

$\sigma_1 \vec{E}$  es debida al movimiento de cargas iónicas a través del fluido extracelular del cerebro. La corriente primaria puede estar concentrada tanto en la corteza como en el volumen cerebral. En las restantes regiones que componen la cabeza sólo puede haber corrientes ohmicas. En este trabajo se desprecian las corrientes primarias en la corteza cerebral.

Se sabe que existe un potencial electrostático  $u$  en  $\Omega_0 = \overline{\Omega_1} \cup \Omega_2$  tal que  $\vec{E} = \nabla u$  ([4],[5]) que satisface

$$(1) \quad \Delta u_1 = f \quad \text{en } \Omega_1,$$

$$(2) \quad \Delta u_2 = 0 \quad \text{en } \Omega_2,$$

$$(3) \quad \begin{aligned} u_1 &= u_2 && \text{en } S_1, \\ \sigma_1 \frac{\partial u_1}{\partial n_1} &= \sigma_2 \frac{\partial u_2}{\partial n_1} && \text{en } S_1, \\ \frac{\partial u_2}{\partial n_2} &= 0 && \text{en } S_2, \end{aligned}$$

donde  $f = -\frac{1}{\sigma_1} \nabla \cdot \vec{J}^p$ ,  $u_i = u|_{\Omega_i}$ ,  $i = 1, 2$  y  $\partial u_i / \partial n_j$  denota la derivada normal de  $u_i$  en  $S_j$  con respecto al vector  $n_j$ ,  $i, j = 1, 2$ . El problema (1)-(3) es llamado Problema de Contorno Electroencefalográfico (PCE) [4].

El Problema Directo Electroencefalográfico (PDE) correspondiente al PCE consiste en encontrar

$$(4) \quad u_2 = V \quad \text{in } S_2,$$

cuando  $f$  es conocida.

El PIEV consiste en hallar un par  $(f, u)$ , que satisfaga el problema (1)-(3) y que además  $u$  satisfaga la condición (4).

**2.1. Resultados conocidos del PIEV.** Para analizar el PIEV a través del PCE consideramos los siguientes espacios de funciones:  $L_2(\Omega_i)$ ,  $L_2(S_i)$  y  $L_2(\Omega_0)$  que son espacios de funciones de cuadrado integrables en  $\Omega_i$ ,  $S_i$  y  $\Omega_0$ , respectivamente,  $i = 1, 2$ , donde  $\Omega_0 = \overline{\Omega_1} \cup \Omega_2$ ;  $L_2^\perp(\Omega_i)$  y  $L_2^\perp(S_i)$  son subespacios de  $L_2(\Omega_i)$  y  $L_2(S_i)$ , respectivamente, de funciones ortogonales a las constantes;  $H^1(\Omega_0)$  el espacio de Sobolev de funciones de  $L_2(\Omega_0)$  para las cuales la primera derivada generalizada pertenece a  $L_2(\Omega_0)$ ;  $H^{1,\perp}(\Omega_0)$  el espacio de funciones de  $H^1(\Omega_0)$  ortogonales a las funciones constantes;  $H^{1/2}(S_i)$  el espacio de Banach de funciones de  $L_2(S_i)$ ,  $i = 1, 2$ , los cuales son trazas de funciones de  $H^1(\Omega_0)$ .

Se sabe que la solución del PCE existe y es única, salvo constantes si  $\int_{\Omega_1} f d\Omega_1 = 0$ , es decir, si  $f \in L_2^\perp(\Omega_1)$ . De esto se tiene que, existe una única solución  $u$  que es ortogonal a las constantes en  $L_2(\Omega_0)$  y se cumple que  $\|u\|_{H^1(\Omega_0)} \leq C \|f\|_{L_2(\Omega_1)}$  donde la constante  $C$  no depende de  $f$  [4].

Se demuestra que, dada una medición exacta  $V$ , la solución del PIEV existe y es única si  $f \in H^\perp(\Omega_1) = \{g \in L_2^\perp(\Omega_1) : \Delta g = 0\}$  [4]. Ha sido probado que  $L_2^\perp(\Omega_1)$  se expresa como suma directa de  $H^\perp(\Omega_1)$  y su complemento ortogonal por medio de planteamientos operacionales [4]. De esto se tiene que para una fuente que pertenece a  $L_2^\perp(\Omega_1)$  sólo podemos recuperar a su componente armónica a partir de una medición exacta  $V$ . Es importante notar que en este trabajo sólo se estudia la componente armónica de la divergencia de la fuente primaria  $J^p$ . El problema de determinar a  $J^p$  a partir de  $div(J^p)$  no es abordado en este trabajo.

### 3. SOLUCIÓN DEL PROBLEMA DE IDENTIFICACIÓN EN EL CASO DE ESFERAS CONCÉNTRICAS

En el caso en que  $S_1$  y  $S_2$  corresponden a esferas concéntricas, dada la medición exacta en  $S_2$

$$V = \sum_{n=0}^{\infty} \sum_{m=-n}^n V_{nm} Y_{nm}(\theta, \phi),$$

donde  $V_{nm}$  son los coeficientes de Fourier de  $V$  y  $Y_{nm}(\theta, \phi)$  son los armónicos esféricos, la solución del PIEV está dada por

$$(5) \quad f = \frac{1}{R_1^2} \sum_{n=1}^{\infty} \sum_{m=-n}^n n(2n+3) \left(\frac{r}{R_1}\right)^n \left[ \frac{n+1}{2n+1} \left(\frac{R_1}{R_2}\right)^n \left(\frac{\sigma_2}{\sigma_1} - 1\right) - \left(\frac{R_2}{R_1}\right)^{n+1} \left(\frac{n}{2n+1} + \frac{\sigma_2}{\sigma_1} \frac{n+1}{2n+1}\right) \right] V_{nm} Y_{nm}(\theta, \phi).$$

Para el caso de mediciones con error  $V_\delta$  con  $\|V - V_\delta\|_{L_2(S_2)} \leq \delta$ , suponiendo la condición a priori sobre los coeficientes de Fourier de la medición exacta  $V$

$$(6) \quad \sum_{n=1}^{\infty} \sum_{m=-n}^n n^5 (2n+1) \left(\frac{R_2}{R_1}\right)^{2(n+1)} |V_{nm}|^2 \leq E,$$

la solución aproximada se busca en la forma:

$$(7) \quad f_N = \frac{1}{R_1^2} \sum_{n=1}^N \sum_{m=-n}^n n(2n+3) \left(\frac{r}{R_1}\right)^n \left[ \frac{n+1}{2n+1} \left(\frac{R_1}{R_2}\right)^n \left(\frac{\sigma_2}{\sigma_1} - 1\right) - \left(\frac{R_2}{R_1}\right)^{n+1} \left(\frac{n}{2n+1} + \frac{\sigma_2}{\sigma_1} \frac{n+1}{2n+1}\right) \right] V_{nm}^\delta Y_{nm}(\theta, \phi),$$

eligiendo  $N = N(\delta)$  como se muestra a continuación.

Se tiene que

$$\begin{aligned} \|f - f_{N(\delta)}\|_{L_2(\Omega_1)}^2 &= \int_0^{R_1} r^2 dr \int_0^{2\pi} d\phi \int_0^\pi |f - f_\delta|^2 \sin\theta d\theta \\ &= \frac{1}{R_1} \sum_{n=1}^N \sum_{m=-n}^n n^2 (2n+3) \left[ \frac{n+1}{2n+1} \left(\frac{R_1}{R_2}\right)^n \left(\frac{\sigma_2}{\sigma_1} - 1\right) - \left(\frac{R_2}{R_1}\right)^{n+1} \left(\frac{n}{2n+1} + \frac{\sigma_2}{\sigma_1} \frac{n+1}{2n+1}\right) \right]^2 |V_{nm} - V_{nm}^\delta|^2 \\ &+ \frac{1}{R_1} \sum_{n=N+1}^{\infty} \sum_{m=-n}^n n^2 (2n+3) \left[ \frac{n+1}{2n+1} \left(\frac{R_1}{R_2}\right)^n \left(\frac{\sigma_2}{\sigma_1} - 1\right) - \left(\frac{R_2}{R_1}\right)^{n+1} \left(\frac{n}{2n+1} + \frac{\sigma_2}{\sigma_1} \frac{n+1}{2n+1}\right) \right]^2 |V_{nm}|^2 \end{aligned}$$

$$\begin{aligned} &\leq \frac{1}{R_1} \sum_{n=1}^N \sum_{m=-n}^n n^2(2n+3) \left[ \left( \frac{R_1}{R_2} \right)^n \left( 1 - \frac{\sigma_2}{\sigma_1} \right) + \left( \frac{R_2}{R_1} \right)^{n+1} \times \right. \\ &\quad \left. \left( 1 + \frac{\sigma_2}{\sigma_1} \right) \right]^2 |V_{nm} - V_{nm}^\delta|^2 + \frac{1}{R_1} \sum_{n=N+1}^{\infty} \sum_{m=-n}^n n^2(2n+3) \times \\ &\quad \left[ \left( \frac{R_1}{R_2} \right)^n \left( 1 - \frac{\sigma_2}{\sigma_1} \right) + \left( \frac{R_2}{R_1} \right)^{n+1} \left( 1 + \frac{\sigma_2}{\sigma_1} \right) \right]^2 |V_{nm}|^2. \end{aligned}$$

Asumiendo la condición adicional (6) obtenemos

$$(8) \quad \|f - f_{N(\delta)}\|_{L_2(\Omega_1)}^2 \leq \frac{12}{R_1} \left( 1 + \frac{\sigma_2}{\sigma_1} \right)^2 \left[ \delta^2 N^3 \left( \frac{R_2}{R_1} \right)^{2n+2} + \frac{E}{N} \right].$$

Se define a  $N(\delta)$  como el entero más cercano al punto donde toma el mínimo la ecuación del lado derecho de (8) con respecto de  $N$ . Tal función es convexa y el punto donde toma el mínimo es en la raíz de la ecuación:

$$(9) \quad 3\delta^2 N^2 \left( \frac{R_2}{R_1} \right)^{2N} + 2\delta^2 N^3 \left( \frac{R_2}{R_1} \right)^{2N} \ln \left( \frac{R_2}{R_1} \right) = \frac{E}{N^2}.$$

De (9) obtenemos:

$$(10) \quad \delta^2 N^3 \left( \frac{R_2}{R_1} \right)^{2N+2} = \frac{E}{N \left( 3 + 2N \ln \left( \frac{R_2}{R_1} \right) \right)}.$$

Sustituyendo (10) en (8):

$$(11) \quad \|f - f_{N(\delta)}\|_{L_2(\Omega_1)}^2 \leq \frac{12E}{NR_1} \left( 1 + \frac{\sigma_2}{\sigma_1} \right)^2 \left[ \frac{E}{\left( 3 + 2N \ln \left( \frac{R_2}{R_1} \right) \right)} + 1 \right] \leq \frac{24E}{NR_1} \left( 1 + \frac{\sigma_2}{\sigma_1} \right)^2.$$

De donde obtenemos

$$(12) \quad N^5(\delta) \left( \frac{R_2}{R_1} \right)^{2N(\delta)} \geq \frac{E}{\delta^2 \left( 3 + 2 \ln \frac{R_2}{R_1} \right) \left( \frac{R_2}{R_1} \right)^2},$$

y aplicando logaritmo a ambos lados de la desigualdad

$$(13) \quad N \geq \frac{\ln \left( \frac{E}{\delta^2 \left( 3 + 2 \ln \frac{R_2}{R_1} \right) \left( \frac{R_2}{R_1} \right)^2} \right)}{\left( 5 + 2 \ln \frac{R_2}{R_1} \right)}.$$

Sustituyendo (13) en (11) se obtiene el siguiente resultado:

$$(14) \quad \|f - f_{N(\delta)}\|_{L_2(\Omega_1)} \leq \frac{2\sqrt{6E} \left( \frac{\sigma_2}{\sigma_1} + 1 \right) \sqrt{5 + 2 \ln \frac{R_2}{R_1}}}{R_1^2 \sqrt{\ln \left[ \frac{E}{\left[ 3 + 2 \ln \frac{R_2}{R_1} \right] \left( \frac{R_2}{R_1} \right)^2} + \ln \frac{1}{\delta^2} \right]}}.$$

Finalmente, se obtiene el siguiente teorema el cual da una estrategia de regularización [8] para el PIEV.

3.1. TEOREMA. Supongamos las medición sin error del problema de identificación asociado al problema (1)-(3) satisfacen la condición a priori (6) y que las mediciones con error  $V_\delta$  satisfacen que  $\|V - V\delta\|_{L_2(S_2)} \leq \delta$ . Una aproximación  $f_N$  a la fuente  $f$  es dado en la forma (7). Sea  $N(\delta)$  el entero más cercano a la única raíz de la ecuación (9). Entonces la expresión (7) con  $N = N(\delta)$  es una estrategia de regularización para el calculo estable de  $f_{N(\delta)}$  respecto a la norma en  $L_2(\Omega_1)$ , si se satisface la condición (6). Eligiendo  $N(\delta)$  de esta forma se tiene el estimado (14).

#### 4. EJEMPLOS NUMÉRICOS

En esta sección, se dan algunos ejemplos para validar la estrategia de regularización de la sección anterior (Teorema 3.1).

Para estos ejemplos se toma  $R_1 = 1$ ,  $R_2 = 1,2$ ,  $\sigma_1 = 3$ ,  $\sigma_2 = 1$ ,  $\delta = 0,1$  y las funciones armónicas:

$$\begin{aligned} f(x, y, z) &= (x^2 - y^2)z = r^3 \cos(2\varphi) \sin^2(\theta) \cos(\theta), \\ g(x, y, z) &= x^4 - 6x^2y^2 + y^4 = r^4 \cos(4\varphi) \operatorname{sen}^4(\theta), \\ h(x, y, z) &= (x^4 - 6x^2y^2 + y^4)z = r^5 \cos(4\varphi) \operatorname{sen}^4(\theta) \cos(\theta), \end{aligned}$$

las cuales pertenecen a  $H^\perp(\Omega_1)$ .

Al resolver el PDE tenemos que las mediciones exactas son:

$$(15) \quad V_1 = \frac{-14\sqrt{2}\pi\sigma_1 R_1^9 R_2^3 [Y_{3,2}(\theta, \phi) + Y_{3,-2}(\theta, \phi)]}{27\sqrt{105} [4R_1^7(\sigma_1 - \sigma_2) + R_2^7(3\sigma_1 + 4\sigma_2)]},$$

$$(16) \quad V_2 = \frac{(-144)\sqrt{2}\pi\sigma_1 R_1^9 R_2^3 [Y_{4,4}(\theta, \phi) + Y_{4,-4}(\theta, \phi)]}{(264)\sqrt{35} [5R_1^9(\sigma_1 - \sigma_2) + R_2^9(4\sigma_1 + 5\sigma_2)]},$$

$$(17) \quad V_3 = \frac{(-176)\sqrt{2}\pi\sigma_1 R_1^{13} R_2^5 [Y_{5,4}(\theta, \phi) + Y_{5,-4}(\theta, \phi)]}{(390)\sqrt{385} [6R_1^{11}(\sigma_1 - \sigma_2) + R_2^{11}(5\sigma_1 + 6\sigma_2)]},$$

que corresponden a los potenciales producidos sobre  $S_2$  por las fuentes  $f$ ,  $g$  y  $h$  respectivamente.

Como las mediciones reales tienen error inherente, éste lo simulamos incluyendo un error aleatorio a los coeficientes de Fourier de la medición exacta  $V_i$ ,  $i = 1, 2, 3$ , eligiéndolo en la forma

$$(18) \quad Error = \sum_{n=1}^{\bar{N}} \sum_{m=-n}^n Error_{nm} Y_{nm}(\theta, \phi),$$

donde

$$(19) \quad Error_{nm} = \frac{a_{nm}}{n\sqrt{2n+1}},$$

$$(20) \quad a_{nm} = \frac{NA\sqrt{6\delta} \max |V_i|}{\pi}.$$

En la ecuación (20)  $NA$  es un número aleatorio que está en  $(-1, 1)$ ,  $|Error_{nm}| < \frac{\sqrt{6\delta}}{\pi}$  y  $\bar{N} = 55$ . Tomando en cuenta que  $\sum_{n=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$  se tiene que  $\|Error\|_{L_2(S_2)} \leq \delta = 0,1$ . Se define  $V_\delta^i = V_i + Error$ . En este caso  $\|V_i - V_\delta^i\|_{L_2(S_2)} < \delta$ . De la condición (6) se toma  $E = 25, 50, 385$ , para cada  $V_i$ ,  $i = 1, 2, 3$ . La parte entera de

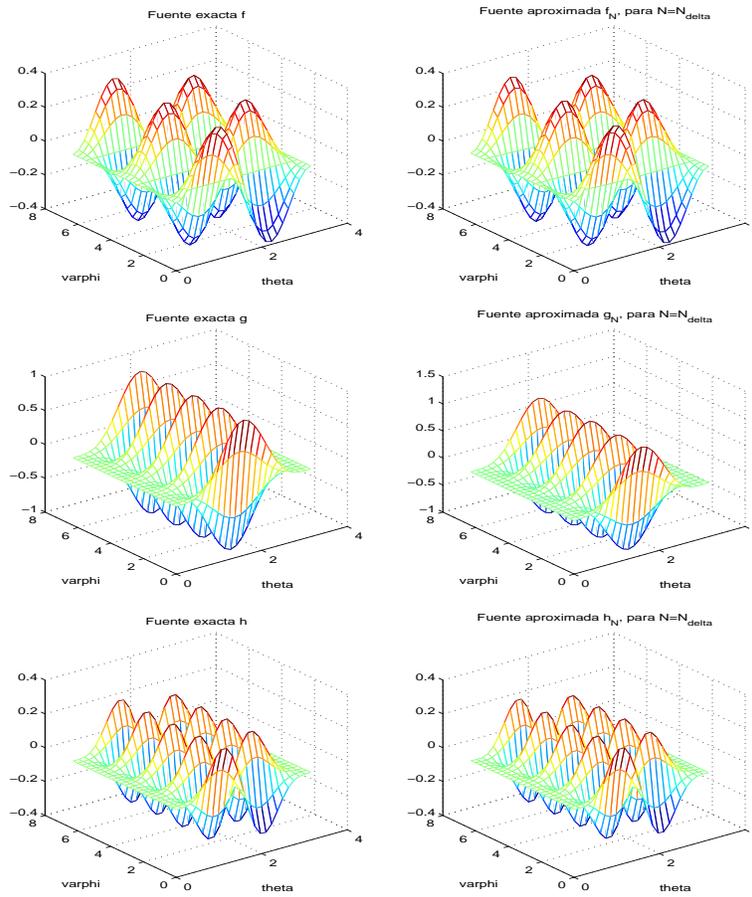


FIGURA 2. Comparación de las fuentes aproximada y exacta para  $r = 1$ .

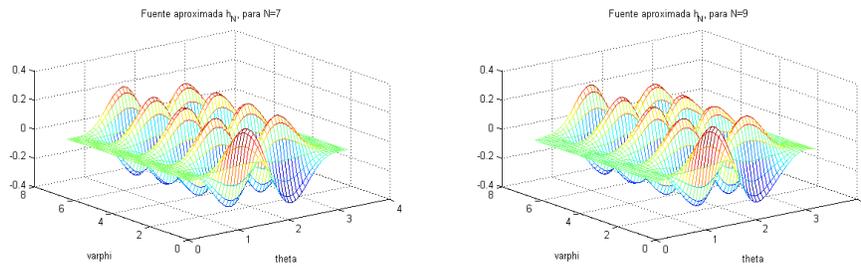


FIGURA 3. Comparación de las fuentes aproximadas para la fuente exacta  $h(r, \theta, \varphi) = r^5 \cos(4\varphi) \text{sen}^4(\theta) \cos(\theta)$  en  $r = 1$ .

la raíz de la ecuación (9) es  $N(\delta) = 3, 4, 5$  para cada caso. Los programas están

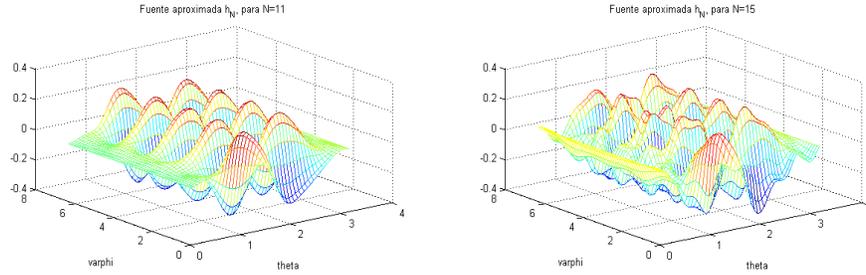


FIGURA 4. Comparación de las fuentes aproximadas para la fuente exacta  $h(r, \theta, \varphi) = r^5 \cos(4\varphi) \text{sen}^4(\theta) \cos(\theta)$  en  $r = 1$ .

hechos en el sistema MATLAB y los errores aleatorios fueron generados por la función rand de MATLAB.

En la figura 2 se comparan las gráficas de las fuentes exactas  $f$ ,  $g$  y  $h$  con las gráficas de las fuentes aproximadas respectivas calculadas por la estrategia de regularización en  $r = 1$ . En la tabla siguiente se muestran los errores relativos para  $N = 3, 4, 5, 7, 9, 11, 15, 30$  y  $55$ . Denotamos al error relativo por  $ER(f, f_N)$  que se define por

$$ER(f, f_N) = \frac{\max_{(r, \theta, \phi) \in \Omega_1} |f - f_N|}{\max_{(r, \theta, \phi) \in \Omega_1} |f|},$$

donde  $f$  es la fuente exacta y  $f_N$  es la fuente aproximada dada en la forma (7).

Tabla I: Error relativo de las aproximaciones.			
$N$	$ER(f, f_N)$	$ER(g, g_N)$	$ER(h, h_N)$
$N(\delta)$	0.0440	0.0297	0.0216
7	0.1349	0.0965	0.0437
9	0.3482	0.1815	0.0728
11	0.3839	0.2258	0.2260
15	1.2156	0.8502	0.8081
30	45.5770	26.1446	12.9389
55	$7.8879 \times 10^3$	$5.1305 \times 10^3$	$2.8147 \times 10^3$

Para cada caso, el error relativo para  $N = N(\delta)$  es el más pequeño.  $N = N(\delta) = 3, 4, 5$  son elegidos usando la estrategia de regularización presentada en la sección 3 para cada  $V_i$ ,  $i = 1, 2, 3$ . Para  $N = 30, 55$  se observa que el mal planteamiento del problema respecto al parámetro  $N$  en todos los casos. En cada caso, podemos considerar que la fuente aproximada para  $N = 7, 9, 11$  es una buena aproximación a la fuente exacta, tal como se ve en las figuras 3 y 4 para el caso de la fuente exacta  $h$  en  $r = 1$ .

## 5. CONCLUSIONES

Los resultados de existencia y unicidad de la solución de PCE conllevan a dar un algoritmo de identificación en el caso de que la cabeza se modela por capas conductoras concéntricas con conductividad constante en cada capa. Este algoritmo se desarrolla suponiendo que conocemos a la medición en toda la superficie  $S_2$ .

En el caso de que el EEG venga dado con error, se da una estrategia de regularización para calcular una fuente aproximada a la fuente correspondiente a mediciones sin error. La estrategia consiste que apartir de información a priori, hallar la única fuente armónica aproximada como una serie truncada, donde el parámetro de regularización es  $N(\delta)$  que se escoge de manera adecuada.

En este trabajo, se ha validado la estrategia de regularización mencionada anteriormente para el Problema Inverso Electroencefalográfico Volumétrico por medio de ejemplos sintéticos.

De los resultados numéricos presentados en este trabajo se observa que cuando se tiene la medición con error debemos aproximar con un polinomio de grado a los más  $N = 7$  (lo que da 48 términos) dando la condición a priori (6) de suavidad de la medición exacta, el cual puede ser proporcionado por los especialistas en el área.

Entre las perspectivas que se tienen en esta línea de investigación se encuentran:

1. Desarrollar este algoritmo estable para una geometría más general.
2. Encontrar la fuente con soporte compacto dentro de  $\Omega_1$  que esté más cercana a la fuente armónica  $f$  encontrada por este algoritmo. Este problema lo podemos ver como un problema de minimización:

$$\min_{g \in F} \|f - g\|_{L_2^\perp(\Omega_1)},$$

donde  $F$  es una clase de funciones de  $L_2^\perp(\Omega_1)$  que tienen soporte compacto en  $\Omega_1$ .

3. Hallar la componente no armónica de la fuente.
4. La interpolación estable de los datos electroencefalográficos.

#### REFERENCIAS

- [1] A. Amir, *Uniqueness of the generators of brain evoked potential maps*, IEEE Transactions on Biomedical Engineering, 41 (1994), 1-11.
- [2] A. El Badiá and T. Ha-Duong, *Some remarks on the problem of source identification from boundary measurements*. Inverse Problems, 14 No. 4 (1998), 883-891.
- [3] A. Fraguera, J. Oliveros, M. Morín and L. Cervantes, *Inverse Electroencephalography for Cortical Sources*. Applied Numerical Mathematics, 55 (2005), 191-203.
- [4] A. Fraguera, M. Morín and J. Oliveros, *Inverse Electroencephalography for Volumetric Sources*, Mathematics and Computers in Simulation, 71 (4) (2007) 481-492.
- [5] M. Morín, J. Conde, J. Oliveros, *A regularization strategy for the identification of the bioelectrical source problem for two-concentric-sphere geometry*, Sometido a la Revista Mexicana de Física.
- [6] A. Fraguera, M. Morín, J. Oliveros, *Modelos Matemáticos en Electroencefalografía Inversa*, Tópicos de la Teoría de la Aproximación II, (Textos Científicos. Benemérita Universidad Autónoma de Puebla), (2007) 73-95.
- [7] L. Heller, *Return Current in Encephalography Variational Principles*, Biophysical Journal, 57 (1990) 601-607.
- [8] A. Kirsch, *An introduction to the mathematical theory of inverse problems*, Springer Verlag, 1996.
- [9] Yu. M. Koptelov and E. V. Zakharov, *Inverse Problems in Electroencephalography and Their Numerical Solving*, III-posed Problems in Natural Sciences, VSP, Utrecht, 92C55 (1992) 543-552.
- [10] V. P. Mijailov, *Partial Differential Equations*, Editorial Mir, 1982.
- [11] P. L. Nunez, *Electric Field of the Brain*, N.Y. Oxford Univ. Press, 1981.
- [12] R. Plonsey and D. G. Fleming, *Bioelectric Phenomena*, N.Y. Mc Graw-Hill, 1969.
- [13] J. Sarvas, *Basic Mathematical and Electromagnetic Concepts of the Biomagnetic inverse Problem*, Phys. Med. Biol., 32 (1) (1987), 11-22.

Facultad de Ciencias Físico Matemáticas de la BUAP  
Postgrado en Ciencias Matemáticas  
Av. San Claudio y 18 sur, 72570 Puebla, Pue.  
juliocondem@yahoo.com.mx  
fraguela@fcfm.buap.mx  
mmorin@ece.buap.mx  
oliveros@fcfm.buap.mx

# DISEÑO DE UN CONTROL ÓPTIMO DIGITAL PARA UN ROBOT MÓVIL.

JOSÉ ELIGIO MOISÉS GUTIÉRREZ ARIAS  
MARÍA MONSERRAT MORÍN CASTILLO  
GELACIO SALAS ORTEGA  
FACULTAD DE CIENCIAS DE LA ELECTRÓNICA - BUAP

RESUMEN. En la Teoría de Control, el concepto de "optimización" es muy importante. La optimización se caracteriza por ser una rama de las matemáticas cuyo objetivo es mejorar una variable con vistas a maximizar un beneficio (o minimizar un coste). Esto se puede aplicar a muchas situaciones prácticas distintas (la variable puede ser una temperatura, un campo de velocidades, una medida de la información, etc.).

En la práctica, un problema de optimización es aquél en el que se desea conducir la solución del sistema a un estado objetivo  $y_d$  y para ello se minimiza la distancia entre  $y$  e  $y_d$ . Así, con este planteamiento, un problema de control se reduce al cálculo de puntos extremos con restricciones (y esto explica la íntima conexión de la Teoría de Control con la Optimización a la que nos referimos)

## 1. INTRODUCCIÓN

La idea de control puede ser expresada como el proceso mediante el cual se ejerce una influencia sobre el comportamiento de un sistema dinámico (que varía con el tiempo) para alcanzar un propósito previamente fijado. Una clase importante de modelos de sistemas dinámicos controlados, son los que se presentan en la mecatrónica y en el área de la robótica a estos sistemas se les puede representar por una ecuación diferencial [2] en  $\mathbb{R}^n$

$$(1) \quad \dot{x}(t) = f(x(t), u(t)), \quad x(t_0) = x_0,$$

donde la dinámica  $f$  es una función que satisface condiciones adecuadas y el control  $u(\cdot)$  pertenece a una familia especial  $\mathcal{U}$  de funciones con valores en un subconjunto  $U$  de  $\mathbb{R}^n$ . Una vez elegido un control  $u \in \mathcal{U}$ , el sistema anterior determina una trayectoria o estado  $x(\cdot)$  con condición inicial  $x_0$  en el momento  $t_0$ .

1.1. EJEMPLO. Si se desea controlar la trayectoria de un avión, con condición inicial  $x(t_0)$ , para lograr una condición final  $x(t_f)$ , el estado del sistema  $x(\cdot)$  podría representar la posición y velocidad del avión y el control  $u(\cdot)$  representaría la fuerza o aceleración necesaria para lograr tal objetivo. Con esta formulación, este ejemplo representa un problema para la Teoría de Control, la cual hace énfasis en el análisis sobre las condiciones necesarias y suficientes para la existencia de los controles adecuados, así como también de la existencia, unicidad, y estabilidad de la trayectoria que garantice el logro de dicho objetivo. Ahora bien, si además se desea lograr tal propósito en un tiempo mínimo, o con mínimo uso de combustible, entonces este

es un problema de control óptimo. En tal caso, se quiere minimizar un funcional que depende del estado del sistema y del control llamado funcional de costo.

$$(2) \quad J^{u(\cdot)}(t_0, x_0) = \ell(x(t_f)) + \int_{t_0}^{t_f} L(x(t), u(t)) dt,$$

donde  $L$  y  $\ell$  son funciones que satisfacen condiciones adecuadas. La función  $L$  representa el costo incurrido por el desplazamiento  $x(\cdot)$  y por la fuerza realizada  $u(\cdot)$ , mientras que la función  $\ell$  representa la penalización por la desviación del estado  $x(t_f)$  en el instante final  $t_f$  de un estado deseado  $x_f$ . En nuestro ejemplo, si queremos minimizar la cantidad de tiempo transcurrido  $t_f$ , debemos tomar  $\ell = 0$ ,  $L = 1$ . Por otro lado, si deseamos minimizar el uso de combustible, podemos tomar  $L(x, u) = u^2$ .

Si el control  $u^*$  es tal que minimiza al funcional de costo, es decir, si

$$(3) \quad J^{u^*(\cdot)}(t_0, x_0) \leq J^{u(\cdot)}(t_0, x_0), \quad \forall u(\cdot) \in \mathcal{U},$$

entonces  $u^*$  se denomina *control óptimo*.

La teoría de control óptimo hace énfasis en el estudio de condiciones necesarias y suficientes para la existencia y unicidad del control óptimo, así como también del desarrollo de metodologías para su determinación. Una de estas metodologías es el Principio del Mínimo (o Máximo) de Pontryagin que presentamos en este trabajo.

## 2. NOMENCLATURA

Variable	Descripción
$v_o$	Velocidad en milisegundos
$a$	Distancia entre ruedas [mts]
$b$	Distancia del centro de masa al eje de las ruedas [mts]
$h$	Distancia del eje de las ruedas al arreglo de sensores infrarrojos [mts]
$m$	Masa del robot [kg]
$\rho$	Radio de las ruedas [mts]
$R$	Radio de inercia del carro [mts]
$\chi$	Fricción viscosa
$\sigma$	Fuerza contraelectromotriz del motor
$\tilde{J}$	Momento de inercia
$\omega$	Velocidad angular
$\alpha$	Ángulo que forma el eje de simetría del móvil $x$ y el eje $\xi$
$F_{r,l}$	Fuerzas activas
$R_{r,l}$	Fuerzas reactivas
$M$	Torque de los motores
$P$	Punto donde se coloca un arreglo de sensores
$T$	Tiempo de muestreo en milisegundos

TABLA 1. Parámetros del robot y variables utilizadas

## 3. PLANTEAMIENTO DEL PROBLEMA

Considerando el siguiente proceso controlable [5]

$$(4) \quad \begin{aligned} \dot{y} &= f(y, u), \\ u(\cdot) &\in U = \{u : u(t) \in \Omega \subseteq \mathbb{R}^r\}, \end{aligned}$$

donde  $y$  es el vector  $\mathbf{n}$ -dimensional que contiene las coordenadas de estado del sistema,  $u$  es un vector  $\mathbf{r}$ -dimensional que representa los controles de entrada. El control es una función vectorial continua a trozos, la cual en cada instante de tiempo, toma sus valores en un conjunto  $\Omega$  convexo, cerrado y acotado. Suponemos que dado algún movimiento  $y^d(t)$  y un control  $u^d(t)$  deseado; satisfacen las siguientes ecuaciones

$$(5) \quad \begin{cases} \dot{y}^d = f(y^d(t), u^d(t)), \\ u(\cdot) \in U, \quad t \in [t_0, t_1]. \end{cases}$$

Se tiene un arreglo de sensores que nos dan información sobre el movimiento que realiza el móvil. Después de procesar dicha información se pueden estimar las desviaciones que ocurren  $x(t) = y(t) - y^d(t)$  para así poder ejercer el control sobre los motores o actuadores.

Dadas las siguientes notaciones:

$$\begin{aligned} \Delta u &= u - u^d \quad \text{control adicional,} \\ x &= y - y^d \quad \text{desviación respecto al movimiento deseado,} \\ \tilde{z} &= \varphi(y) - \varphi(y^d) \quad \text{vector de la información que se recibe sobre la desviación.} \end{aligned}$$

Las ecuaciones diferenciales que gobiernan las desviaciones  $x(t) = y(t) - y^d(t)$  para algún movimiento deseado  $y(t) = y^d(t)$  y un control deseado  $u(t) = u^d(t)$ , pueden escribirse

$$(6) \quad \dot{x} = A(t)x + B(t)\Delta u,$$

con

$$A(t) = \frac{\partial f[y^d(t), u^d(t)]}{\partial y}, \quad B(t) = \frac{\partial f[y^d(t), u^d(t)]}{\partial u}.$$

El problema de diseño puede plantearse de la siguiente manera: Determinar el control óptimo  $u^0(t)$  sobre  $[0, N]$  tal que el índice de desempeño

$$(7) \quad J = G[x(N), N] + \sum_{k=0}^{N-1} F[x(k), u(k), k],$$

sea mínimo, sujeto a la restricción de igualdad,

$$(8) \quad x(k+1) = f[x(k), u(k), k].$$

El término  $G[x(N), N]$  que aparece en la ecuación (7) es el costo final del índice de desempeño, y se requiere como restricción final sobre la condición en el extremo sólo si  $x(N)$  no es fijo.

## 4. ECUACIONES DINÁMICAS DEL ROBOT MÓVIL

Considerando la clase de robots móviles autónomos que consisten de tres ruedas, dos activas y una pasiva, con restricciones no-holónicas, que aparecen como consecuencia de la hipótesis de no deslizamiento. Las velocidades del centro de las ruedas son denotadas como  $v_r$  y  $v_l$ , ver Figura 1.

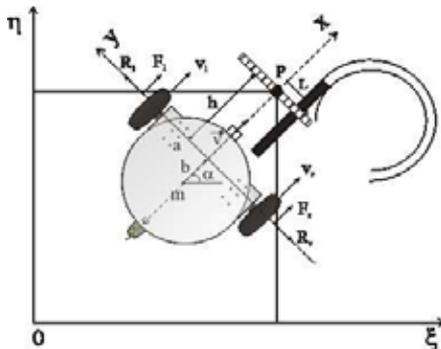


FIGURA 1. Sistema de coordenadas fijo  $\xi\eta$  y sistema de referencia relativo al punto  $P$ .

La posición del robot con respecto al sistema de referencia inercial  $\xi\eta$ , están dadas por las siguientes relaciones

$$(9) \quad \begin{aligned} \dot{\alpha} &= w, \\ \dot{\xi} &= v \cos \alpha - h\omega \sin \alpha, \\ \dot{\eta} &= v \sin \alpha + h\omega \cos \alpha. \end{aligned}$$

Para obtener las ecuaciones dinámicas, se consideran principalmente fuerzas activas  $\bar{F}_r, \bar{F}_l$ . Se deducen las ecuaciones del sistema relativo al punto  $P_{xy}$  mediante los teoremas principales de la mecánica, obteniendo las siguientes relaciones dinámicas

$$(10) \quad \begin{aligned} m(\dot{v} + b\omega^2) &= F_r + F_l, \\ \tilde{J}\dot{\omega} + mb\omega v &= (F_r - F_l)a. \end{aligned}$$

Sustituyendo las fuerzas activas por los torques de los motores y los voltajes que son aplicados a los mismos, hallamos:

$$(11) \quad M = F\rho,$$

donde  $M$  es el torque del motor, y el modelo más simple del motor es

$$M = \chi u - \sigma\dot{\varphi},$$

donde el miembro derecho  $\sigma\dot{\varphi}$  es la suma de la fricción viscosa y la fuerza contra-electromotriz. Entonces para la rueda derecha

$$(12) \quad \begin{aligned} F_r &= \frac{\chi u_r - \sigma\dot{\varphi}_r}{\rho}, \\ \varphi_r &= \frac{v_r}{\rho} = \frac{v + wa}{\rho}. \end{aligned}$$

Sustituyendo en la ecuación para cada rueda, y realizando operaciones en (10) se obtienen las ecuaciones dinámicas:

$$(13) \quad \begin{aligned} m\dot{v} + mbw^2 + \frac{2\sigma}{\rho^2}v &= \frac{\chi}{\rho}(u_r + u_l), \\ \tilde{J}\dot{w} + mbwv + \frac{2\sigma a}{\rho^2}w &= \frac{\chi}{\rho}(u_r - u_l). \end{aligned}$$

Finalmente se obtienen las ecuaciones del movimiento:

$$(14) \quad \begin{cases} \dot{\xi}_c = v \cos \alpha - b \omega \sin \alpha, \\ \dot{\eta}_c = v \sin \alpha + b \omega \cos \alpha, \\ \dot{\alpha} = \omega, \\ m \dot{v} = -mb\omega^2 - \frac{2\sigma}{\rho^2}v + \frac{\chi}{\rho}(u_r + u_l), \\ \tilde{J} \dot{w} = mb\omega v - \frac{2a^2\omega\sigma}{\rho^2} + \frac{\chi a}{\rho}(u_r - u_l), \end{cases}$$

que permiten simular la dinámica de este robot.

## 5. TRAYECTORIAS PROGRAMADAS Y ECUACIONES EN DESVIACIONES

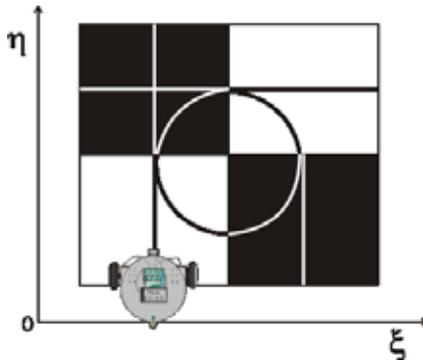


FIGURA 2. Tablero que muestra posibles trayectorias y combinaciones de las mismas.

Una trayectoria deseada se puede presentar considerando la configuración de líneas como se representa en la Figura 2, se observa que el robot puede realizar su actividad por una línea recta horizontal o vertical en sus dos sentidos y por un semicírculo; en cuyo caso, se tienen ocho posibles configuraciones.

La tabla 2 muestra el conjunto de trayectorias deseadas, cuando el movimiento se realiza a lo largo de un segmento de línea paralela al eje  $0\xi$  o al eje  $0\eta$ .

En la Figura 2 se muestra un tablero, donde tenemos cuatro cuadrantes contados en dirección contraria a las manecillas del reloj. Se tienen ocho posibles configuraciones, de las cuales únicamente describimos la correspondiente al cuadrante I (sentido del reloj).

Trayectoria	$\xi^d$	$\eta^d$	$\alpha^d$	$v^d$	$\omega^d$
1. Línea horizontal sentido positivo	$v_0 t + \xi_0$	0	0	$v_0$	0
2. Línea horizontal sentido negativo	$v_0 t + \xi_0$	0	$\pi$	$v_0$	0
3. Línea vertical sentido negativo	0	$v_0 t + \eta_0$	$\frac{\pi}{2}$	$v_0$	0
4. Línea vertical sentido positivo	0	$v_0 t + \eta_0$	$-\frac{\pi}{2}$	$v_0$	0

TABLA 2. Trayectorias programadas de las líneas horizontal y vertical

5.1. **Ecuaciones lineales en desviaciones.** Si  $u^d(t)$  es una entrada nominal al sistema descrito por las ecuaciones (14) y  $y^d$  es una trayectoria nominal de dicho sistema, entonces el sistema de ecuaciones lineales en desviaciones, considerando el sistema lineal para la línea horizontal en (*sentido positivo*) como trayectoria deseada, esta dada por:

$$(15) \quad \begin{cases} \dot{\xi} = -v, \\ \dot{\eta} = -v_0 \alpha - h \omega, \\ \dot{\alpha} = \omega, \\ \dot{v} = -\frac{2\sigma}{\rho^2 m} v + \frac{\chi}{\rho} u_1, \\ \dot{\omega} = -\frac{mbv_o}{\tilde{J}} \omega - \frac{2a^2\sigma}{\rho^2 \tilde{J}} \omega + \frac{\chi a}{\tilde{J} \rho} u_2. \end{cases}$$

Si la trayectoria deseada es un semicírculo, se obtienen ocho sistemas en desviaciones lineales, al sustituir cada uno de los movimientos deseados (*en este caso, cada sistema depende de  $\alpha^d$* ). El siguiente sistema corresponde a la trayectoria deseada descrita al inicio de esta sección:

$$\begin{aligned} \dot{\xi} &= -(\omega_0 R \sin \alpha^d + h \omega_0 \cos \alpha^d) \alpha + v \cos \alpha^d - \omega h \sin \alpha^d, \\ \dot{\eta} &= (-\omega_0 R \cos \alpha^d - h \omega_0 \sin \alpha^d) \alpha + v \sin \alpha^d + \omega h \cos \alpha^d, \\ \dot{\alpha} &= \omega, \\ \dot{v} &= -\frac{2\sigma}{\rho^2} v - 2b\omega_0 \omega + \frac{\chi}{m\rho} (u_r + u_l), \\ \dot{\omega} &= \frac{mb\omega_o}{\tilde{J}} v + \frac{mb\omega_o R}{\tilde{J}} \omega - \frac{2a^2\sigma}{\rho^2 \tilde{J}} \omega + \frac{\chi a}{\rho} (u_r - u_l). \end{aligned}$$

Los sistemas en desviaciones para las trayectorias restantes son análogos.

5.2. **Reducción del sistema de ecuaciones para una trayectoria deseada y sistema lineal.** En este trabajo analizaremos el caso para una trayectoria de una línea horizontal ( $0\xi$ ) en sentido positivo (*caso 2 del cuadro 2*). Por tanto el sistema de ecuaciones se reduce a un sistema de solo tres ecuaciones no lineales.

$$(16) \quad \begin{cases} \dot{\eta} = v \sin \alpha + h \omega \cos \alpha, \\ \dot{\alpha} = \omega, \\ \tilde{J} \dot{\omega} = mb\omega v - \frac{2a^2\omega\sigma}{\rho^2} + \frac{\chi a}{\rho} (u_r - u_l), \end{cases}$$

donde  $\eta, \alpha, \omega$  son las únicas variables que intervienen en esta trayectoria.

La linealización de nuestro sistema de ecuaciones (16) la hacemos alrededor de la trayectoria deseada [1], [3],

$$\begin{pmatrix} \eta \\ \alpha \\ \omega \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Aplicamos el jacobiano a nuestro sistema de ecuaciones y evaluamos, para obtener nuestras variables de estado:

$$A = \left. \frac{\partial f(x, u)}{\partial x} \right|_{(x^0, u^0)}$$

$$B = \left. \frac{\partial f(x, u)}{\partial u} \right|_{(x^0, u^0)}$$

siendo  $f = (f_1, f_2, f_3)$  donde,

$$\begin{cases} f_1 = v \sin \alpha + h \omega \cos \alpha, \\ f_2 = \omega, \\ f_3 = [mbv - \frac{2a^2\sigma}{\rho^2}] \frac{\omega}{\tilde{J}} + \frac{\chi a}{\tilde{J} \rho} u, \end{cases}$$

y por lo tanto nuestras matrices jacobianas quedan de la siguiente forma:

$$A = \begin{pmatrix} 0 & v_0 \cos \alpha - h \omega \sin \alpha & h \cos \alpha \\ 0 & 0 & 1 \\ 0 & 0 & [mbv - \frac{2a^2\sigma}{\rho^2}] \frac{1}{\tilde{J}} \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ \frac{\chi a}{\tilde{J} \rho} \end{pmatrix}.$$

Sustituyendo los valores numéricos de los parámetros mostrados en la tabla 3 y considerando la trayectoria deseada,

Variable	Valor
$v_o$	1.5
$a$	0.40
$b$	0.05
$h$	0.10
$m$	4.5
$\rho$	0.08
$R$	0.35
$\chi$	0.01
$\sigma$	0.009
$\tilde{J}$	0.2868

TABLA 3. Valores de los parámetros del robot.

tenemos que las matrices A y B son:

$$A = \begin{pmatrix} 0 & 1.5 & 0.10 \\ 0 & 0 & 1 \\ 0 & 0 & -0.3922 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ 0.1743 \end{pmatrix}$$

Nuestro sistema lineal de *tiempo continuo* expresado en variables de estado queda de la forma:

$$(17) \quad \dot{x} = A x(t) + B u(t).$$

## 6. PRINCIPIO DISCRETO DEL MÍNIMO

Al igual que con el multiplicador de Lagrange, ahora se define un vector de co-estado  $p(k)$  de  $(n \times 1)$ . Con esto, el problema de optimización equivale a minimizar, al funcional  $J_c$  dado por:

$$(18) \quad J_c = G[x(N), N] + \sum_{k=0}^{N-1} F[x(k), u(k), k] + \langle p(k+1), [x(k+1) - f(x, u, k)] \rangle$$

El siguiente paso consiste en definir la función escalar  $H[x(k), u(k), p(k+1), k]$  conocida como el Hamiltoniano, en la forma:

$$(19) \quad H[x(k), u(k), p(k+1), k] = F[x(k), u(k), k] + \langle p(k+1), f[x(k), u(k), k] \rangle,$$

que es la base del principio discreto del Mínimo [2], [6].

Por tanto, al resumir, la condición necesaria para que  $J_c$  tenga un extremo es:

$$\begin{aligned} \frac{\partial H(k)}{\partial x(k)} &= p(k), \\ \frac{\partial H(k)}{\partial p(k+1)} &= x(k+1), \\ \frac{\partial H(k)}{\partial u(k)} &= 0, \\ \frac{\partial G(N)}{\partial x(N)} &= p(N). \end{aligned}$$

Las primeras dos relaciones representan  $(2n)$  ecuaciones de diferencias conocidas como las ecuaciones canónicas de estado. La tercera relación proporciona el control óptimo  $u^o(k)$ , mientras que la última relación es la condición de transversalidad cuando  $x(N)$  no es fijo.

## 7. DISEÑO DEL CONTROL MEDIANTE EL PRINCIPIO DEL MÍNIMO DISCRETO

Deseamos obtener un sistema discreto para poder aplicar el principio del mínimo [2], [4], procedemos a discretizar el sistema (17), y obtener así al sistema:

$$(20) \quad \dot{x}[(k+1)T] = G(T) x(kT) + H(T) u(kT),$$

donde  $T = 100$ .

Entonces, las matrices resultantes son:

$$G = \begin{pmatrix} 1 & 0.01500 & 0.010154669 \\ 0 & 1 & 0.09806409 \\ 0 & 0 & 0.96153343 \end{pmatrix}, \quad H = \begin{pmatrix} 0.00009035 \\ 0.00086040 \\ 0.01709625 \end{pmatrix},$$

Nuestro sistema ya discretizado es:

$$(21) \quad \dot{x}[k+1] = G x(k) + H u(k)$$

Ahora deseamos minimizar al funcional

$$(22) \quad J_c = \sum_{k=0}^{N-1} F[x(k), u(k), k] - \langle p(k+1), [x(k+1) - f(x, u, k)] \rangle$$

sujeto a la restricción de igualdad (21).

Para ello definimos ahora al Hamiltoniano de la siguiente forma:

$$\begin{aligned} H &= [x_1^2(k), x_2^2(k), x_3^2(k) + u^2(k)] + \\ &+ \left\langle \begin{pmatrix} P_1(k+1) \\ P_2(k+1) \\ P_3(k+1) \end{pmatrix} \begin{pmatrix} 1 & 0.01500 & 0.01015466 \\ 0 & 1 & 0.09806409 \\ 0 & 0 & 0.96153343 \end{pmatrix} \begin{pmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{pmatrix} + \right. \\ &\left. + \begin{pmatrix} 0.00009035 \\ 0.00086040 \\ 0.01709625 \end{pmatrix} u(k) \right\rangle. \end{aligned}$$

Si derivamos con respecto de  $p(k+1)$  la expresión anterior, obtendremos el siguiente sistema de ecuaciones:

$$\begin{aligned} P_1(k+1) - P_1(k) &= -x_1(k) \\ P_1(k+1)(0.01500) + P_2(k+1) - P_2(k) &= -x_2(k) \\ P_1(k+1)(0.01054) + P_2(k+1)(0.09806) + P_3(k+1)(0.96153) - P_3(k) &= -x_3(k) \end{aligned}$$

las condiciones finales del sistema anterior se toman son todas iguales a 0.

También derivamos con respecto de  $u(k)$ , que nos proporciona la relación para el control óptimo, y queda expresado como sigue,

$$u(k) = -[P_1(k+1)(0.000090) + P_2(k+1)(0.000860) + P_3(k+1)(0.01709)]$$

De esta manera, resolviendo el sistema (21) junto con el sistema que involucra las variables  $P$  se determina la ley de control.

## 8. CONCLUSIONES

La aplicación del principio del Mínimo para nuestro sistema discreto resulta menos complicada que en el caso continuo. El control sintetizado tiene una expresión relativamente simple, en la situación que consideramos una trayectoria sencilla (línea recta horizontal), lo cual resulta hace más idóneo de aplicar.

Una extensión de este trabajo es considerar el planteamiento análogo para las trayectorias restantes.

## REFERENCIAS

- [1] Aníbal Ollero Baturone, *Robótica Manipuladores y Robots Móviles*, Marcombo, 2001.
- [2] Benjamin C. Kuo, *Sistemas De Control Digital*, CECSA, 2000, pp. 609 - 723.
- [3] Donald E. Kirk, *Optimal Control Theory An Introduction*, Prentice-Hall, 1970.
- [4] Jean Jaques E. Slotine y Weiping Li., *Applied Nonlinear Control*, Pearson Education, Republic of China, 2004.
- [5] V.V. Alexandrov, L. Guerra, *Optimization and computer - aided testing of stabilization precision, Mathematical modeling of complex information processing systems*, Editado por la Benemérita Universidad Autónoma de Puebla en colaboración con la Moscow State University, 2001, pp. 49 - 60.
- [6] Pablo Pedregal, *Introduction to Optimization*, Springer, 2004, 195-236.

Edificio 182, Ciudad Universitaria, Av. San Claudio y 18 Sur,  
Col. Jardines de San Manuel Puebla, Pue. México.  
Teléfono: 2295500 Ext: 7400 C.P. 72570  
jmgutierrez@ece.buap.mx  
mmorin@ece.buap.mx  
gelacio.salas@fce.buap.mx

# CONTROL ÓPTIMO Y UNA PROPUESTA DE SOLUCIÓN NUMÉRICA PARA UN PVF

AGUSTÍN MÉNDEZ ANDRADE  
EVODIO MUÑOZ AGUIRRE  
UNIVERSIDAD VERACRUZANA

RESUMEN. En este trabajo se muestra la necesidad de resolver numéricamente un problema de valores en la frontera que surge de una aplicación del Principio del Máximo de Pontryagin para calcular la dosis de droga usada en el tratamiento de pacientes cero positivos del VIH. Enseguida se expone el método numérico conocido como método del disparo como una propuesta para calcular el control óptimo. Se inicia con el caso lineal, posteriormente se trata el caso no lineal, debido a que el modelo final se presenta en esta forma.

## 1. INTRODUCCIÓN

Johan Bernoulli fué el primero en tratar el Principio de Optimalidad, en conexión con el problema de la braquistocrona que surgió en 1696. Este problema fue resuelto por los hermanos Bernoulli e I. Newton, quedando claro que la  $op$  es una propiedad fundamental en cuanto a sistemas naturales se refiere. Se investigaron varios principios de optimalidad, el principio de tiempo mínimo de P. Fermat (1600), empleado en óptica, y los trabajos de Hamilton y L. Euler (1744).

Estos principios de optimalidad son todos principios mínimos. A comienzos de 1900, A. Einstein demostró que, referido al sistema de coordenadas espacio-tiempo en cuatro dimensiones, la moción de los sistemas ocurre siempre de manera que se desea maximizar el tiempo.

Durante 1958, L.S. Pontryagin desarrolló su principio del máximo, el cual resolvía los problemas de control óptimo basándose en el cálculo de variaciones desarrollado por L. Euler (1707-1783). El verdadero elemento habilitado para el uso de la Teoría del Control Óptimo era el computador digital, el cual estuvo disponible comercialmente en los 50's. al final de los 50's y principios de los 60's, Lawden, Leitmann, Miele y Breakwell mostraron posibles usos del Cálculo de Variaciones en la optimalidad de rutas de vuelo aeroespacial usando algoritmos de disparo, mientras que Kelley y Bryson utilizaron el algoritmo de gradiente, que eliminaba la inestabilidad inherente de los algoritmos de disparo. En 1960, R. Kalman y sus cooperadores publicaron tres escritos muy importantes. En uno de ellos se reconocía la importancia del trabajo desarrollado por Lyapunov referente al control en el dominio del tiempo de sistemas no lineales. En otro, se discutía sobre el control óptimo de sistemas, proporcionando una serie de ecuaciones para el regulador cuadrático lineal (LQR). El tercer escrito discutía acerca del filtrador óptimo y la teoría de estimación, proporcionando las ecuaciones para el filtro discreto de Kalman. El filtro continuo de Kalman se desarrollaría poco después por Kalman y Bucy (1961).

En el periodo de un año, se hicieron patentes las limitaciones de la teoría de control clásica, comenzaron a introducirse nuevas herramientas teóricas y comenzó una nueva era en la teoría de control que se dio en llamar la era del control moderno [13].

La teoría de Control Óptimo se basa en minimizar o maximizar cierto índice de funcionamiento, teniendo en cuenta las restricciones del sistema. Las características principales de un sistema de control son; las variables de estado, las variables de salida y las variables de control. El principal resultado de Control Óptimo es el Principio del Máximo de Pontryagin, este principio se utilizará para encontrar la dosis óptima del problema que se estudia, por lo que se describirá con detalle más adelante. Además, la Teoría de Control Óptimo se utiliza para desarrollar sistemas automotores marinos, manufactura, sistemas robóticos, procesos industriales y químicos, modelos socioeconómicos, sistemas eléctricos y biomédicos. Es en ésta última clase de sistemas en donde se aplicará el Principio del Máximo de Pontryagin.

Más concretamente, el número de individuos VIH-cero positivos (VIH: Virus de Inmunodeficiencia Humana) está en continuo aumento alrededor del mundo, y esto constituye un problema serio en términos de dirección de salud.

Según un reporte publicado por la ONU (Organización de Naciones Unidas), aproximadamente 40.3 millones de personas están infectadas por VIH hasta el 2005. Esta estadística indica un incremento de aproximadamente un 7 por ciento de nuevos casos sobre el 2003, cuando el número de personas infectadas fue alrededor de 37.5 millones [14].

Desde la perspectiva inmunológica, el virus del síndrome de la inmunodeficiencia humana tiene características particulares que lo hacen muy interesante: el tiempo promedio entre una infección primaria y la aparición del sida es alrededor de 10 años. Este lapso de tiempo no ha sido, a la fecha, satisfactoriamente explicado, aunque es muy posible que se encuentre relacionado con el número de linfocitos  $T$  presentes en el torrente sanguíneo. Estos linfocitos  $CD_4+T$  son uno de los principales blancos del VIH. Este tipo de glóbulos blancos segregan factores de crecimiento y diferenciación que necesitan otras células del sistema inmune; por ello a estos linfocitos  $CD_4+T$  se les denomina "células  $T$  cooperadoras o inductoras". La carga normal de células  $CD_4+T$  es de  $1000\text{ mm}^{-3}$ . En un paciente infectado con VIH esta concentración puede disminuir hasta  $200\text{ mm}^{-3}$  o menos. De hecho, cuando una persona cuyo conteo de linfocitos  $CD_4+T$  llegue a esta concentración, se le diagnostica como clínicamente enferma de Síndrome de Inmunodeficiencia Adquirida (SIDA). El papel regulatorio tan importante que los linfocitos  $T$  cooperadores juegan en la respuesta inmune, produce que su disminución tenga efectos perniciosos muy amplios en el funcionamiento de la respuesta inmune que conlleva a la inmunodeficiencia que caracteriza al SIDA. Se utilizará el Principio del Máximo de Pontryagin para establecer una estrategia de tratamiento que balancee la terapia y los efectos colaterales, tomando un índice de funcionamiento ya conocido [4].

Se sabe que por su complejidad, muchas ecuaciones diferenciales ordinarias no pueden resolverse mediante fórmulas o métodos analíticos, tal es el caso de las

ecuaciones obtenidas en el presente análisis, es por eso que se propone un método numérico para la solución aproximada al problema de control que surge en la tercera sección.

El trabajo se desarrolla como sigue:

En la sección 2 se describe a grandes rasgos el Principio del Máximo de Pontryagin; en la sección 3 se describirá el modelo a utilizar, así como los elementos que lo componen y con ayuda de éste se plantea el problema de control para el modelo del VIH. En esta sección se llega al problema de valores en la frontera. Por último, en la sección 4 se desarrolla el método numérico del disparo para problemas de valores en la frontera bi-puntuales como el obtenido en la sección anterior, y se explica la relación que existe con el problema del apartado anterior.

## 2. TEORÍA DE CONTROL

Consideremos el problema autónomo:

$$(1) \quad \dot{x} = f(x, u), \quad x(t) \in \mathfrak{R}^n, \quad u(t) \in \mathfrak{R}^m,$$

con tiempo inicial fijo y estado inicial  $x(t_o) = x_o$ , estado objetivo  $\tau(t) \equiv x_1$  (posiblemente fijo); donde  $x(t) \in \mathfrak{R}^n$  es el estado del sistema y  $u(t) \in \mathfrak{R}^m$  es el control.

La teoría de control óptimo consiste en buscar un control  $u = u(t)$  dentro de una clase de controles admisibles  $U$ , con el fin de conducir un estado inicial  $x_0$  a un estado objetivo  $x_1$  mediante el sistema de ecuaciones diferenciales (1), que minimiza un determinado índice de rendimiento (costo funcional):

$$J = \int_a^b L(x, u) dt.$$

Particularmente, el costo funcional puede ser determinado por:

$$C[u(\cdot)] = \int_{t_o}^{t_1} f^o(x(t), u(t)) dt,$$

con controles medibles  $u(\cdot)$  que toman sus valores en un conjunto acotado fijo  $\Psi \subset \mathfrak{R}^m$ , se supone además que  $(u(\cdot), x(\cdot))$  es un par de respuesta de control óptimo. Para un control dado  $u(\cdot)$  y respuesta asociada  $x(\cdot)$ , se define la variable *Costo Dinámico*

$$x^o(t) = \int_{t_o}^t f^o(x(t), u(s)) ds$$

Si  $u(\cdot)$  es exitoso, lo cual significa que  $\exists t_1 \geq 0$  tal que  $x(t_1; x_o, u(\cdot)) \in \tau(t_1)$ , entonces  $x(t_1) = x_1$  para algún  $t_1 \geq t_o$ , y el costo asociado es

$$x^o[t_1].$$

Si se incrementa la dimensión del problema definiendo el vector  $(n + 1)$ , se tiene

$$\hat{x} = (x^o, x^T)^T \quad y \quad \hat{f}(t, \hat{x}) = (f^o, f^T)^T$$

Consideramos el sistema lineal de dimensión  $(n + 1)$ , conocido como el sistema adjunto

$$\hat{w}' = -\hat{f}_{\hat{x}}(x(t), u(t))^T \hat{w}(t), \quad \text{casi donde quiera.}$$

Se define:

$$H(\hat{w}, \hat{x}, u) \equiv \langle \hat{w}, \hat{f} \rangle \equiv \sum_{j=0}^n w^j(t) f^j(x(t), u(t))$$

Entonces  $H$  es un Hamiltoniano para el sistema y el sistema adjunto, es decir,

$$\begin{aligned} \hat{x}' &= \text{grad}_{\hat{w}} H(\hat{w}, \hat{x}, u) = \left( \frac{\partial H}{\partial w^0}, \frac{\partial H}{\partial w^1}, \dots, \frac{\partial H}{\partial w^n} \right)^T \text{ c.d.q.} \\ \hat{w}' &= -\text{grad}_{\hat{x}} H(\hat{w}, \hat{x}, u) = -\left( \frac{\partial H}{\partial x^0}, \frac{\partial H}{\partial x^1}, \dots, \frac{\partial H}{\partial x^n} \right)^T \text{ c.d.q.} \end{aligned}$$

Una herramienta que ayuda a calcular el control óptimo y trayectoria óptima, es el siguiente teorema [9], [7], [1].

**2.1. TEOREMA. (El Principio del Máximo de Pontryagin)** Sea el problema de control (1) con controles medibles  $u(\cdot)$  tomando valores en un conjunto acotado fijo  $\psi \subset \mathbb{R}^m$ . Supóngase que  $(u(\cdot), \hat{x}(\cdot))$  es un par de respuesta de control óptimo. Entonces existe una función absolutamente continua  $\hat{w}(\cdot)$  que resuelve la adjunta casi donde quiera sobre  $[t_o, t_1]$  con

$$\begin{aligned} H(\hat{w}(t), x(t), u(t)) &= M(\hat{w}(t), x(t)) \text{ c.d.q.} \\ M(\hat{w}(t), x(t)) &= 0 \text{ sobre } [t_o, t_1], \\ w^o(t) = w^o(t_o) &\leq 0 \quad \text{y} \quad \hat{w}(t) \neq \hat{o} \text{ sobre } [t_o, t_1] \end{aligned}$$

$$M(\hat{w}(t), x(t)) = \sup_{v \in \psi} H(\hat{w}(t), x(t), v(t)).$$

En la siguiente sección se muestra un ejemplo en donde se aplica el teorema anterior.

### 3. MODELO DEL VIH

El VIH puede ser encontrado en concentraciones significativas de sangre y en algunas secreciones del cuerpo humano, tales como esperma y fluido vaginal, eventualmente puede ser encontrado en el sistema nervioso central.

Al VIH se le llama retrovirus porque su información genética es codificada en una cadena de RNA (Ácido Ribonucleico) y ésta necesita ser convertida al formato DNA (Ácido Deoxyrebonucleico) para ser asimilada en el núcleo de las células receptoras. En el caso del VIH, las células receptoras son del tipo CD4+T que participan en el sistema inmunológico. El VIH necesita un encima conocida como transcripción inversa que realiza la codificación de DNA a RNA. Por lo tanto, inhibir la transcripción inversa es el mecanismo que se utilizará para controlar la multiplicación del virus. Otro mecanismo es la inhibición de la encima de proteasa, requerido para la segmentación de la cadena de polypeptides sintetizado en las partículas virales.

Para el modelo dinámico que describe la variación de células  $CD4 + T$  y las partículas virales, se usará una versión adaptada del modelo propuesto por Tao and Wu. El modelo está dado por las ecuaciones [4]:

$$(2) \quad \begin{aligned} \dot{x}_1 &= S + \lambda x_1 - x_1 \{ \mu_1 + k_1 x_4 \} \\ \dot{x}_2 &= w k_1 x_4 x_1 - x_2 \{ \mu_2 + k_2 \} \\ \dot{x}_3 &= (1 - w) k_1 x_4 x_1 + k_2 x_2 - \mu_3 x_3 \\ \dot{x}_4 &= N \mu_3 x_3 - x_4 \{ k_1 x_1 + \mu_v \}, \end{aligned}$$

donde

$$\begin{aligned}
 S = S(x_4) &= \frac{S\theta}{\theta + x_4} \\
 \lambda = \lambda(x_1, x_2, x_3) &= r\left[1 - \frac{x_1 + x_2 + x_3}{T_{max}}\right] \\
 N = N(t) &= \beta_2 - (\beta_2 - N_0) \exp^{-\beta_1 t} \\
 k_1(m_1(t)) &= k_{10} \exp^{-\alpha_1 m_1(t)} \\
 k_2(m_2(t)) &= k_{20} \exp^{-\alpha_2 m_2(t)},
 \end{aligned}$$

con:

$x_1$ : densidad de células CD4+T no infectadas;  $x_2$ : densidad de células CD4+T latentemente infectadas;  $x_3$ : densidad de células CD4+T infectadas;  $x_4$ : densidad de virus libre de VIH;  $s$ : índice de generación de  $x_1$ ;  $\theta$ : concentración viral necesaria para decrementar  $s$ ;  $\lambda$ : índice de generación de nuevas  $x_1$  en presencia de VIH;  $r$ : índice de crecimiento simultáneo de  $x_1$ .  $T_{max}$ : nivel de población máximo de células T;  $\mu$ : índice de mortandad de  $x_1$ ;  $k_1$ : índice de infección de  $x_1$  a  $x_2$  por virus;  $k_{10}$ : constante de ajuste para la función  $k_1(t)$ ;  $\alpha_1$ : sensibilidad de la acción del inhibidor de transcripción inversa;  $m_1$ : dosis de inhibición de transcripción inversa;  $\omega$ : índice de células que se vuelven activas;  $\mu_2$ : índice de muerte de  $x_2$ ;  $k_2$ : índice de conversión de  $x_2$  a  $x_3$ ;  $k_{20}$ : constante de ajuste para la función de  $k_2(t)$ ;  $\alpha_2$ : sensibilidad de la acción del inhibidor de proteasa;  $m_2$ : dosis de inhibidor de proteasa;  $\mu_3$ : índice de muerte de  $x_3$ ;  $N$ : número de virus infectado producido por una célula T activa;  $N_0$ : número de virus al inicio del tratamiento;  $\beta_2$ : el término se adapta de modo que el número de virus tenga una disminución exponencial;  $\mu_v$ : índice de muertes de  $x_4$ .

El costo funcional usado está dado por:

$$(3) \quad J[m] = \int_{t_0}^{t_f} \{\zeta + \psi + \nu\} dt$$

donde

$$\begin{aligned}
 \zeta &= \phi_1(1 - \epsilon_1 m_1(t) \exp^{-\alpha_1 m_1(t)}) \\
 \psi &= \phi_2(1 - \epsilon_2 m_2(t) \exp^{-\alpha_2 m_2(t)}) \\
 \nu &= \frac{\gamma_1}{x_1^2(t)} + \gamma_2 x_2^4(t)
 \end{aligned}$$

Aquí,  $\Phi_1$  es la carga que refleja la importancia de los efectos colaterales;  $m_1$  es la dosis del inhibidor de transcripción inversa;  $\Phi_2$  es la carga que refleja la importancia de los efectos colaterales,  $m_2$  es la dosis del inhibidor de proteasa;  $\epsilon_1, \epsilon_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2$  son constantes;  $x_1 = x_1(t)$  el número de células CD4+T;  $x_4 = x_4(t)$  es la cantidad de virus libre de VIH.

El problema de control óptimo es encontrar la función  $m(t) = (m_1(t), m_2(t))^T$  que minimice el costo funcional sujeto a condiciones iniciales tomadas de [4]:

$$(4) \quad \begin{aligned} x_1(t_i) &= 30 \\ x_2(t_i) &= 300 \\ x_3(t_i) &= 110 \\ x_4(t_i) &= 407000 \end{aligned}$$

De acuerdo al Principio del Máximo de Pontryagin de la sección anterior, el sistema adjunto se obtiene por:

$$\hat{w}' = -\text{grad}_{\hat{x}} H(\hat{w}, \hat{x}, u) = -\left(\frac{\partial H}{\partial x^0}, \frac{\partial H}{\partial x^1}, \dots, \frac{\partial H}{\partial x^n}\right)^T \text{ c.d.q}$$

donde

$$H = \langle \hat{w}, \hat{f} \rangle,$$

con

$$(5) \quad \hat{f} = \begin{pmatrix} S + \lambda x_1 - x_1\{\mu_1 + k_1 x_4\} \\ w k_1 x_4 x_1 - x_2\{\mu_2 + k_2\} \\ (1-w)k_1 x_4 x_1 + k_2 x_2 - \mu_3 x_3 \\ N \mu_3 x_3 - x_4\{k_1 x_1 + \mu_v\} \end{pmatrix}$$

y  $\hat{w} = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ .

Realizando la diferenciación correspondiente, se llega al sistema adjunto:

$$(6) \quad \begin{aligned} \dot{\lambda}_1 &= -\lambda_1 r + \frac{2\lambda_1 r x_1}{T_{max}} + \lambda_1(\mu_1 + k_1 x_4 - \lambda_3(1-w)k_1 x_4 + \lambda_4 x_4 k_1) - \frac{2\gamma_1}{x_1^3} \\ \dot{\lambda}_2 &= \frac{\lambda_1 r x_1}{T_{max}} + \lambda_2(\mu_2 k_2) - \lambda_3 k_2 \\ \dot{\lambda}_3 &= \frac{\lambda_1 r x_1}{T_{max}} + \lambda_3 \mu_3 - \lambda_3 \beta_4 \mu_3 + \lambda_4(\beta_2 - N_0) \exp^{-\beta_1 t} \mu_3 \\ \dot{\lambda}_4 &= \frac{\lambda_1 s \theta}{(\theta + x_4)^2} + \lambda_1 x_1 k_1 - \lambda_2 w k_1 x_1 - \lambda_3(1-w)k_1 x_1 + \lambda_4 k_1 x_1 \\ &\quad + \lambda_4 \mu_v + 2x_4 \gamma_2. \end{aligned}$$

Como el problema de control óptimo es con punto objetivo final libre, las condiciones finales para el sistema adjunto pueden ser, según [7]:

$$(7) \quad \begin{aligned} \lambda_1(t_f) &= 0 \\ \lambda_2(t_f) &= 0 \\ \lambda_3(t_f) &= 0 \\ \lambda_4(t_f) &= 0 \end{aligned}$$

Nótese que no se ha escrito la variable costo dinámico para el sistema adjunto, sin embargo sí se utiliza para poder aplicar el Principio del Máximo de Pontryagin, aunque no es necesario realizar los cálculos para ésta, ya que ésta última no es de interés para el problema, puesto que lo principal es encontrar la dosis óptima y en ella no se inmiscuye dicha variable.

El problema de control óptimo inicial se ha reducido a un problema de valores en la frontera con 8 ecuaciones diferenciales ordinarias, 4 condiciones iniciales y 4 condiciones finales. Es muy difícil resolverlo explícitamente, si no es que sea imposible, así lo mejor es abordarlo desde el punto de vista numérico, el método apropiado se expone en la siguiente sección.

#### 4. MÉTODO NUMÉRICO DEL DISPARO PARA PROBLEMAS DE VALOR EN LA FRONTERA BI-PUNTUALES

A continuación se explica el método del disparo para la solución de problemas de valores en la frontera, el cual consiste en transformar el problema en uno de valores iniciales completando la condición inicial, y se ajusta para que la curva pase por el otro extremo [7]. Tomaremos en cuenta el método para la solución numérica de los siguientes tipos de problemas:

- Se requiere resolver  $n$  ecuaciones de primer orden sobre el intervalo  $[t_0, t_f]$ , donde  $t_0$  es el punto inicial, y  $t_f$  es el punto final;
- Se especifican  $r$  condiciones de frontera en  $t_0$ ;
- Se especifican  $(n - r)$  condiciones de frontera en  $t_f$ ;

Sin pérdida de generalidad se tomará el problema de la forma:

$$(8) \quad \begin{aligned} \dot{y}_i &= g_i(y_1, y_2, \dots, y_n, t) \quad i = 1, \dots, n \\ y_i(t_0) &= c_i, \quad i = 1, \dots, r \\ y_i(t_f) &= c_i, \quad i = r + 1, \dots, n \end{aligned}$$

donde cada  $g_i$  es dos veces diferenciable con respecto a  $y_i$ .

##### 4.1. Método del Disparo Lineal.

Primero consideremos el caso lineal

$$(9) \quad \dot{y}(t) = A(t)y(t) + f(t), \quad t_0 \leq t \leq t_f,$$

donde

$$\begin{aligned} y_i(t_0) &= c_i, \quad i = 1, \dots, r \\ y_i(t_f) &= c_i, \quad i = r + 1, \dots, n \end{aligned}$$

con  $c$  y  $f(t)$  dados.

El sistema adjunto está definido como la solución de la ecuación homogénea.

$$(10) \quad \dot{z}(t) = -A^T(t)z(t)$$

Como siempre, la solución general de (9) se escribe como

$$(11) \quad y(t_f) = X(t_f)y(t_0) + \int_{t_0}^{t_f} X(t_f)X(s)^{-1}f(s)ds$$

donde

$$\dot{X} = A(t)X, \quad X(t_0) = I.$$

Además, la solución de (10) puede escribirse

$$(12) \quad z(t_f) = X(t_f)^{-T}z(t_0)$$

y tomando (11) se tiene

$$(13) \quad z(t_f)^T y(t_f) = z(t_f)^T X(t_f)y(t_0) + z(t_f)^T \int_{t_0}^{t_f} X(t_f)X(s)^{-1}f(s)ds$$

Sustituyendo la transpuesta de (12) en (13):

$$z(t_f)^T y(t_f) = z(t_0)^T y(t_0) + \int_{t_0}^{t_f} z(t_f)^T X(t_f)X(s)^{-1}f(s)ds$$

Ahora, otra vez por (12)

$$z(s) = X(s)^{-T} x(t_f)^T z(t_f),$$

y

$$(14) \quad z(t_f)^T y(t_f) = z(t_0)^T y(t_0) + \int_{t_0}^{t_f} z(s)^T f(s) ds,$$

o en forma de componentes,

$$(15) \quad \sum_{i=1}^n z_i(t_f) y_i(t_f) - \sum_{i=1}^n z_i(t_0) y_i(t_0) = \int_{t_0}^{t_f} \sum_{i=1}^n z_i(s) f_i(s) ds$$

Las ecuaciones (14) y (15) son las identidades básicas en el método de las adjuntas, el primero de los métodos del disparo.

El método inicia integrando las ecuaciones adjuntas (10) hacia atrás en el tiempo  $n - r$  veces, con la condición de frontera terminal.

$$z^{(1)}(t_f) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} rth;$$

de igual manera

$$(16) \quad z^{(n-r)}(t_f) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

Esto nos da  $n - r$  funciones  $z^{(1)}(t)$ ,  $z^{(2)}(t)$ ,  $\dots$ ,  $z^{(n-r)}(t)$ ,  $t_0 \leq t \leq t_f$  (claro en la práctica se integra sobre un conjunto discreto de puntos). Entonces

$$z^{(m)}(t_f)^T y(t_f) = \sum_{i=1}^n z_i^{(m)}(t_f) y_i(t_f) = y_{r+m}(t_f) = c_{r+m}, \quad m = 1, \dots, n - r$$

y convirtiendo (15), con algún reordenamiento,

$$(17) \quad \sum_{i=r+1}^n z_i^{(m)}(t_0) y_i(t_0) = c_{r+m} - \sum_{i=1}^r z_i^{(m)}(t_0) y_i(t_0) - \int_{t_0}^{t_f} \sum_{i=1}^n z_i^{(m)}(s) f_i(s) ds$$

para  $m = 1, \dots, n - r$ .

El conjunto de ecuaciones (17) es un conjunto de  $n - r$  ecuaciones algebraicas en  $n - r$  incógnitas  $\{y_{r+1}(t_0), y_{r+2}(t_0), \dots, y_n(t_0)\}$  que se resuelve fácilmente, y así de esta manera, se puede encontrar un conjunto completo de condiciones iniciales en  $t = t_0$ .

Nótese que (17) se puede escribir como:

$$(18) \quad \begin{pmatrix} z_{r+1}^{(1)}(t_0) & z_{r+2}^{(1)}(t_0) & \cdots & z_n^{(1)}(t_0) \\ z_{r+1}^{(2)}(t_0) & z_{r+2}^{(2)}(t_0) & \cdots & z_n^{(2)}(t_0) \\ \vdots & \vdots & \ddots & \vdots \\ z_{r+1}^{(n-r)}(t_0) & z_{r+2}^{(n-r)}(t_0) & \cdots & z_n^{(n-r)}(t_0) \end{pmatrix} \begin{pmatrix} y_{r+1}(t_0) \\ y_{r+2}(t_0) \\ \vdots \\ y_n(t_0) \end{pmatrix}$$

$$= \begin{pmatrix} c_{r+1} - \sum_{i=1}^r z_i^{(1)}(t_0)c_i - \int_{t_0}^{t_f} \sum z_i^{(1)}(t)f_i(t)dt \\ c_{r+2} - \sum_{i=1}^r z_i^{(2)}(t_0)c_i - \int_{t_0}^{t_f} \sum z_i^{(2)}(t)f_i(t)dt \\ \vdots \\ c_n - \sum_{i=1}^r z_i^{(n-r)}(t_0)c_i - \int_{t_0}^{t_f} \sum z_i^{(n-r)}(t)f_i(t)dt \end{pmatrix}$$

La inversa de la matriz  $z$  existe ya que  $\{z^{(1)}(t_f), \dots, z^{(n-r)}(t_f)\}$  son linealmente independiente, y consecuentemente lo son

$$\{z^{(1)}(t), \dots, z^{(n-r)}(t)\}$$

para algún  $t \in [t_0, t_f]$ . Éste es un hecho conocido por las ecuaciones lineales diferenciales ordinarias.

Resumiendo, se tiene el siguiente algoritmo para este primer caso.

- a:** Sea  $m = 1$
- b:** Integrar las ecuaciones adjuntas hacia atrás desde  $t_f$  a  $t_0$  para el  $m$ -ésimo conjunto de condiciones iniciales (16)
- c:** Evaluar la  $m$ -ésima fila de (18)
- d:** Si  $m = n - r$ , resolver el sistema algebraico (18) para las condiciones iniciales  $\{y_{r+1}(t_0), \dots, y_n(t_0)\}$ ; ir al ítem 6.
- e:** Si  $m < n - r$ , poner  $m = m + 1$ ; regresar al ítem 2.
- f:** Usando el conjunto total de condiciones iniciales  $\{y_i(t_0)\}_{i=1}^n$ , se integra (1) para obtener la solución al problema de valores en la frontera.

**4.2. Método del Disparo No Lineal.** Los problemas no lineales en la frontera bi puntuales se resuelven por un proceso iterativo.

Iniciaremos con una condición inicial supuesta

$$\{y_{r+1}^{(0)}(t_0), y_{r+2}^{(0)}(t_0), \dots, y_n^{(0)}(t_0)\}.$$

Esto nos permite resolver la ecuación (1) o hallar  $y^{(0)}(t)$ ,  $t_0 \leq t \leq t_f$ .

Entonces, nuestra iteración es según el siguiente esquema.

Sea

$$(19) \quad \delta y^{(k)}(t) = y^{(k+1)}(t) - y^{(k)}(t), \quad t_0 \leq t \leq t_f$$

Por la expansión usual de Series de Taylor, se tiene una primera aproximación

$$\delta \dot{y}_i^{(k)}(t) = \sum_{j=1}^n \frac{\partial g_i}{\partial y_j}(y^{(k)}(t)) \delta y_j^{(k)}(t), \quad i = 1, \dots, n$$

esto es,

$$(20) \quad \delta \dot{y}^{(k)} = J(y^{(k)}) \delta y^{(k)}$$

(donde  $J$  es el gradiente de  $g$ ), que es sólo un conjunto de ecuaciones diferenciables con coeficientes variables.

Además, se puede calcular cada condición inicial para los deltas de la forma:

$$(21) \quad \begin{aligned} \delta y_i^{(k)}(t_0) &= 0, \quad i = 1, \dots, r \quad k = 0, 1, 2, \dots \\ \delta y_i^{(k)}(t_f) &= c_i - y_i^{(k)}(t_f), \quad i = r+1, \dots, n \quad k = 0, 1, 2, \dots \end{aligned}$$

Las ecuaciones (20) y (21) definen un problema lineal de valor en la frontera bi puntual para el vector corrección  $\delta y^{(k)}(t)$ ,  $t_0 \leq t \leq t_f$ . Este se resuelve por el método anterior, el método de adjuntas; es decir, como antes, se define el sistema adjunto al sistema (20) como :

$$(22) \quad \dot{z}^{(k)}(t) = -J_k^T(t) z^{(k)}(t),$$

donde  $J_k = J(y^{(k)}(t))$ , y resolviendo (22) hacia atrás  $n - r$  veces con condiciones iniciales en el punto final

$$rth \quad \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Se designa la solución así obtenida en la  $k$ -ésima iteración por

$$\{z^{(1)}(t), z^{(2)}(t), \dots, z^{(n-r)}(t)\}_{(k)},$$

entonces la identidad fundamental es:

$$(23) \quad \begin{pmatrix} z_{r+1}^{(1)}(t_0) & z_{r+2}^{(1)}(t_0) & \dots & z_n^{(1)}(t_0) \\ z_{r+1}^{(2)}(t_0) & z_{r+2}^{(2)}(t_0) & \dots & z_n^{(2)}(t_0) \\ \vdots & \vdots & \vdots & \vdots \\ z_{r+1}^{(n-r)}(t_0) & z_{r+2}^{(n-r)}(t_0) & \dots & z_n^{(n-r)}(t_0) \end{pmatrix} \begin{pmatrix} \delta y_{r+1}^{(k)}(t_0) \\ \delta y_{r+2}^{(k)}(t_0) \\ \vdots \\ \delta y_n^{(k)}(t_0) \end{pmatrix} \\ = \begin{pmatrix} \delta y_{r+1}^{(k)}(t_f) \\ \delta y_{r+2}^{(k)}(t_f) \\ \vdots \\ \delta y_n^{(k)}(t_f) \end{pmatrix}.$$

Resolviendo (23), se llega a un conjunto completo de condiciones iniciales  $\delta y^{(k)}(t_0)$  y se obtiene el siguiente cambio para la condición inicial en  $y$  por el conjunto

$$y^{(k+1)}(t_0) = y^{(k)}(t_0) + \delta y^{(k)}(t_0).$$

Ahora resolvemos para  $y^{(k+1)}(t)$ ,  $t \in [t_0, t_f]$ , y se regresa a (20) para calcular  $J(y^{(k+1)})$ ,  $\delta y^{(k+1)}$ , etc.

Terminamos cuando

$$\max\{\delta y_i^{(k)}(t_f) : i = r, r + 1, \dots, n\}$$

es suficientemente pequeña o  $k$  es muy grande.

Esto es lo que se conoce como el método del disparo; se supone el valor inicial desconocido, se resuelven las ecuaciones, sobre las bases de estas soluciones se hacen correcciones a los valores iniciales previos. En el método del disparo el objetivo es encontrar los datos iniciales faltantes.

Con esto se muestra que el método del disparo es un caso especial del método de Newton-Raphson. También se puede ver que si  $J_k$  es no singular y el intervalo  $[t_o, t_f]$  no es muy grande, la aproximación converge de forma cuadrática.

Resumiendo, el método de las adjuntas para ecuaciones diferenciales ordinarias no lineales se lleva a cabo como sigue:

- i:** Determinar analíticamente el gradiente  $(\frac{\partial g_i}{\partial y_j})$ .
- ii:** Iniciar el contador sobre el proceso iterativo. Sea  $k = 0$ .
- iii:** Para  $k = 0$ , suponer las condiciones iniciales  $y_i^{(0)}(t_0)$ ,  $i = r + 1, \dots, n$ .
- iv:** Integrar (1) de la sección con condiciones iniciales

$$\begin{aligned} y_i^{(k)}(t_0) &= c_i, \quad i = 1, 2, \dots, r \\ y_i^{(k)}(t_0), & \quad i = r + 1, \dots, n \end{aligned}$$

y almacenamos  $y^{(k)}$ .

- v:** Poner el contador sobre la integración de las ecuaciones adjuntas,  $m = 1$ .
- vi:** Calcular  $z_i^{(t_0)}$ ,  $i = r + 1, \dots, n$  integrando las ecuaciones adjuntas (22) hacia atrás para  $t_f$  a  $t_0$ , con datos finales  $z_i^{(m)}(t_f)$  de la ecuación (16) de la subsección 1,  $i = 1, \dots, n$ .

Notemos que en esta integración los perfiles almacenados  $y^{(k)}(t)$  son usados para evaluar las derivadas parciales  $\frac{\partial g_i}{\partial y_j}$ ,  $i, j = 1, \dots, n$ .

- vii:** Para la  $m$ -ésima fila de (23) para la posición derecha de (23) se sustrae el valor terminal específico  $y_i(t_f) = c_i$  de el valor calculado en  $y_i^{(k)}(t_f)$ ,  $i = r + 1, \dots, n$  fundado en el ítem iv:.
- viii:** Si  $m < n - r$ , se pone  $m = m + 1$  y regresa al ítem a vi:.
- ix:** Formar el conjunto de  $n - r$  ecuaciones algebraicas lineales (23), resolver para  $\delta y_i^{(k)}(t_0)$ ,  $i = r + 1, \dots, n$ .
- x:** Formar el siguiente conjunto de valores de prueba dados por

$$(24) \quad y_i^{(k+1)}(t_0) = y_i^{(k)}(t_0) + \delta y_i^{(k)}(t_0), \quad i = r + 1, \dots, n.$$

- xi:** Poner  $k = k + 1$ ; regresar al ítem iv:.

**xii:** Terminamos, cuando  $\max\{\delta y_i^{(k)}(t_f) : i = r + 1, \dots, n\}$  es suficientemente pequeño, o cuando  $k$  excede un valor máximo.

A manera de conclusión, el último algoritmo se adapta al problema de valores en la frontera formado por las ecuaciones diferenciales (2) y (6), las cuatro condiciones iniciales (4) y las cuatro condiciones finales (7), lo que permite calcular numéricamente la dosis óptima. De hecho, un trabajo a futuro es la simulación numérica de este algoritmo mediante algún software matemático y comparar la solución  $m(t)$ .

#### REFERENCIAS

- [1] Dimitri P. Bertsekas, *Dynamics programming and Optimal Control*, Vol I, Athena Scientific, Massachusetts, 1995.
- [2] Burden R. L., Fares J. D. *Análisis Numérico*, CENGAGE, México, 2003.
- [3] Arthur E. Bryson and Yu-Chi Ho *Applied Optimal Control*, Blaisdell Publishing Company, 1969.
- [4] Denise C. Cardoso Hardt, Marco A. Leonel Caetano and Takashi Yoneyama, *On the Optimal Doses of Transcriptase and Protease Inhibitors in the Treatment of HIV-Seropositive Patients*, Instituto Superior Técnico, Lisboa, Portugal, 2006.
- [5] Curtis F. Gerald, Patrick O. Weathley *Análisis Numérico con aplicaciones*, Addison Wesley, New York, 2000.
- [6] Chang, *Control Theory of bifurcation*, Springer Verlag, New York, 2003.
- [7] Greg Knowles, *An Introduction To Applied Optimal Control*, Academic Press, New York, 1981.
- [8] Magnus R. Hestenes, *Calculus of Variations and Optimal Control Theory*, Springer Verlag, New York, 1980.
- [9] Jack Macki and Aaron Straus, *Introduction To Optimal Control*, Springer Verlag, New York, 1982.
- [10] Shoichiro Nakamura. *Análisis Numérico visual. Gráficas con Matlab*, Prentice Hall, México, 1997.
- [11] Sharma J. N. *Numerical Methods for Egginers and Scientists*, Alpha Science International, India, 2005.
- [12] Jiongmin Yong and Xun Yu Zhou, *Stochastic Controls*, Springer Verlag, New York, 1999.
- [13] *Historia de la Teoría de Control*, <http://profefblog.es/blog/procesos/files/2008/03/historia-de-la-tmc.pdf>.
- [14] SIDA: Aspectos de la Salud Pública, <http://books.google.com.mx/books>.
- [15] Module for the Shooting Method for Boundary Value Problems. <http://math.fullerton.edu/mathews/n2003/web/ShootingMod/ShootingMod.html>. Última visita, nov de 2009.
- [16] A Maple Implementacion of the Simpe Shooting Method. <http://www.math.sc.edu/meade/maple/shoot/mtn-shoot4.html> Última visita, noviembre del 2009.

Facultad de Matemáticas, zona universitaria  
 Circuito Aguirre Beltrán s/n, c.p. 91090.  
[andradeagus@gmail.com.mx](mailto:andradeagus@gmail.com.mx)  
[evo2003001@yahoo.com.mx](mailto:evo2003001@yahoo.com.mx)

# VALIDACIÓN NUMÉRICA Y EXPERIMENTAL DE UN ALGORITMO DE IDENTIFICACIÓN DE FUENTES BIOELÉCTRICAS TIPO DIPOLO

MARÍA MONSERRAT MORÍN CASTILLO<sup>1</sup>

JACOBO OLIVEROS OLIVEROS<sup>2</sup>

BLANCA MINERVA RODRÍGUEZ RAMÍREZ<sup>1</sup>

<sup>1</sup>FACULTAD DE CIENCIAS DE LA ELECTRÓNICA - BUAP

<sup>2</sup>FCFM - BUAP

RESUMEN. El problema de identificación de fuentes bioeléctricas en el cerebro consiste en determinar dichas fuentes a través del potencial que producen midiéndolo sobre el cuero cabelludo a través de electrodos. A esta medición se le conoce como Electroencefalograma.

Por medio de un modelo de medio conductor se obtiene un problema de valores en la frontera que permite el análisis de este problema de identificación.

En este trabajo se presenta el análisis de los problemas directo e inverso electroencefalográfico para el caso correspondiente a una fuente dipolar, así como el planteamiento de un problema de minimización para recuperar los parámetros de la fuente dipolar por medio de las mediciones.

Debido al mal planteamiento de este problema, es necesario aplicar algún método de regularización que permita identificar de manera estable tales parámetros.

Para la validación experimental se construirá un sistema físico, que consiste en colocar un dipolo eléctrico dentro de una esfera conductora, con esto se simulara una fuente dipolar dentro de la cabeza.

Los potenciales experimentales se miden utilizando un arreglo de electrodos colocados sobre la superficie de la esfera. Además para la validación numérica es necesaria la construcción de ejemplos sintéticos.

## 1. INTRODUCCIÓN

El método de la Electroencefalografía (EEG) es el más conocido entre los métodos no invasivos en la investigación del cerebro y se basa en el registro de su actividad eléctrica. Así, mediante este método se puede detectar la presencia de algunas anomalías como: Tumores cerebrales, enfermedades infecciosas, retardo mental y últimamente muerte cerebral. Además, los potenciales obtenidos como respuesta a algún estímulo (potenciales evocados) se muestran prometedores en el diagnóstico y tratamiento de enfermedades del sistema nervioso central.

Para el análisis del mismo se utilizan modelos matemáticos que permiten realizar el análisis de existencia y unicidad del problema de identificación. En estos modelos se introducen hipótesis fisiológicas y anatómicas sobre la naturaleza y el comportamiento de las fuentes cerebrales; en particular los focos epilépticos que se pueden caracterizar matemáticamente por medio de dipolos.

En este trabajo se estudia el problema de identificación de fuentes en forma de dipolo al que se le asocian dos problemas matemáticos a seguir:

El Problema Directo Electroencefalográfico (PDE), consiste en determinar las mediciones de potencial electrostático generadas por una distribución espacial de agregados de neuronas (fuentes), utilizando la técnica del EEG.

El Problema Inverso Electroencefalográfico (PIE), consiste, a grosso modo, en determinar a partir de las mediciones de potencial electrostático, la distribución espacial de los agregados de neuronas (fuentes) [4], [5], que generan dichas mediciones.

## 2. PROBLEMA DIRECTO E INVERSO ELECTROENCEFALOGRÁFICO

Partiremos de un modelo utilizado en [4], [5], [8], [9]. En estos trabajos se supone a la cabeza humana  $\Omega$ , está compuesta por dos zonas disjuntas, a saber:

- (1)  $\Omega_1$  Cerebro,
- (2)  $\Omega_2$  Restantes regiones que componen a la cabeza.

Suponemos que cada una de las componentes  $\Omega_i$  tiene una conductividad constante  $\sigma_i$ ,  $i = 1, 2$ , con  $\sigma_1 \neq \sigma_2$ . Mediante  $S_i$  se denotarán las superficies que componen las fronteras de las regiones  $\Omega_i$ ,  $i = 1, 2$ :  $\partial\Omega_1 = S_1$ ,  $\partial\Omega_2 = S_1 \cup S_2$ , donde  $S_1$  representa la corteza cerebral y  $S_2$  al cuero cabelludo.

El potencial  $u$  producido en  $\Omega = \overline{\Omega}_1 \cup \overline{\Omega}_2$  ( $\overline{\Omega}_1$  denota la cerradura de  $\Omega_1$ ) por una fuente de corriente primaria  $J^p$  satisface el siguiente problema de contorno (1) (2):

$$(1) \quad \begin{aligned} \Delta u_1 &= f & \text{en } \Omega_1, \\ \Delta u_2 &= 0 & \text{en } \Omega_2, \end{aligned}$$

$$u_1 = u_2 \quad \text{en } S_1,$$

$$(2) \quad \begin{aligned} \sigma_1 \frac{\partial u_1}{\partial n_1} &= \sigma_2 \frac{\partial u_2}{\partial n_1} & \text{en } S_1, \\ \frac{\partial u_2}{\partial n_2} &= 0 & \text{en } S_2, \end{aligned}$$

donde  $f = \text{div} \left( \frac{J^p}{\sigma} \right)$ ,  $u_i = u|_{\Omega_i}$ ,  $i = 1, 2$ ,  $n_i$  es el vector normal unitario exterior a  $\Omega_i$  en  $S_i$ ,  $i = 1, 2$ ,  $\frac{\partial u_i}{\partial n_j}$  representa la derivada normal de  $u_i$  en  $S_j$  con respecto al vector  $n_j$ .

Las condiciones de contorno (2) corresponden a la continuidad del potencial eléctrico y de las componentes normales de la corriente en cada superficie  $S_j$ ,  $j = 1, 2$ .

Al problema (1) con las condiciones (2) lo llamaremos **Problema de Contorno Electroencefalográfico** (PCE).

Este modelo ha sido utilizado en [4], [5]. A partir del problema (1)-(2) se dan las siguientes dos definiciones:

2.1. DEFINICIÓN. (Problema Directo asociado al PCE) Llamaremos Problema Directo (**PDE**) asociado al PCE al problema que consiste en hallar la solución  $u(x)$  del PCE cuando está dada  $f(x)$ .

2.2. DEFINICIÓN. (Problema Inverso Electroencefalográfico asociado al PCE) (**PIE**) Dada una función  $V$  definida sobre  $S_2$  encontrar  $f(x)$  de manera que para la solución  $u(x)$  del problema directo correspondiente a  $f(x)$  se cumpla que:  $u|_{S_2} = V$ .

3. CASO DE UNA SOLA REGIÓN CONDUCTORA

Para simplificar la exposición supondremos en lo que sigue que  $\Omega$  corresponde a una bola de radio uno centrada en el origen con conductividad constante  $\sigma$ . Consideremos el problema de contorno:

$$(3) \quad \Delta u = f \quad \text{en } \Omega,$$

$$(4) \quad \frac{\partial u}{\partial n} = 0 \quad \text{en } \partial\Omega,$$

donde  $f$  representa a la fuente dipolar la cual puede representarse por ([2], [8])

$$(5) \quad f(P) = \frac{\text{div}(\rho(t)v(a,t)\delta(P-a))}{\sigma},$$

donde  $v = (v_1, v_2, v_3)$ , llamado momento dipolar, es un vector con origen en el punto  $a$ , que representa la velocidad de cargas en movimiento,  $\rho(t)$  representa una densidad de cargas y  $\delta(P-a)$  es la función delta de Dirac. En [3] y [5] se ha estudiado el PIE para el caso en que las fuentes están concentradas en la corteza cerebral [1], [4], [7] y [8], para fuentes en el volumen cerebral.

Debido a que la diferencia de dos cualesquiera soluciones del problema (3)-(4) es una constante, diremos que este problema tiene solución única salvo constantes. La solución de (3)-(4) se expresa a través de la función de Green.

3.1. DEFINICIÓN. La función  $G(P, P_0)$  es llamada la función de Green del problema (3)-(4) si satisface el problema.

$$\Delta_p G(P, P_0) = \delta(P - P_0) - \frac{1}{m(\Omega)} \quad P, P_0 \in \Omega,$$

$$\frac{\partial G}{\partial n_p}(P, P_0) = 0 \quad P \in \Omega, P_0 \in \partial\Omega,$$

donde  $\delta(P - P_0)$  es la función delta de Dirac y  $m(\Omega)$  es el volumen de  $\Omega$ . La solución del problema (3)-(4) está dada por ([9])

$$(6) \quad u(P_0) = \int_{\Omega} G(P, P_0) f(P) dP.$$

La función de Green, para la bola unitaria está dada por:

$$(7) \quad G(P, P_0) = \frac{1}{4\pi |P - P_0|} + \alpha R^2(P) + g_1(P, P_0),$$

donde  $P = (x, y, z)$ ,  $P_0 = (x_0, y_0, z_0)$ ,  $\alpha = -\frac{1}{6m(\Omega)}$ ,

$$|P - P_0| = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}, \quad R^2 = x^2 + y^2 + z^2$$

y la función  $g_1(P, P_0)$  satisface el problema,

$$\begin{aligned} \Delta_p g_1(P, P_0) &= 0, & P, P_0 \in \Omega \\ \frac{\partial g_1}{\partial n_p}(P, P_0) &= -\alpha \frac{\partial R^2}{\partial n}(P) - \frac{1}{4\pi} \frac{\partial}{\partial n} \left( \frac{1}{|P - P_0|} \right), & P_0 \in \partial\Omega, \quad P \in \Omega. \end{aligned}$$

En el caso en que  $f$  está dada por (5) con  $\rho(t) = q$ ,  $v = \vec{d}$  la expresión (6) toma la forma [2]:

$$(8) \quad u(P_0) = \left[ -\frac{q\vec{d}}{\sigma} \cdot \nabla_p G(P, P_0) \right] \Bigg|_{P=a}.$$

A través de (8) se halla el potencial producido por una fuente dipolar que tiene la forma (5).

Cuando  $P_0$  está sobre el eje  $Z$ , es decir, cuando  $P_0 = (0, 0, z_0)$  se tiene que [10]

$$(9) \quad G(P, P_0) = \frac{1}{4\pi |P - P_0|} + \frac{1}{z_0} \frac{1}{4\pi |P - P'_0|} - \frac{w(P, P_0)}{4\pi} - \frac{R^2(P)}{8\pi} + \tilde{C},$$

donde  $P'_0 = (0, 0, z'_0)$ ,  $z'_0 = \frac{1}{z_0}$ ,  $\tilde{C}$  es una constante que se elige de la condición:

$$(10) \quad \int_{\Omega} G(P, P_0) dP = 0,$$

y

$$w(P, P_0) = \ln \left( z'_0 - z + |P - P'_0| \right).$$

Nótese que cuando  $z_0 = 1$ , entonces  $P_0$  corresponde con el polo norte de la esfera. Tendremos entonces el potencial en ese punto. Para el caso en que  $P_0$  no coincida con el polo norte, podemos hacer el cambio de coordenadas para colocar el eje  $z$  en la dirección del punto  $P_0$  y que éste coincida con el polo norte. Así tendremos el **potencial teórico** en cada punto.

#### 4. PROBLEMA DE MINIMIZACIÓN

En la sección anterior se calculó el potencial teórico producido por una fuente dipolar dada por (5).

Para determinar, a partir de las mediciones del potencial  $V_1, V_2, \dots, V_n$  tomadas en los puntos  $P_1, P_2, \dots, P_n$  de la esfera, se utiliza el funcional de mínimos cuadrados:

$$(11) \quad \sum_{k=1}^n |u(P_k) - V_k|^2,$$

donde  $u(P_i)$  está dado por (8).

Supongamos que conocemos la posición del momento dipolar  $a$ . En este caso el funcional (11) puede escribirse en la forma:

$$(12) \quad \| AX - Y \|^2$$

donde  $X = (qd_1, qd_2, qd_3)^t$ ,  $Y = (V_1, V_2, \dots, V_n)^t$  y  $A = (a_{ij})$  está dada por

$$a_{1j} = \frac{\partial G}{\partial x}(a, p_j), \quad j = 1, 2, \dots, n,$$

$$a_{2j} = \frac{\partial G}{\partial y}(a, p_j) \quad j = 1, 2, \dots, n,$$

$$a_{3j} = \frac{\partial G}{\partial z}(a, p_j) \quad j = 1, 2, \dots, n,$$

y  $\| \cdot \|$  denota la norma euclidiana en  $\mathfrak{R}^n$ .

Si conocemos la ubicación del dipolo  $a$ , debemos minimizar (12) para determinar  $q\vec{d}$  el cual es llamado momento dipolar de la fuente dipolar (5). En [1] ha sido probado que el momento dipolar puede determinarse de manera única a partir de la medición sobre la frontera de la región conductora.

Debido a que el problema es mal planteado ya que pequeños errores en la medición pueden producir variaciones sustanciales en la localización del momento dipolar, debemos considerar algún método de regularización tal como el de Tijonov el cual está descrito por [6]:

$$(13) \quad \| AX - Y \|^2 + \alpha \| X \|^2,$$

donde  $\alpha > 0$  es llamado el parámetro de regularización de Tijonov y debe elegirse apropiadamente en función del error en la medición, para garantizar que cuando dicho error tiende a cero la solución aproximada converge a la solución exacta (para datos sin error).

## 5. SIMULACIÓN DE UNA FUENTE EN FORMA DIPOLAR

Las fuentes correspondientes a potenciales evocados o a estímulos de los sentidos básicos pueden ser aproximados de manera eficiente por dipolos de corriente, hecho que deseamos corroborar construyendo un dipolo eléctrico (que simulará a la fuente dipolar) dentro de un medio conductor homogéneo y midiendo los potenciales sobre la superficie de éste medio conductor utilizando la técnica del EEG se obtienen los datos experimentales.

Con esta construcción podremos colocar un dipolo con parámetros conocidos y medir sobre la esfera el potencial producido por esta fuente (problema directo); dada ésta medición podremos identificar, a través del problema de minimización de la sección anterior, los parámetros que caracterizan al dipolo (problema inverso).

Un dipolo lo forman dos cargas de igual magnitud pero de signo contrario separadas por una distancia  $d$ .

Se supone a una carga  $-q$  ubicada en un punto  $r'$  y una carga  $q$  ubicada en un punto  $r' + d$ , como se muestra en la figura (1).

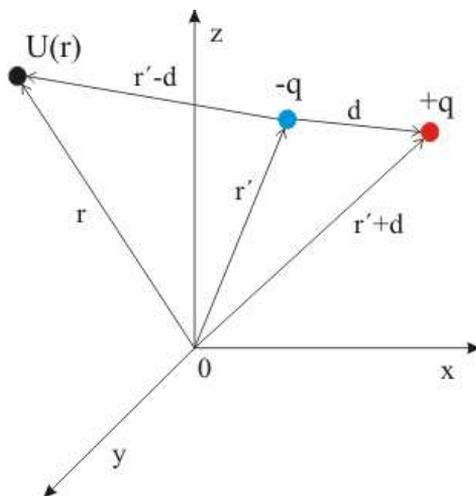


FIGURA 1. Dipolo eléctrico.

Un dipolo puntual no tiene carga neta, no tiene extensión y se caracteriza completamente por su momento dipolar, que es el límite de  $qd$  a medida que  $d$  tiende a cero. Se utiliza el símbolo  $p$  para representar el momento dipolar eléctrico.

$$(14) \quad p = qd.$$

En función del momento dipolar el Campo Eléctrico se expresa como

$$(15) \quad E(r) = \frac{1}{4\pi\epsilon_0} \left\{ \frac{3(r-r') \cdot p}{|r-r'|^5} (r-r') - \frac{p}{|r-r'|^3} \right\}.$$

La distribución del potencial producida por un dipolo puntual es importante ya que esto se mide de manera experimental y se expresa con la siguiente ecuación.

$$(16) \quad U(r) = \frac{1}{4\pi\epsilon_0} \frac{p \cdot (r-r')}{|r-r'|^3}.$$

Para construir el dipolo utilizamos dos Generadores de Van de Graff uno para generar carga positiva y el otro para generar carga negativa; el principio de operación de estas máquinas electrostáticas es el frotamiento de dos materiales diferentes; así se generan grandes cantidades de carga eléctrica que se estará almacenando en el interior de una esfera metálica hueca (de aluminio); para la implementación física del dipolo que nos interesa, se realiza la transferencia de carga eléctrica positiva y negativa respectivamente hacia dos esferas de plata de diámetro menor que representarán a la fuente dipolar, figura (2).

En la parte derecha de la figura (2) se muestra a las dos esferas separadas a una distancia  $d = 5 \text{ cm}$  que tienen una carga  $q_+ = 13.3 \times 10^{-3} \text{ C}$  y  $q_- = -13.3 \times 10^{-3} \text{ C}$

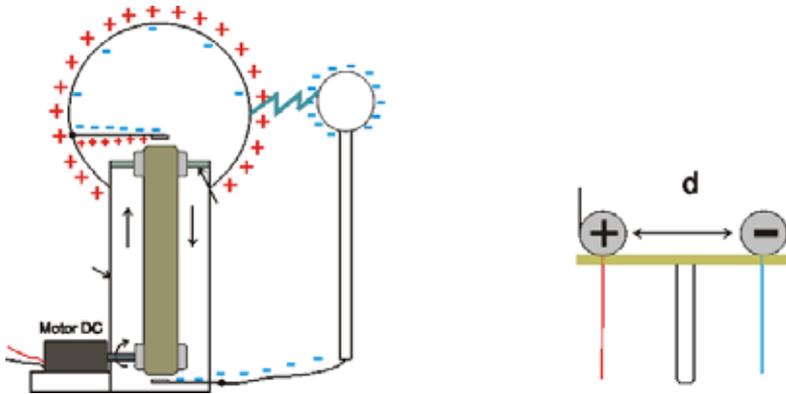


FIGURA 2. Figura de la izquierda Generador de Vann de Graff. Derecha dipolo implementado.

respectivamente, así en términos de los parámetros del sistema físico obtenemos al momento dipolar dado por (14), además conocemos la conductividad del medio  $\sigma = 3 \times 10^{-11} S/m$ ; con esta información y aplicando la definición descrita en (3.1) damos solución al problema directo (PDE) descrito en la definición (2.1) que nos dice que si conocemos  $f(x)$  hallamos  $U(x)$ , que esta dada por la ecuación (8).

La fuente dipolar se introduce en una esfera llena de líquido con conductividad  $\sigma$ , los potenciales experimentales (voltajes) generados por la fuente y la reacción del medio se miden utilizando electrodos comerciales (de los que utilizan para medir EEG) colocados sobre la superficie de la esfera conductora. El diagrama del sistema físico implementado se muestra en la figura (3). De esta manera simulamos la medición sobre el cuero cabelludo (EEG).

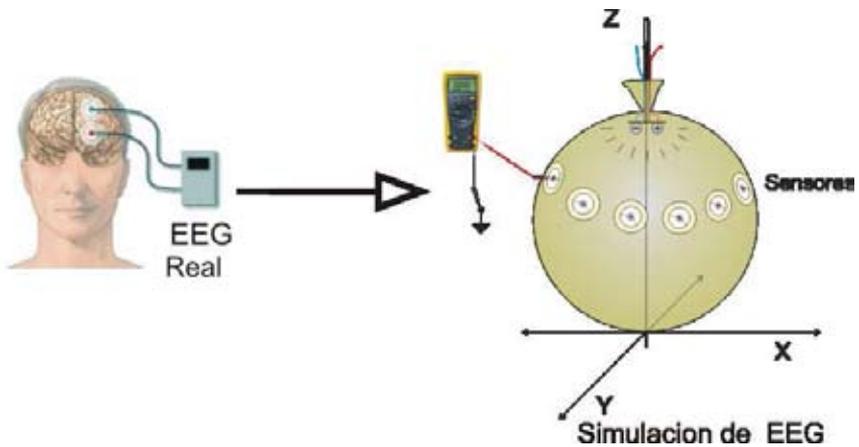


FIGURA 3. Simulación de un EEG real utilizando el sistema físico implementado.

Como se aprecia en la figura (3)(lado derecho) la fuente dipolar esta ubicada en la parte superior del eje de las  $z$  y los puntos donde se hacen las mediciones (sensores) están colocados alrededor de la esfera conductora en puntos  $a_i$ ,  $i = 1, 2, \dots, 10$ .

En la Tabla 1 se muestran los potenciales experimentales medidos, se crea una base de datos que posteriormente se utiliza para obtener el potencial teórico del problema directo (PDE) utilizando la ecuación (8).

Sensores	Potenciales
$a_1$	203 mV
$a_2$	207 mV
$a_3$	202 mV
$a_4$	270 mV
$a_5$	280 mV
$a_6$	205 mV
$a_7$	201 mV
$a_8$	290 mV
$a_9$	280 mV
$a_{10}$	202 mV

TABLA 1. Potenciales(voltajes) medidos sobre la esfera conductora

Los potenciales experimentales varían dependiendo de la ubicación del electrodo registrador con respecto a la fuente; así cuando este se encuentra próximo a la fuente se detecta mayor cantidad de potencial, por el contrario si el electrodo se encuentra lejos de la fuente el potencial detectado será menor como se observa en la Fig 4.



FIGURA 4. Gráfico de potenciales producto de la fuente dipolar

Dado que la unidad de potencial en el Sistema Internacional es el volt (V) las mediciones o potenciales experimentales se obtienen en el orden de milivolt.

## 6. CONCLUSIONES

En este trabajo se plantea el problema de identificar el momento dipolar que corresponde a la localización de la fuente(dipolo) concentrada en el interior del cerebro y para el cual ha sido probado que hay unicidad. Para llegar a este planteamiento se utiliza un modelo de medio conductor en el que la cabeza se considera como un medio homogéneo con conductividad constante.

El potencial teórico producido por esta fuente dipolar (fija) es la solución de un problema de contorno que se expresa por medio de la función de Green la cual se da de forma explícita así como los potenciales experimentales (voltajes) medidos en la superficie de la esfera conductora del sistema físico implementado (Problema Directo).

Para identificar el momento dipolar (ubicación de la fuente) dentro de la esfera conductora se utiliza el funcional de mínimos cuadrados el cual se construye con el potencial teórico antes descrito y el potencial experimental (Problema Inverso).

Se considera un problema mal planteado en el sentido de que pequeños errores en los datos pueden producir variaciones sustanciales en la localización del momento dipolar debe aplicarse un método de regularización como el de Tijonov para obtener de forma estable el momento dipolar.

Al disponer de un sistema físico real es posible realizar mediciones sobre el cuero cabelludo del potencial producido por esa fuente y aplicar el funcional de mínimos cuadrados para identificar el momento dipolar(ubicación de la fuente).

En un futuro se espera que la validez fisiológica de los potenciales evocados que se mide a través de un EEG, pueda ser comprobada mediante la localización de la fuente en el interior del cerebro con ayuda del algoritmo de identificación predefinido en este trabajo, por ello existe un gran interés en seguir buscando la validación física y experimental de estos sistemas dedicados al estudio de la actividad neuronal.

## REFERENCIAS

- [1] A. Amir, *Uniqueness of the generators of brain evoked potential maps.*, IEEE Transactions on Biomedical Engineering. Vol. 41, pp. 1-11., 1964.
- [2] A. Fraguela, M. Morín, Oliveros J. Oliveros *Planteamiento del Problema Inverso de Localización de los Parámetros de una Fuente de Corriente Neuronal en Forma de Dipolo.* Aportaciones Matemáticas, Serie Comunicaciones 25, pp. 41-55. Sociedad Matemática Mexicana 1998.
- [3] A. Fraguela, Oliveros J. Oliveros, Grebennikov A. Grebennikov *Planteamiento Operacional y Análisis del Problema Inverso Electroencefalográfico.* Revista Mexicana de Física, Vol. 47, N. 2, pp. 162-174.,2001.
- [4] A. Fraguela, M. Morín, Oliveros J. Oliveros *Inverse Electroencephalography for volumetric sources.*, Mathematics and Computers in Simulation, Vol. 78, pp. 481-492., 2007..
- [5] A. Fraguela, Oliveros J. Oliveros, M. Morín, Cervantes L. Cervantes *Inverse Electroencephalography for Cortical Sources.*, Applied Numerical Mathematics, Vol. 55, pp.191-203., 2005.
- [6] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems.* Springer Verlag.,1966.
- [7] R. Pascual-Marqui, 1999. *Review of Methods for Solving the EEG Inverse Problems.* International Journal of Bioelectromagnetism. N. 1, Vol. 1., 1999.
- [8] J. Sarvas, *Basic Mathematical and Electromagnetic Concepts of the Biomagnetic inverse Problem.* Phys. Med. Biol., Vol. 32, No. 1, pp. 11-22., 1987.

- [9] S. L. Sobolev, *Partial Differential Equations of Mathematical Physics*. Addison-Wesley Publishing Company., 1964.

Facultad de Ciencias de la Electrónica, BUAP.  
Edificio 182 Ciudad Universitaria Av. Jardines de San Manuel.  
Puebla Pue, México. C. P. 72570.  
Facultad de Ciencias Físico Matemáticas, BUAP.  
Avenida San Claudio y 18 Sur, Colonia San Manuel  
Puebla, Pue., México. C.P. 72570.  
mmorin@ece.buap.mx  
oliveros@fcfm.buap.mx  
blami@ece.buap.mx

# Sesión de Enseñanza e Historia

Quinta Gran Semana Nacional de la Matemática



# UNA ESTRATEGIA DE ENSEÑANZA-APRENDIZAJE

OLGA LETICIA FUCHS GÓMEZ  
MARÍA GUADALUPE RAGGI CÁRDENAS  
FCFM - BUAP

RESUMEN. Pretendemos provocar, tanto en maestros, como en estudiantes, con un problema sencillo, una reflexión sobre su manera de enseñar. Ilustramos, una posible estrategia para una clase de geometría elemental. Este ejemplo es para estudiantes de enseñanza media o media superior. Utilizamos la técnica de preguntas generadoras, con la finalidad de resolver un problema concreto. Además, con esto, contribuimos con los lineamientos del Modelo Universitario Minerva (MUM).

## 1. INTRODUCCIÓN

Con el objetivo de que los alumnos hagan explícitos sus conocimientos al aplicarlos en la solución de problemas y que desarrollen su razonamiento, el presente trabajo propone la implementación de una estrategia de enseñanza, que genere en el alumno no sólo un simple cambio conceptual, sino también metodológico, actitudinal y que ayude a desarrollar su cognición. Para la realización de este trabajo es importante considerar el concepto de cognición como “el conjunto de procesos mentales que tiene lugar entre la recepción de estímulos y la respuesta a éstos”, teniendo en cuenta que estos son procesos estructurales inconscientes y derivan de experiencias del pasado, facilitando la interpretación de estímulos. Existen diferentes definiciones para el aprendizaje. Algunas dicen que es un proceso de adquisición de conocimientos y habilidades de manera que dicho conocimiento se manifieste en un futuro. También se entiende como un producto, resultado de la experiencia o cambio que acompaña a la práctica; un proceso en el que el comportamiento se cambia, y como una función en la que el cambio se origina cuando el sujeto interactúa con la información.

El aprendizaje es un proceso que se desarrolla por niveles, vease [4]. En el caso del adulto podemos hablar de:

1. Sus saberes o maneras de saber hacer en los campos específicos.
2. Sus capacidades (sus métodos y técnicas de trabajo y aprendizaje).
3. Sus recursos estratégicos (conocimiento de sí mismo y su relación dinámica con el entorno).
4. Su motivación, actitudes respecto al aprendizaje.
5. Estos niveles de aprendizaje deben desarrollarse en los alumnos y, la organización del proceso de enseñanza - aprendizaje debe garantizar la atención de la diversidad de intereses y de aptitudes de los alumnos, fomentando hábitos personales de aprendizaje activo.

En el aspecto relacionado con la enseñanza es importante considerar una enseñanza que proporcione a los alumnos experiencias significativas e interesantes que les permitan adquirir conocimientos que de otra forma les resultarían alejados de su

realidad. Los conocimientos que sólo se encuentran en los libros, en lecciones aisladas y desconectadas de la realidad, son conocimientos memorísticos y poco significativos para los alumnos, en este tipo de enseñanza lo menos importante es que el alumno comprenda los conceptos.

El uso de una estrategia requiere de otros componentes cognitivos. Las estrategias precisan disponer de recursos cognitivos para ejercer control más allá de la ejecución de técnicas, así como cierto grado de reflexión consciente (metacognición). Recordemos que la metacognición es el conocimiento sobre la propia cognición, siendo capaz de tomar conciencia de nuestra manera de aprender y comprender los factores que explican que los resultados de una actividad sean positivos o negativos en la actividad intelectual y que las estrategias utilizadas son las idóneas para cada situación de aprendizaje.

Este metaconocimiento es necesario para tres tareas esenciales:

- (i) La selección y planificación de procedimientos eficaces.
- (ii) El control de su ejecución y puesta en marcha.
- (iii) La evaluación del éxito o fracaso obtenido tras la aplicación de la estrategia.

Las estrategias implican una actividad deliberada y controlada por parte del alumno y constan de cuatro fases para su puesta en marcha:

- Fijar metas.
- Elegir la secuencia de acción.
- Aplicación de la estrategia.
- Evaluación de las metas.

Hay que considerar además otros procesos psicológicos que son necesarios para utilizar una estrategia en un dominio de conocimiento dado sin conocimientos temáticos específicos en el área a la que se aplicará la estrategia. El conocimiento conceptual específico es factor determinante de la eficacia en el uso de estrategias de razonamiento y aprendizaje, vease [5]. Las estrategias de control de variables en el razonamiento científico se aprenden paralelamente al contenido al cual se aplican, en el aprendizaje de la física se hace uso de ellas.

Otro tipo de estrategias son las llamadas estrategias de apoyo, que se caracterizan por enfocarse a procesos auxiliares que apoyan el aprendizaje al mejorar las condiciones materiales y psicológicas que producen el aprendizaje, ya que las condiciones ambientales son más favorables, estimula la motivación y la autoestima.

Pero, ¿a qué llamamos estrategia? Una estrategia es el uso deliberado y planificado de una secuencia compuesta de procedimientos dirigida a alcanzar una meta establecida.

Las condiciones didácticas que influyen en el aprendizaje del alumno haciéndolo rutinario o estratégico es un factor importante, si el tipo de tareas de aprendizaje/enseñanza a las que se enfrentan en clase varían en aspectos relevantes. Esto implica una práctica reflexiva por parte del alumno haciéndolo que planifique, seleccione y revise su propia actividad de aprendizaje. Dichas tareas implican situaciones novedosas que requieren nuevos planteamientos, verdaderos problemas que los alumnos deben afrontar de modo estratégico.

Pozo y Postigo (1994) sugieren algunos criterios para hacer que las tareas escolares se planteen como problemas en lugar de como simples ejercicios, estos criterios los dividen en 3 momentos:

- En el planteamiento del problema.

- Plantear tareas abiertas.
- Modificar la manera de definir el problema.
- Diversificar los contextos de aplicación de la estrategia.
- Plantear las tareas en escenarios cotidianos y significativos para el alumno.
- Adecuar la definición del problema a los objetivos de la tarea.
- Utilizar los problemas con fines diversos durante el desarrollo o secuencia didáctica de un tema, evitando que las tareas prácticas aparezcan como demostración o ejemplo de contenidos presentados al alumno previamente.
- Durante la solución del problema.
- En la evaluación.
  - Dar mayor importancia a los procesos de solución seguidos por el alumno que a la respuesta obtenida. Es más importante, evaluar que corregir.
  - Valorar la planeación previa del proceso de solución, así como la reflexión y la autoevaluación del alumno del proceso seguido.

Los criterios antes mencionados consideran que en la medida que las tareas resulten imprevisibles y novedosas, el alumno tenderá a percibir las más como problemas y reflexionará sobre lo que está haciendo.

## 2. DESARROLLO

Para lograr que los alumnos comprendan algunos conceptos geométricos se utilizan preguntas generadoras para conocer y analizar las ideas de los alumnos con respecto a éstos. A estas ideas se les conoce como ideas previas o alternativas y son importantes porque a partir de ese conocimiento, se deben elaborar las estrategias de enseñanza para dar cuenta del progreso conceptual de los alumnos. La construcción de las ideas previas está asociada a explicaciones causales (Pozo, 1989) y a la construcción de esquemas relacionales. Buena parte de las ideas previas son elaboradas a partir de un razonamiento causal directo y pueden ser contradictorias cuando se aplican a contextos diferentes. Las ideas previas interfieren con lo que se enseña en la escuela teniendo como resultado que el aprendizaje sea deficiente, con importante pérdida de coherencia, sin embargo es posible modificarlas por medio de estrategias orientadas al cambio conceptual. Con esta propuesta pretendemos se logre un cambio conceptual con una participación activa y gradual por parte del estudiante y cuidando que el aprendizaje sea significativo además de contribuir al desarrollo cognitivo de ellos propiciando el desarrollo del razonamiento.

## 3. PROBLEMA

### **Un carpintero quiere construir un taburete con la forma de un prisma hexagonal.**

Para el mejor desarrollo del tema, ilustramos diversos tipos de taburetes, incluyendo el de nuestro problema.

## 4. PREGUNTAS GENERADORAS

Pretendemos con las siguientes preguntas, motivar a los estudiantes a participar activamente en la clase, con el fin de que ellos mismos, en el transcurso de la clase, vayan encontrando las respuestas a las siguientes preguntas.

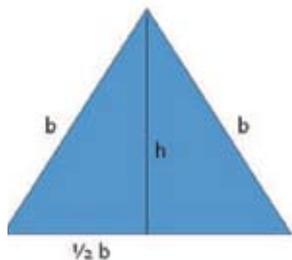
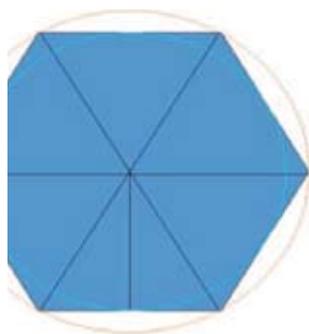
1. ¿Qué forma tiene un prisma hexagonal?
2. ¿Cuáles pueden ser los datos mínimos que debemos tener para poder construir el taburete?
3. ¿Cuánta madera será necesaria?

## 5. CONCEPTOS PREVIOS INVOLUCRADOS

Para dar respuesta a las preguntas anteriores, el estudiante debería ya conocer ciertos conceptos, como son:

- Longitud de segmentos de recta.
- Perpendicularidad entre rectas.
- Figuras planas como: rectángulos, triángulos: equiláteros e isósceles, circunferencias.
- Áreas de esas figuras.
- Hexágono, Prisma, Prisma hexagonal (P.h.)
- Ángulos.

La figura 1 muestra las especificaciones de un hexágono, el cual se usa para formar la base y la tapa del taburete.



**Seis lados iguales**

**Seis triángulos equiláteros**

**Circunferencia circunscrita (C.c.)**

**Circunferencia inscrita (C.i.)**

**Apotema =  $h$  = radio de la C.i.**

**Radio de la C.c. =  $b$  = longitud del lado.**

FIGURA 1. Especificaciones del hexágono

Para contestar a nuestra segunda pregunta, necesitamos conocer:

- La longitud del lado o del apotema y
- La altura  $A$  del prisma (P.h.)

## 6. DISEÑO Y CONSTRUCCIÓN

La forma del taburete obtenido se ilustra en la figura 2.

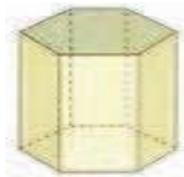


FIGURA 2. Taburete

Con estos datos diseñaremos y construiremos:

- i: Los hexágonos base y tapa del taburete.
- ii: Lados rectangulares del taburete.

Procedamos:

- Dividimos al hexágono en seis triángulos iguales:  $\alpha = 360^\circ/6 = 60^\circ$
- La suma de los ángulos internos de un triángulo es  $\alpha = 180^\circ$ , luego, cada ángulo interno es de  $60^\circ$ . Entonces el triángulo es equilátero.

Si  $b$  = base = longitud del lado y  $h$  = altura = apotema, la relación entre la base y el apotema nos la proporciona el **Teorema de Pitágoras**:

$$h^2 + \left(\frac{b}{2}\right)^2 = b^2.$$

Si el dato que tenemos es  $h$ , despejamos  $b$ .

Calculemos ahora el área de cada triángulo equilátero,

$$A(T) = (\text{base} \times \text{altura})/2.$$

Luego, el área del hexágono  $H$  es:

$$A(H) = \left(\frac{6\sqrt{3}}{4}\right)b^2 = \left(\frac{3\sqrt{3}}{2}\right)b^2.$$

Los seis lados del prisma son rectángulos  $R$  y son iguales. El área de cada rectángulo es  $A(R) = b \times A$ .

Finalmente, tenemos que, el área total del prisma es la suma de las áreas que calculamos:

$$A(T) = 2A(H) + 6A(R) = (3\sqrt{3})b^2 + 6(b \times A).$$

Finalmente, para poder responder la tercera pregunta, necesitamos conocer las dimensiones de las hojas de la madera que utilizaríamos y la mejor manera de aprovecharla, es decir, resolver un problema de optimización. La solución de este problema no lo trataremos aquí.

## 7. CONCLUSIONES

Se desarrolló una estrategia que permite al docente monitorear el cambio conceptual de sus estudiantes partiendo de la identificación de las ideas previas de cada uno de ellos y, promoviendo la reflexión en cada fase de la estrategia. El problema elegido es un problema de la vida cotidiana por lo que es significativo para los estudiantes. Con esto se pretende que esta nueva habilidad que adquiere el estudiante no esté alejada de su realidad y sea un conocimiento que pueda serle útil. Por otra parte se promueven las habilidades de pensamiento partiendo de la observación y elección de variables, de parámetros y análisis de las componentes geométricas. Con este trabajo se pretende crear un banco de estrategias que nos permitan desarrollar actividades constructivistas en el aula, de manera congruente con el Modelo Universitario Minerva, recién instaurado en la BUAP.

## REFERENCIAS

- [1] Pozo, J.I. y Postigo, *La solución de problemas*, Santillana. Madrid, 1994.
- [2] B. Pozo, J. I. *Teorías Cognitivas del aprendizaje*, Madrid, España. Morata Ediciones. 1989.
- [3] Pozo, J.I, M. A. Gómez Crespo, *Aprender y Enseñar Ciencia*, Madrid, Esp. Morata Ediciones. 1998.
- [4] Bartolomé, A. Alonso, C.M. *Análisis y Diagnóstico de los Estilos de Aprendizaje de los Universitarios*, Universidad Complutense. Madrid, 1991.
- [5] Pozo J.I. Perez Echeverría M, *Psicología del Aprendizaje Universitario*, Madrid, Esp. Morata Ediciones. 1994.

Facultad de Ciencias Físico Matemáticas, BUAP.  
Av. San Claudio y 18 Sur  
[www.fcfm.buap.mx](http://www.fcfm.buap.mx)  
[letyfuchs@yahoo.com.mx](mailto:letyfuchs@yahoo.com.mx)  
[gperaggi@fcfm.buap.mx](mailto:gperaggi@fcfm.buap.mx)

# TEOREMA DE EQUIPOTENCIA DE BERNSTEIN

SERGIO ATAYAN GARCÍA BALÁN  
MANUEL IBARRA CONTRERAS  
FCFM - BUAP

RESUMEN. Sean  $M$  y  $N$  dos conjuntos. Si cada uno de ellos es equipotente a un subconjunto del otro, entonces  $M$  y  $N$  son “equipotentes”. Enunciado por Cantor y demostrado por Bernstein, este hermoso teorema ha sido fundamental en el estudio de los números cardinales. Aquí lo demostraremos siguiendo a un autor clásico, E. Kamke, situándonos en el primer tercio del siglo XX, es decir, daremos más importancia a la manera “intuitiva” con la que se comunicaban en ese entonces algunos conceptos fundamentales que al rigor matemático.

## 1. INTRODUCCIÓN

Félix Bernstein nació en 1878 en la misma ciudad que vio nacer a la teoría de conjuntos, Halle, en Alemania. En 1896, como participante de un seminario que impartía Cantor en la Universidad de Halle, Bernstein se ofreció para corregir las pruebas del célebre trabajo de Cantor *Beiträge Begründung der transfiniten Mengenlehre*; notó que la justificación del teorema que nos ocupa no era una prueba rigurosa y cuando Bernstein le mostró a Cantor su demostración, éste quedó tan impresionado que inmediatamente se lo comunicó a Émile Borel quien decidió publicarlo en sus *Leçons sur la théorie des fonctions* en 1898. Ernst Schröder también publicó una prueba de este teorema en ese mismo año de manera independiente. Después se descubrió que contenía un error pero que la idea básica se podía corregir para construir una prueba.

## 2. TEOREMA DE EQUIPOTENCIA

2.1. DEFINICIÓN. Diremos que dos conjuntos  $\mathcal{A}$  y  $\mathcal{B}$  son equipotentes (equivalentes), si y sólo si existe una biyección entre ellos. Este hecho lo denotaremos por  $\mathcal{A} \sim \mathcal{B}$ .

2.2. TEOREMA. Sean  $\mathcal{M}$  y  $\mathcal{N}$  dos conjuntos. Si cada uno de ellos es equipotente a un subconjunto del otro, entonces  $\mathcal{M}$  y  $\mathcal{N}$  son equipotentes.

DEMOSTRACIÓN. La prueba del teorema se reduce a demostrar la siguiente proposición:

2.3. PROPOSICIÓN. **(P)** Si  $\mathcal{M}$  es equipotente a un subconjunto  $\mathcal{M}_2$  ( $\mathcal{M}_2 \subseteq \mathcal{M}$ ), entonces  $\mathcal{M}$  es también equipotente a cada conjunto  $\mathcal{M}_1$  “entre”  $\mathcal{M}$  y  $\mathcal{M}_2$ , es decir, a cada conjunto  $\mathcal{M}_1$  que cumple  $\mathcal{M}_2 \subseteq \mathcal{M}_1 \subseteq \mathcal{M}$ .

En efecto, veamos la demostración del teorema:

Supongamos que  $\mathcal{M}_1$ ,  $\mathcal{N}_1$  son subconjuntos de  $\mathcal{M}$  y  $\mathcal{N}$  respectivamente, tal que  $\mathcal{M} \sim \mathcal{N}_1$  y  $\mathcal{N} \sim \mathcal{M}_1$ . Entonces existen funciones  $\alpha$ ,  $\beta$  biyectivas,  $\alpha : \mathcal{N} \rightarrow \mathcal{M}_1$ ,  $\beta : \mathcal{M} \rightarrow \mathcal{N}_1$ . Sea  $\mathcal{M}_2 \subseteq \mathcal{M}_1$  la imagen de  $\mathcal{N}_1$  bajo  $\alpha$ , entonces tenemos que  $\mathcal{M}_2 \subseteq$

$\mathcal{M}_1 \subseteq \mathcal{M}$  y  $\mathcal{M} \sim \mathcal{N}_1 \sim \mathcal{M}_2$ . Ahora, si usamos **(P)** tenemos que  $\mathcal{M} \sim \mathcal{M}_1$  y por tanto  $\mathcal{M} \sim \mathcal{N}$ . □

Por lo tanto, sólo nos resta demostrar **(P)**. Supongamos que  $\mathcal{M}_2 \subset \mathcal{M}_1 \subset \mathcal{M}$  y  $\mathcal{M} \sim \mathcal{M}_2$ . Denotemos  $\mathcal{M}_2 = A$ ,  $\mathcal{M}_1 - \mathcal{M}_2 = B$ ,  $\mathcal{M} - \mathcal{M}_1 = C$ .

La proposición **(P)** quedaría de esta forma:

2.4. PROPOSICIÓN. (**P\***) Si  $A, B, C$  son conjuntos ajenos, entonces  $A \cup B \cup C \sim A \Rightarrow A \cup B \cup C \sim A \cup B$ .

Como  $A \cup B \cup C \sim A$ , existe  $\gamma$  función biyectiva,  $\gamma : A \cup B \cup C \rightarrow A$ . Sean  $A_1, B_1, C_1$  los subconjuntos de  $A$  que son las imágenes de  $A, B, C$  respectivamente bajo  $\gamma$ , es decir,  $\gamma(A) = A_1$ ,  $\gamma(B) = B_1$ ,  $\gamma(C) = C_1$ . De esta manera:

$$(1a): A_1 \cup B_1 \cup C_1 = A$$

$$(1b): A \sim A_1, B \sim B_1, C \sim C_1$$

$A_1, B_1, C_1$  son ajenos.

Ahora bien, como  $\gamma(A) = A_1$ , se sigue de (1a) que  $A_1, B_1, C_1$  son enviados mediante  $\gamma$  a subconjuntos  $A_2, B_2, C_2$  de  $A_1$  respectivamente. Y ocurre que:

$$(2a): A_2 \cup B_2 \cup C_2 = A_1$$

$$(2b): A_1 \sim A_2, B_1 \sim B_2, C_1 \sim C_2$$

$A_2, B_2, C_2$  son ajenos.

El siguiente paso nos conduce a los conjuntos  $A_3, B_3, C_3$  que cumplen:

$$(3a): A_3 \cup B_3 \cup C_3 = A_2$$

$$(3b): A_2 \sim A_3, B_2 \sim B_3, C_2 \sim C_3$$

$A_3, B_3, C_3$  son ajenos.

Etcétera...

Debido a que  $A \sim A_1 \sim A_2 \sim \dots$ , el proceso no termina. Notemos que, en particular,

$$C \sim C_1 \sim C_2 \sim \dots \quad (I).$$

Sea  $D = A_1 \cap A_2 \cap \dots$  ( $D$  pudiera ser vacío), entonces

$$\begin{aligned} A \cup B \cup C &= D \cup B \cup C \cup B_1 \cup C_1 \cup B_2 \cup C_2 \cup \dots \\ A \cup B &= D \cup B \cup C_1 \cup B_1 \cup C_2 \cup B_2 \cup C_3 \cup \dots \end{aligned}$$

En el lado derecho de la primera igualdad, todos los términos son mutuamente excluyentes, y esto mismo también es cierto para la segunda igualdad. (**P\***) quedará demostrado si logramos establecer una biyección entre cada término del lado derecho de la primera igualdad, y el término situado precisamente abajo en la segunda igualdad. Sin embargo, esto queda garantizado por (I).

#### REFERENCIAS

- [1] E. Kamke, *Theory of Sets*, Dover Publications, Inc., New York. 1950.

Facultad de Cs. Físico Matemáticas, BUAP.  
Río Verde y San Claudio, Col. San Manuel,  
Puebla, Pue. C.P. 72570.  
darkbizhop@hotmail.com  
mibarra@fcfm.buap.mx

# LA SIMETRÍA DINÁMICA EN LA OBRA DE JOSÉ CLEMENTE OROZCO

EMMA GARRIDO SÁNCHEZ<sup>1</sup>

IVÁN RUIZ<sup>2</sup>

PEDRO TOLENTINO ESLAVA<sup>3</sup>

<sup>1</sup>COLEGIO DE HISTORIA - FACULTAD DE FILOSOFÍA Y LETRAS - BUAP

<sup>2</sup>PROGRAMA DE SEMIÓTICA Y ESTUDIOS DE LA SIGNIFICACIÓN - BUAP

<sup>3</sup>FCFM - BUAP

RESUMEN. La geometría no sólo es el puente entre las matemáticas y el arte, sino también proporciona un criterio de belleza denominado “sección áurea”, conocido desde la antigua cultura griega. Esta proporción se encuentra en la naturaleza, en la arquitectura y en la pintura. En los años 20 del siglo pasado, Jay Hambidge estudió con detalle la coherencia de estas proporciones geométricas en la cerámica clásica griega, llamando a esta teoría *simetría dinámica*. José Clemente Orozco, influenciado por este autor, como lo reconoce en su *Autobiografía*[1], sigue fielmente los cálculos de las proporciones geométricas en la elaboración de algunos de sus murales. En este trabajo se presenta un recorrido histórico sobre la influencia de esta teoría en el pintor y se muestra cómo se puede apreciar la simetría dinámica en su obra mural. Para ello se utiliza el Análisis del Discurso como método de análisis del texto visual.

## 1. INTRODUCCIÓN

El interés del presente trabajo surgió en un primer acercamiento al Análisis del Discurso en el Colegio de Historia, cuyas enseñanzas hacen reflexionar sobre el trabajo del historiador. El fundamento del método del Análisis del Discurso es el texto, y el método aspira a llegar al nivel más profundo de la significación a través de la deconstrucción del texto sea éste visual o verbal, para luego llevar a cabo una reconstrucción del mismo. Entendiéndose como deconstrucción la descomposición analítica de los elementos que constituyen una estructura conceptual.

Al seleccionar previamente a *Catarsis* como objeto de estudio en el curso, fue necesario investigar al pintor y también conocer otros murales que ayudaran a definir el objeto de análisis. Aún conociendo otras obras consideradas por críticos e historiadores del arte como las más importantes del artista, se decidió que *Catarsis* fuese la obra que aislaría como texto de análisis, no nada más por el tema o su composición y discurso pictórico, sino también por la osadía de Orozco de pintar un mural cuyos motivos iconográficos pudieran resultar agresivos y grotescos para el espectador, en uno de los más trascendentales Museos y Teatro Nacional de la ciudad de México inaugurado en 1934: Bellas Artes.

El primer autor, Emma Garrido, decidió utilizar el método del Análisis del Discurso para analizar a *Catarsis* ya que el estudio de la imagen es realizado por críticos o historiadores del arte, pero no por historiadores que sólo la utilizan para legitimar lo escrito o como simples ilustraciones, práctica que ella realizaba antes



FIGURA 1. Mural *Catarsis*. 1934-1935[2].

de conocer el Análisis del Discurso cuyos principios la hicieron reflexionar sobre la actividad historiográfica. Estudiar a *Catarsis*, es tratar de entender, de comprender a ese “Otro”, desde un tiempo presente, de penetrar en su interior, de tratar de mostrar ese pasado que subordina y emociona descubrirlo, es tratar de darle un sentido. . . para que no pierda su condición de arte.[3]

Al plantearse la hipótesis de que Orozco no abandonó la simetría dinámica después de los años treinta, conllevó a investigar también sobre la simetría dinámica para poder corroborar o desmentir la hipótesis.

El presente trabajo forma parte de la investigación a desarrollar como tesis de Licenciatura en Historia titulada *Historia, Ideología y Discurso Pictórico en el mural Catarsis*, proyecto que concluirá el año próximo.

## 2. ALGUNAS CONSIDERACIONES SOBRE EL MÉTODO

La historiografía ha ido evolucionando de manera gradual desde su nacimiento (siglo XIX); sin embargo, no se ha puesto a la vanguardia respecto a métodos y técnicas en diversas instituciones educativas, sobre todo por la reticencia en historia a utilizar métodos de otras disciplinas, avocándose sólo a la interpretación o actualización de las *fuentes*, por lo que es necesario considerar lo planteado por François Furet, de que el historiador contemporáneo no debe constreñirse a delimitar su objeto de análisis, definir sus hipótesis, constituir, describir o interpretar sus fuentes, tampoco limitarse a realizar un análisis de una sola serie cronológica, sino reagrupar varias series que le puedan servir para proponer la “interpretación de un sistema, o de un subsistema. . . , para poder realizar el análisis, comprensión, interpretación y explicación, que conlleva a desarrollar un trabajo científico interdisciplinario” [4].

Al utilizar *documentos* de otras disciplinas, como son la estadística, la lingüística, la psicología, la semiótica, etc., para hacer historia, hasta el más mínimo *vestigio* hecho por el hombre debe ser considerado documento para la historia. . . , ya que la historia se edifica, se construye, sin exclusión, con todo lo que el ingenio de los hombres pueda inventar y combinar para suplir *el silencio, los ocultos de los textos*, o en su caso, los estragos del *olvido*. . . [5] de los autores, o del pintor en este caso.

Cabe mencionar que el Análisis del Discurso tiene como soporte la semiótica cuyo objetivo es el estudio de los sistemas de lenguaje, ya sea visual o escrito, sin embargo, sin pretender ser semiotista, es necesario apoyarse en los planteamientos de la semiótica de lo visual, cuyo interés se centra en el estudio de las relaciones

internas de los documentos, sean éstos impresos o visuales. La semiótica de lo visual está orientada hacia lo social, pero sobre todo hacia lo fenomenológico y lo cognitivo, por lo que sus *corpus* son muy extensos, entre los que se encuentran: el cine, la televisión, y por supuesto la pintura. . . así como temas relacionados con la publicidad como son: la foto, la ropa, la caricatura, el cuerpo y los espectáculos naturales. . . [6] Entendiéndose por *corpus* el “conjunto de datos o textos de un mismo tipo que sirven de base a una investigación” [7]. Así pues, estudiar los sistemas de lenguaje del texto conlleva al análisis del discurso y por consiguiente a la escritura.

La obra de Orozco es muy amplia, por lo que, habiendo seleccionado el objeto de estudio que es el mural *Catarsis*, fue necesario conocer otras obras entre las que se encuentran: los murales de la Escuela Nacional Preparatoria hoy Ex-Colegio de San Ildefonso, *Omniciencia* ubicado en la Casa de los Azulejos, hoy Sanborns, en el Distrito Federal; los murales del Palacio de Gobierno, Ex-Hospicio Cabañas y Paraninfo de la Universidad de Guadalajara en Jalisco, sin embargo, se confirmó aislar a *Catarsis* como texto de análisis.

*Catarsis* pertenece a un *corpus* de obras de Orozco, por lo que, para situarlo en un periodo, se consideró hacer la delimitación para el presente trabajo, de la siguiente manera: delimitación espacial, ciudad de México, periodo 1920-1939 y el Palacio de Bellas Artes, como la institución productora y expositora del mismo, donde se encuentra ubicado físicamente desde 1934, fecha de su elaboración.

Como lo plantea De Certeau, la historiografía consiste en construir representaciones con material del pasado, situarse en la frontera del presente, incluir la tradición y explotarla con métodos nuevos, por lo que al analizar a *Catarsis* se pretende reconstruir la historia a través del método seleccionado que es el trabajo del historiador; de tal modo, que el estudio de una imagen mediante el método del Análisis del Discurso Histórico plantea que desde el inicio se interrogará al mural *Catarsis*, cuestionándolo, utilizando una forma inversa a la que se acostumbra hacer historia, [8] ya que por lo general, cuando se estudia una imagen, se analiza y “se lee la iconografía basándose en fuentes escritas, convirtiendo a la imagen como simples ilustraciones. . . y no para dar nuevas respuestas o plantear nuevas cuestiones.” Se intentará el camino inverso: “hacer hablar” a *Catarsis*, y buscar el apoyo en textos escritos [9] para obtener respuestas a las preguntas planteadas e ir corroborando o desmintiendo las hipótesis, lo cual condujo a investigar sobre la simetría dinámica.

### 3. LA SIMETRÍA DINÁMICA.

Debido a que el tema que nos ocupa es la simetría dinámica en la obra de Orozco, el pintor cita en su *Autobiografía*, que en 1930 utilizó los principios geométrico-estéticos de la simetría dinámica desarrollados en las investigaciones de Jay Hambidge. Hambidge fue un geómetra americano financiado por la Universidad de Yale para realizar sus investigaciones en Grecia y museos de toda Europa. Realizó investigaciones sobre la simetría dinámica en el Partenón, el templo de Apolo en Arcadia, el templo de Zeus en Olimpia, los templos de Egina y de Sunium cerca de Atenas. De igual manera llevó a cabo estudios sobre vasijas griegas [10]. Sus obras más importantes son: *The Elements of Dynamic Symmetry*, *Dynamic Symmetry: The Greek Vase* y la revista *The Diagonal*. Hambidge murió en 1928 [11]; sin embargo, sus planteamientos influyeron en arquitectos y pintores que posteriormente desarrollaron importantes obras.

Pero, ¿en qué consiste la simetría dinámica? La simetría dinámica está basada en la sección áurea o proporción divina. Esta proporción, que corresponde a la partición más simple de una magnitud en dos partes desiguales o partición más lógica, es lo que Euclides en el libro VI, proposición 30, de los *Elementos*, plantea como “dividir una recta dada en extrema y media razón”, y define así al inicio del libro (Def.3): “Se dice que una recta está dividida en extrema y media razón, cuando la totalidad del segmento es al segmento mayor como el segmento mayor es al menor” [12].

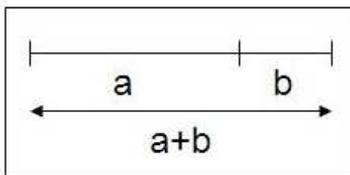


FIGURA 2. Cálculo de la proporción dorada.

En la figura 2, si hacemos  $a = 1$ ,  $a + b = \Phi$ ,  $b = \Phi - 1$ , entonces  $\frac{a+b}{a} = \frac{a}{b}$  se transforma en  $\frac{\Phi}{1} = \frac{1}{\Phi-1}$ .

Simplificando esta última igualdad se obtiene la ecuación cuadrática  $\Phi^2 - \Phi - 1 = 0$ , cuya solución, aplicando la fórmula general para resolver ecuaciones de segundo grado, es:  $\Phi_1 = \frac{1-\sqrt{5}}{2}$ ,  $\Phi_2 = \frac{1+\sqrt{5}}{2}$ .

De esta manera, escogiendo la raíz o solución positiva porque  $\Phi$  representa una longitud, el número de oro es  $\Phi = \frac{1+\sqrt{5}}{2} = 1.618$ ,  $a = 1$ ,  $a + b = 1.618$  y  $b = 0.618$

Cuando la totalidad del segmento constituye la unidad, la longitud del segmento mayor es 0,618 y la del segmento menor es 0,382. [13]

Una manera más ilustrativa de obtener esta proporción es partiendo del segmento mayor como muestra la figura 3.

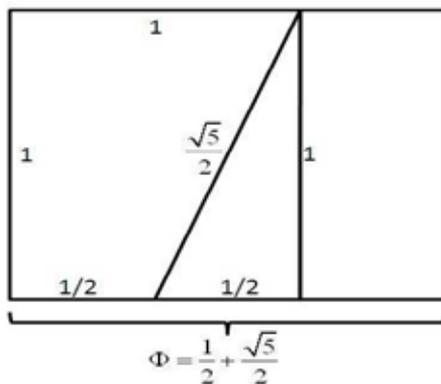


FIGURA 3. Construcción de la sección áurea a partir del segmento mayor.

La relación  $a/b$ , resultante de la “división de una recta en media y extrema razón”, ha recibido diferentes denominaciones en el transcurso del tiempo, pero

las definitivas le fueron otorgadas en el Renacimiento. Luca Pacioli la calificó como la *Divina Proporción*, en su obra del mismo nombre publicada en Venecia en 1509. Según Pacioli, Leonardo de Vinci fue el ilustrador de *Divina Proporción* y es precisamente a él a quien se le atribuye la otra denominación con que es conocida esta proporción: *sectio aurea* (sección áurea) de donde provienen los nombres de sección de oro o número de oro.[14]

Uno de los primeros documentos escritos sobre las proporciones humanas se pueden encontrar en el siglo I con Marcos Vitruvius Pollio, arquitecto y escritor romano, quien en su obra *Diez libros sobre arquitectura* planteó “que los templos para ser magníficos, se construyan análogos al cuerpo humano bien formado, en el cual, dice, existe una perfecta armonía entre todas las partes.”[15]

En el siglo XII destaca la figura de Leonardo de Pisa, llamado Fibonacci, autor de un importante tratado, el *Liber Abaci*, donde, entre otros problemas teóricos y prácticos, aparece una serie de números, la serie de Fibonacci, en la que cada término es igual a la suma de los dos precedentes, propiedad aditiva que comparte con la serie  $\Phi$ , con la que le unen otros lazos pues la razón entre dos de sus términos, consecutivos tiende hacia un límite, que es precisamente  $\Phi$ . Esta serie aritmética y geométrica precisa está relacionada con problemas de crecimiento y desarrollo vegetal y animal.[16]

En el periodo del Renacimiento, Leonardo da Vinci ilustró el hombre de Vitruvio con las proporciones áureas agregadas donde muestra cómo las partes del cuerpo comparten proporciones comprendidas en el rango de la sección áurea y del triángulo pitagórico, ayudando a comprender la conexión, la armonía y el ritmo que existe entre las partes óseas y mentales de los seres humanos. Posteriormente Alberto Durero también realizó estudios sobre las proporciones humanas en niños y adultos.[17]

De lo anterior se puede mencionar que, en los siglos XX y XXI, arquitectos, ingenieros, pintores, historiadores del arte e historiadores se están planteado observar que la naturaleza es proveedora de formas y patrones geométricos que han ayudado al ser humano a resolver problemas constructivos y de diseño. Los historiadores tratan de acercarse tanto a la naturaleza, a los animales y a los seres humanos no nada más como entes del pasado para exhumarlos sino como objetos de estudio presentes, de poder fijar su mirada en lo más profundo que lo que a primera vista pueda captar. Es detenerse a observar lo que lo rodea tratando de plantearse interrogantes y poder obtener respuestas.

#### 4. OROZCO Y LA SIMETRÍA DINÁMICA

La teoría matemática basada en la sección áurea o la sección de oro es lo que Hambidge renombró como Simetría Dinámica, la cual en 1920 y 1930 tuvo auge tanto en Europa como en Estados Unidos. Orozco, influenciado por Hambidge, aplicó estrictamente los principios de la Simetría Dinámica en los murales de la New School for Social Research realizados en 1930, sin embargo también menciona que:

“Después de la pintura de la New School abandoné los métodos tan rigurosos y científicos de la simetría dinámica, pero guardé lo que había de fundamental e

inevitable en lo aprendido, para forjar nuevos métodos de trabajo. Tuve la explicación de muchos errores anteriores encontré, nuevos caminos.”[18]

Considerando lo antes citado por el artista, en 1930 pintó al fresco un mural en la New School for Social Research o Escuela de Investigaciones Sociales, en Nueva York, donde representa la mesa de la fraternidad universal cuyo objetivo fue incluir a personas de todas las razas. En los muros laterales pintó alegorías de la revolución mundial donde se puede observar a Gandhi, Carrillo Puerto y Lenin. Así también un grupo de esclavos y un grupo de obreros entrando a su casa después de su jornada laboral. En muro exterior del salón, una alegoría de las ciencias y las artes donde aplicó los principios de la Simetría Dinámica. De 1932 a 1934 pintó el mural con el tema *Quetzalcóatl* en el Darmouth College[19], sin embargo, en los años antes citados desarrolló algunos bosquejos donde se puede observar la sección dorada. Ver figura 4.

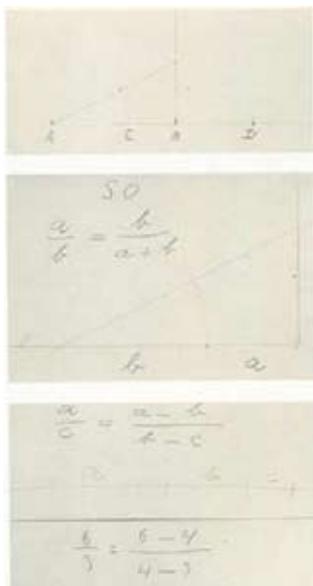


FIGURA 4. Apuntes de Orozco sobre la sección de oro.[20]

De 1934 a 1935 con motivo de la inauguración del Palacio de Bellas Artes, por invitación de don Antonio Castro Leal, Orozco pintó al fresco en el Palacio de Bellas Artes, el mural *Katharsis* o *La Katharsis*[21]. Dado que esta obra es el objeto de análisis, siguiendo la disposición de los motivos iconográficos, es decir de las figuras, se procedió a trazar el cuadrado y su diagonal y la diagonal de la mitad del cuadrado, encontrándose que en la distribución se cumple la proporción dorada: La razón de los segmentos  $AB/AC$ . La figura 5 muestra por separado las diagonales debido a que los murales son considerados patrimonio histórico y no se permite deformar o rayar la imagen.

En 1935 se mudó a Guadalajara, cerca de su ciudad natal Zapotlán El Grande hoy Ciudad Guzmán, Jal. De 1936 a 1939 lleva a cabo los murales en el Paraninfo de la

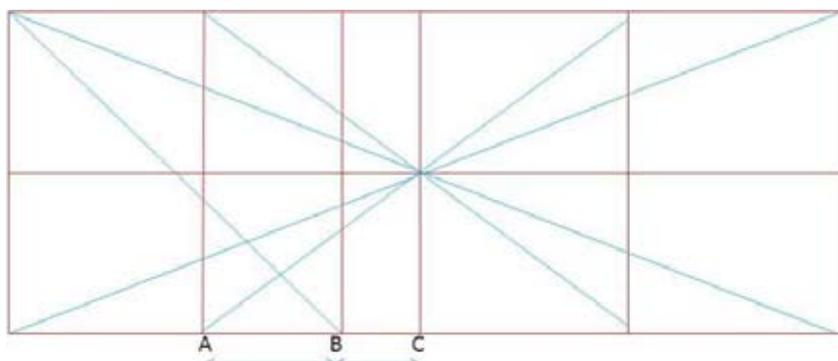


FIGURA 5. Diagonales para el mural *Catarsis*

Universidad de Guadalajara, Palacio de Gobierno y en el Ex-Hospicio Cabañas[22], cuyas obras son consideradas las más significativas del pintor debido a la composición pictórica y el contenido temático, dejando evidencias de la aplicación de las lecciones de Hambidge, como se puede observar en la figura 6. Remitiendo al lector al libro *Los elementos de Simetría Dinámica*[23], para hacer un comparativo visual, se puede constatar cómo Orozco utilizó las lecciones 1 y 8: El rectángulo raíz de dos inscrito en una modificación del rectángulo raíz de tres.



FIGURA 6. Detalle de El Maestro. 1936-1939[24].

Si bien Orozco hace alusión en su *Autobiografía* que en 1930 “abandonó” la simetría dinámica, se invita al lector a observar los murales realizados por el artista en la Escuela Nacional Preparatoria hoy Ex-Colegio de San Ildefonso, en los cuales están contenidos los principios de la simetría dinámica y también a considerar que fue en los años veinte y treinta que la simetría dinámica se diseminó por todo Europa y Estados, por lo que se puede inferir que, como dibujante y perito topógrafo, la geometría, el dibujo aprendido durante su formación en la Escuela de Ingenieros

de San Jacinto y en la Academia de San Carlos, su capacidad crítica como caricaturista, su conocimiento de la historia y de la sociedad, fueron determinantes para que pudiera pintar esas magnas obras que lo sitúan como uno de los pintores mexicanos más importantes, sin omitir mencionar el contenido filosófico del cual están impregnadas estas obras como lo explica Fausto Ramírez.[25]

### CONCLUSIONES

La fascinación de la civilización por las formas geométricas ha llevado a la humanidad a la búsqueda de criterios de belleza universales y ha encontrado uno en la proporción áurea. Ésta ha sido permanentemente estudiada debido a su presencia en la naturaleza, las ciencias y las artes.

Jay Hambidge fue un entusiasta promotor de la sección áurea y con sus investigaciones en los años veinte le dió un gran impulso tanto en Europa como en Estados Unidos. Llamó simetría dinámica a la teoría basada en esta proporción y Orozco la volvió a llamar sección de oro (SO) como se aprecia en sus trazos presentados en la figura 4.

Así pues, se muestra que Orozco no abandonó la simetría dinámica ya que, de hacerlo, significaba abandonar la Geometría misma que es la base para el desarrollo de la pintura y otras artes plásticas. Su deseo de lograr un arte universal lo mantiene fiel a los conocimientos obtenidos en el estudio de la simetría dinámica y a una tendencia a la estilización geométrica presente en gran parte de sus obras.

### AGRADECIMIENTOS

Los autores expresan su agradecimiento al Centro Nacional de Investigación, Documentación e Información de Artes Plásticas (CENIDIAP), la Biblioteca de las Artes y al Departamento Jurídico del INBA por la autorización otorgada para la publicación de las imágenes: Mural *Catarsis* y detalle de *El Maestro Creador y Rebelde*. Los autores (EGS e IR) agradecen a la BUAP el apoyo financiero para el desarrollo de esta investigación mediante el Proyecto VIEP *Abstracción y gestualidad: Las marcas del pintor*.

### REFERENCIAS

- [1] J. C. Orozco, *Autobiografía*, Ediciones Era, México, 1999, pp. 99-102
- [2] "REPRODUCCION AUTORIZADA POR EL INSTITUTO NACIONAL DE BELLAS ARTES Y LITERATURA, 2009". Permiso No. 119
- [3] M. De Certeau, "La operación historiográfica" en *La Escritura de la Historia*, Universidad Iberoamericana, México, 2006, p. 19.
- [4] F. Furet, "Lo cuantitativo de la historia" en *Hacer la Historia: Nuevos Problemas*, Editorial Laia/Barcelona, 1974, p. 66
- [5] L. Febvre, *Combates por la historia*, Editorial Ariel, S.A., Barcelona, 1974, p. 30
- [6] J.-M. Klinkenberg, "La semiótica visual: grandes paradigmas y tendencias de línea dura" en *Semiótica de lo visual*, Tópicos del Seminario, No. 13, BUAP, Puebla, 2005, pp. 20-21
- [7] *Diccionario Panhispánico de Dudas*, Real Academia Española, Santillana Ediciones Generales, S.L. Bogotá, p. 183.
- [8] De Certeau, Op Cit. p. 68-70
- [9] P. Burke, *Visto y no Visto, el uso de la imagen como documento histórico*. Editorial Crítica, Madrid, p. 12
- [10] Orozco, Op Cit. p. 100
- [11] <http://www.hambidge.org>
- [12] P. H. Michel (apud Bonell, Carmen, p. 16)

- [13] C. Bonell, *La Divida Proporción. Las formas geométricas*, Alfaomega Grupo Editor, S.A. de C.V., México, 2006, p. 16
- [14] L. Pacioli, (apud Bonell, Carmen, p. 17 )
- [15] G. Doczi, *El poder de los límites*. Editorial Troquel, Buenos Aires, 2003, p. 93
- [16] C. Bonell, Op Cit. p. 28
- [17] G. Doczi, Op Cit. p. 95-96
- [18] J. C. Orozco, Op Cit. p 104.
- [19] *Ibidem* p 99.
- [20] R. Tibol, José Clemente Orozco. CUADERNOS. CULTURA-SEP. México, 1983, p. 146
- [21] J. C. Orozco, Op Cit. p 112.
- [22] *Idem*
- [23] J. Hambidge, *The Elements of Dynamic Symmetry*, Dover Publications, 1967. pp 18, 48.
- [24] “REPRODUCCION AUTORIZADA POR EL INSTITUTO NACIONAL DE BELLAS ARTES Y LITERATURA, 2009”. Permiso No. 133
- [25] F. Ramírez, “Artistas e iniciados en la obra mural de Orozco” en *OROZCO: UNA RELECCION*. UNAM, México, 1983, pp. 61-102 passim

<sup>1</sup>Juan de Palafox y Mendoza No. 211, Col. Centro

C. P. 72000, Puebla, Pue.

<sup>2</sup>3 Oriente No. 312, Col. Centro

C. P. 72000, Puebla, Pue.

<sup>3</sup>Av. San Claudio y 18 Sur, Ciudad Universitaria

C. P. 72570, Puebla, Pue.

`garrido_emma@hotmail.com`

`tolenti@fcfm.buap.mx`

`soldadero@gmail.com`



# EL CÁLCULO

RAÚL LINARES GRACIA  
FCFM - BUAP

RESUMEN. Presentamos un estudio acerca de la noción de conjunto de los números reales y del infinito, haciendo un análisis histórico, ligando problemas, conocimientos y herramientas disponibles en determinados períodos históricos. Partiendo de la transformación que sufrió la geometría y el análisis, al ir eliminando de los razonamientos las ideas intuitivas que le atribuían a los números reales, por ejemplo la de ser continuos como lo hace Cardano al hallar las raíces positivas de una ecuación cúbica, para considerar el valor demostrativo en los elementos del pensamiento puro, concatenamos esto y el estudio del infinito para hablar de otro sistema axiomático, el de los conjuntos. Considero que el tipo de análisis que se realiza es de utilidad en la enseñanza de la matemática ya que para tener un mejor entendimiento acerca de una teoría matemática es necesario regresar a las ideas iniciales, así el estudiante descubrirá las razones históricas y lógicas que dan origen a los conceptos, favoreciendo la adquisición de una mayor objetividad.

## 1. INTRODUCCIÓN

Durante la práctica docente muchas veces tenemos que hacer uso de los números reales y de alguna interpretación geométrica de estos o de otros hechos, por ejemplo: al calcular la raíces de una ecuación, cuando buscamos los máximos y mínimos de una función o cuando interpretamos el teorema de Bolzano, asumimos que la recta numérica es continua y que los números reales son completos, es decir no tienen agujeros, además de establecer una correspondencia entre números reales y puntos. En este trabajo hacemos un análisis histórico del sistema de los números reales, en dirección de su completitud y del infinito.

## 2. ANTECEDENTES

Gracias a la escritura cuneiforme sobre tablillas de arcilla, se tiene conocimiento de que en la antigua Babilonia (entendiendo esta por la civilización desarrollada entre los ríos Tigris y Eufrates, entre el 2000 a. C. y el 200 a. C.) se resolvieron problemas matemáticos concretos y casos especiales diversos sin ningún tipo de formulación general. La actividad intelectual desarrollada por esta civilización perdió su amplio impulso mucho antes de la era cristiana. A pesar de esto, la cultura babilónica ya había sembrado el germen científico en las culturas a lo largo del mediterráneo y de ellas la cultura helénica fue la que más destacó. De la época de Tales de Mileto a Euclides de Alejandría, los griegos construyeron el imperio científico más importante que a la fecha existe. Este es el imperio de las matemáticas. Sabemos de la existencia de dos problemas críticos concernientes a la relación entre lo discreto y lo continuo: uno fue el descubrimiento de los incommensurables y el otro fueron las paradojas de Zenón. El descubrimiento de los incommensurables, condujo a la teoría de la proporción, que se le acredita a Eudoxo,

este descubrimiento es una de las contribuciones más significativas en la historia, y a la primera fundamentalización de las matemáticas. Sin embargo, a mi parecer son las paradojas de Zenón las que son centrales en el desarrollo de los antiguos fundamentos de las matemáticas y que con muchos siglos de diferencia ocupan un lugar muy importante en la segunda fundamentalización de las matemáticas.

La historia de las paradojas de Zenón es en gran medida la historia de los conceptos de continuidad, de infinito e infinitesimal. En un análisis más cuidadoso, debemos examinar el concepto de continuidad para Aristóteles y también llegar al concepto de infinito.

### 3. EL CONTINUO PARA ARISTÓTELES

En diferentes lugares de los libros “Física” y “Metafísica”, Aristóteles trata el continuo; aquí sólo menciono algunos. En su libro VI de “Física”, define lo continuo: “Por continuo me refiero a aquello que es divisible en partes siempre divisibles.”

Para Aristóteles los indivisibles no pueden ser divididos en partes, lo que significa que no tiene partes, un indivisible no puede tener un extremo o un límite, ya que eso sería una parte.

De aquí que indivisibles no pueden satisfacer la condición de ser contiguos, es decir, no tienen límites para estarse tocando, y consecuentemente no pueden ser continuos. Aristóteles concluye; el continuo no puede estar hecho de indivisibles (libros V, VI, “Física”), es imposible que algún continuo exista a partir de componentes divisibles, como por ejemplo una línea a partir de puntos, si es que la línea es continua y el punto indivisible.

### 4. ARITMÉTICA Y GEOMETRÍA

Aristóteles clasifica las matemáticas en aritmética y geometría, ciencias separadas y no subordinadas requiriendo definiciones y teoremas por separado. Para él, la aritmética es la ciencia del número y la geometría es la ciencia de la magnitud. Estas dos ciencias no están relacionadas en el sentido de que una es subordinada de la otra; él explica esta independencia sobre la base de la independencia del número y la magnitud.

Sin embargo existe una excepción: si una magnitud tiene una medida asignada, puede ser pensada como un número. En este caso, los teoremas de la aritmética pueden ser aplicados a magnitudes.

La utilización de los números entraba en el ámbito de la aritmética y allí se trataba de fenómenos discretos, mientras que la geometría se ocupaba de magnitudes continuas.

### 5. EL INFINITO ARISTOTÉLICO

Para Aristóteles es en el continuo que el infinito aparece por primera vez.

Debido a que Aristóteles deriva los conceptos matemáticos a partir de la experiencia, es que para él no existe una entidad matemática infinita en extensión. A pesar de esto, Aristóteles trata el tema del infinito en diferentes momentos, en “Física” III, capítulo 6, dice:

“Una cantidad es infinita si es de tal manera que siempre podemos tomar una parte afuera de lo que ya ha sido tomado.”

Las magnitudes, por lo tanto, son infinitas en división, o en dirección de decrecer. La característica central del concepto Aristotélico del infinito es que es un proceso, procedimiento de un paso al siguiente: siempre existe un siguiente paso.

Aristóteles finalmente sólo acepta el infinito potencial y halla su aprobación en la posición asumida por Euclides (aprox. 321 a .c.-265 a. c.) quién siempre consideró figuras geométricas limitadas, susceptibles de ser ampliadas.

## 6. EUCLIDES DE ALEJANDRÍA Y LA CONSTITUCIÓN DE LOS “ELEMENTOS”

El libro I empieza con una serie de definiciones (de las cuales sólo examinaremos dos):

- Una línea es longitud sin anchura.
- Una superficie es lo que sólo tiene longitud y anchura.

Euclides presenta a continuación cinco postulados (sólo enunciaremos tres):

- Postúlese el trazar una línea recta desde un punto cualquiera hasta un punto cualquiera.
- El prolongar continuamente una recta finita en línea recta.
- El describir un círculo con cualquier centro y distancia.

Finalmente tenemos una relación de nociones comunes (sólo enunciaremos la quinta):

- El todo es mayor que la parte.

Como podemos observar, las definiciones muestran con claridad la conformidad de Euclides con nociones y formulaciones anteriores. Se mueve dentro del marco general que Aristóteles evoca en su “Metafísica”: “lo completamente indivisible según la cantidad y carente de posición se llama unidad, y cuando es completamente indivisible con posición se llama punto.”

## 7. EL CONCEPTO EUCLIDIANO DE UNIDAD

Tenemos que Euclides sigue la tradición Aristotélica al considerar a la unidad indivisible. No existen las fracciones en “Los Elementos”, él nunca divide la unidad; uno de los rasgos peculiares de las matemática griegas teóricas es que no hay fracciones. En lugar de fracciones, relaciones entre los números fueron usados, y estas relaciones son lo que llamamos razones y proporciones. Utiliza, sin embargo, el término mitad en situaciones como la bisección de la línea, pero nunca lo utiliza de modo sustancial en una proposición.

Al discutir las relaciones entre las partes de un número, Euclides usa la teoría de proporciones. Desde luego que esta teoría puede reemplazarse con la teoría de fracciones. Las fracciones son una parte de la matemática griega, pero no lo fueron de la matemática teórica y filosófica dada por Euclides en los Elementos.

Euclides sigue la tradición que no incluía la unidad entre los números, aún cuando en términos de la sustancia matemática este hecho tiene poca importancia. De hecho se refería a todos los números como unidades y números.

## 8. EUCLIDES Y EL INFINITO

Euclides hace un uso del infinito que coincide con el de Aristóteles. Por ejemplo, en el segundo postulado de “Los Elementos” se afirma explícitamente la posibilidad de prolongar un segmento. Esto indica que en el transcurso del siglo IV a. C., el período de formación de la geometría en el que resaltan o aglutinan sus conceptos

fundamentales, también era discutido el problema del infinito y afrontada la legitimidad de su empleo en el campo matemático, o la famosa prueba de que hay una infinidad de números primos:

“Los números primos son más que cualquier magnitud asignada de números primos.”

La fórmula lingüística que utiliza Euclides dice lo mismo que la utilizada por Aristóteles en la “Física”: siempre toma algo aparte de lo que ya se ha tomado.

Euclides usa el infinito en el mismo sentido en el que Aristóteles habló de él. Es decir existe una infinidad potencial de números primos.

Podemos detectar fácilmente la posición de Euclides al respecto gracias a nuestra propia definición de un conjunto infinito:

“Un conjunto equivalente con un subconjunto propio.”

La noción común: “el todo es mayor que las partes”, se opone a tal definición. Nuestra definición del infinito nos dice que el todo no necesariamente es mayor que una parte. Esta noción común restringe la discusión sólo a colecciones finitas de objetos.

No sólo en los Elementos de Euclides sino en el Método de Arquímedes, donde usa el método de exhaución de Eudoxo, aparece el infinito del único modo previsto por Aristóteles, evitando cualquier consideración que se refiera a una presunta existencia actual del mismo. Sin embargo, el mismo examen de estos métodos permite encontrar una concepción distinta del infinito pues como sabemos, los trabajos de este matemático griego se reconocen como una de las raíces del cálculo moderno pues allí aparece el concepto de límite como salida conceptual que permite encapsular los procesos infinitos para obtener resultados específicos; el método exhaustivo involucra, de manera oculta, un tratamiento infinitesimal, constituyéndose así en el primer paso hacia la adopción del infinito como concepto matemático.

## 9. EL DESPERTAR DE EUROPA

Sabemos que Europa se apropió de las creaciones de los griegos y, sobre ella levantó una civilización grande y de orientación científica, pero ¿cómo ocurrió esto?

Hacia el siglo V d. C. la iglesia católica ya era una organización fuerte; poco a poco fue aumentando el número de seguidores, estableció escuelas por toda Europa y perpetuó e impuso la organización legal y política de Roma.

Se establecieron en toda Europa ciudades y pequeños Estados gobernados por jefes poderosos. Creció el comercio entre ciudades, lo que produjo la riqueza suficiente para apoyar la educación, salvo que esta se dedicó casi por entero a entender la palabra de Dios tal y como lo creían, explicaban y prescribían los padres de la Iglesia. Las obras griegas que sobrevivieron al ímpetu destructor de romanos, cristianos y musulmanes permanecieron prácticamente olvidados.

Una nueva civilización se levantó en Europa. Desgraciadamente fue infecunda en cuanto a la prosecución del saber y de la creación en el terreno matemático, se difundió la enseñanza de principios éticos, se fomentó la arquitectura gótica y se produjeron incomparables pinturas religiosas; pero no se estableció ningún concepto científico, ni técnico, ni matemático; podemos decir que, en ninguna de las civilizaciones que han dado su aportación a la edad moderna estuvo la ciencia matemática reducida a tan bajo nivel.

Al impulso de fuerzas provenientes de fuera de Europa, el mundo cristiano sufrió cambios de grandes proporciones. Los árabes fueron la primera influencia tendiente

a transformar el pensamiento y la vida de la Europa medieval. Los árabes lograron plantar su propia civilización y cultura en el sur de Europa, el norte de África y el Cercano Oriente, asimilaron rápidamente las matemáticas y demás ciencias de los griegos y los hindúes, y construyeron centros culturales en España y el Cercano Oriente. Tradujeron al árabe las obras griegas y añadieron sus propias contribuciones a los trabajos de matemáticas, astronomía, medicina, óptica, meteorología y demás ciencias.

## 10. EL SIGLO XV

Sus principales exponentes se encuentran en el campo de las artes aunque también se produjo una renovación en el de las letras y la ciencia.

## 11. EL SIGLO XVI

El desarrollo del álgebra en Europa, durante el siglo XVI y parte del XVII siguió apoyándose en los significados geométricos. En el libro de Zariski (Zariski, 1926) se muestra como Girolano Cardano(1501-1576) hace uso explícitamente de la idea de completitud para examinar raíces positivas de una ecuación cúbica. Así que nos preguntamos ¿Qué conjunto numérico está usando Cardano, que le permite asegurar la existencia de una raíz?

Con respecto al infinito, uno de los personajes que más influyeron en el pensamiento científico fue Nicolás Copérnico considerado como el fundador de la astronomía moderna. Su obra maestra, “De Revolutionibus Orbium Coelestium”, representó la ruptura con la ideología religiosa medieval, la sustitución de un cosmos cerrado y jerarquizado, con el hombre como centro, por un universo homogéneo e infinito situado alrededor del sol. La Iglesia católica europea inició su intento de acabar con tales herejías. Giordano Bruno (1549-1600), argumentó vehementemente a favor de un universo infinito en “Sobre el universo y los mundos” (1584), “Dios es omnipotente y perfecto y el universo es infinito...”

Fue llevado ante la inquisición y quemado en la hoguera en 1600.

## 12. EL SIGLO XVII

**12.1. El desarrollo del Cálculo.** En el siglo XVII la noción de curva era la del lugar geométrico de un punto móvil. Isaac Newton (1642-1727), en su “Tratado sobre la cuadratura de las curvas” nos dice:

“En este trabajo considero las magnitudes matemáticas constituidas no por partes arbitrariamente pequeñas, sino por el movimiento continuo de puntos, las superficies, por el movimiento de líneas, los sólidos por el movimiento de superficies, los ángulos por la rotación de sus lados, los tiempos, por el flujo continuo, y así en otros casos semejantes.”

Así tenemos que el Cálculo que inventan Newton y Leibniz es sobre curvas y no sobre funciones (la definición de función se dará en el siglo XVIII por otro gran matemático: L. Euler, (1707-1783)), de donde el cálculo tiene un tratamiento esencialmente geométrico al igual que toda la matemática hasta esa época y la pregunta sigue siendo la misma ¿sobre qué sistema numérico?

## 13. EL SIGLO XVIII

En el siglo XVIII aumentó considerablemente el número de aplicaciones del cálculo pero el uso impreciso de las cantidades infinitas e infinitesimales, así como la intuición geométrica, causaban todavía confusión y controversia sobre sus fundamentos.

## 14. SIGLO XIX

A lo largo del siglo XIX, con la geometría de Grassmann, Riemann y Lobachevski, se habían rebasado ampliamente los límites de la experiencia física. Louis Weber decía:

“Las geometrías no euclidianas habían tenido el efecto de despojar a la intuición espacial de ese carácter apodíctico que la volvía absoluta y eternamente necesaria para todos los espíritus, la desaparición de los a priori era no sólo visible en las acreditadas geometrías antes mencionadas”.

Con satisfacción se tomaba nota de que a lo largo del siglo XIX la geometría y el análisis fueron objeto de una paciente elaboración que debía eliminar cada vez más la intuición, y depositar el valor demostrativo de esas ciencias en los elementos del pensamiento puro.

Los números no podían escapar a esta nueva filosofía. R Dedekind (1831-1916) decía que los números son libres creaciones de la mente humana, de este modo, los números irracionales, fueron definidos por él como cortaduras del cuerpo racional. Cuando estamos en relación con una cortadura no producida por algún número racional, nosotros creamos un nuevo número irracional consideramos completamente definido por esta cortadura.

También. G. Cantor (1845-1918), al igual que Dedekind, creaba conceptos, sobre todo, el de los números transfinitos y la definición de continuo.

El paso fundamental que dieron estos matemáticos fue reconocer que se estaban utilizando evidencias geométricas para justificar algunos pasos en el cálculo diferencial.

15. LA DEFINICIÓN AXIOMÁTICA DE  $\mathbb{R}$ 

**15.1. Los axiomas de Hilbert.** En 1899 se publicó el libro “Fundamentos de Geometría” de David Hilbert (1862-1943). En esta obra él asume la tarea de establecer un conjunto completo de axiomas para la presentación de la geometría. Introdujo los elementos geométricos mediante cinco grupos de axiomas: los de incidencia, los de orden, los de congruencia, el de las paralelas, y los axiomas de continuidad.

Los axiomas de continuidad para la recta son dos: el axioma de Arquímedes para segmentos y el axioma de continuidad para la recta (es imposible extender en un conjunto de puntos a la recta con sus relaciones de orden y congruencia de modo que se preserven las relaciones existentes entre los elementos del conjunto original y las propiedades fundamentales del orden y la congruencia de la recta que se siguen de los axiomas I, II, III y VI).

Con las mismas ideas, y por la misma fecha, Hilbert introdujo axiomas para definir al conjunto de los números reales. Presentó cuatro grupos de axiomas: los de composición, los de cálculo, los de orden y los de continuidad. Los tres primeros grupos corresponden a lo que hoy llamamos cuerpo ordenado.

Los axiomas de continuidad son: el axioma de Arquímedes (propiedad Arquimedianidad) y el axioma de completitud (los números forman una colección de objetos que no puede ampliarse sin que deje de cumplirse alguno de los axiomas precedentes).

En la presentación axiomática de Hilbert se define a los números reales como un conjunto de objetos que satisfacen ciertas propiedades. Él necesita explicitar la arquimedianidad pues no se desprende de los otros grupos de axiomas. Mientras que en las construcciones hechas por Dedekind y Cantor, es una propiedad que satisfacen los números reales.

## 16. EL SIGLO XX: TEORÍA DE CONJUNTOS Y LA HIPÓTESIS DEL CONTINUO

En esta etapa, el estudio matemático del infinito se mueve hacia la teoría de conjuntos, inventada por Cantor, cuyo trabajo ha sido fundamental para comprender las matemáticas de nuestro tiempo. En el último tercio del siglo XIX Cantor descubrió una estructura del infinito que no había sido prevista hasta entonces y que a él mismo le costaba creer

Años más adelante, en 1904, E. Zermelo (1871-1953) justificó la suposición de Cantor.

Lo que hizo Zermelo para evitar las paradojas fue desarrollar una teoría axiomática: la teoría de conjuntos de Zermelo, donde se establecen principios para formar conjuntos.

## 17. CONCLUSIONES

Hemos hecho un recorrido histórico por más de veinticinco siglos para describir la génesis de dos sistemas axiomáticos que en la actualidad juegan un papel predominante en el desarrollo de las matemáticas. Nos han servido para validar hechos matemáticos y para resolver problemas que se tenían pendientes. Considero que, el tipo de análisis que se realiza es de utilidad en la enseñanza de la matemática ya que la comprensión de un fenómeno no puede ser completa sin una vuelta a sus orígenes, a las ideas iniciales. El estudiante descubrirá los porqués históricos y lógicos que dan origen a los conceptos, favoreciendo en el estudiante la adquisición de una mayor objetividad.

## REFERENCIAS

- [1] Apostol, T., *Calculus*, Ed. Reverté, 1995 1966.
- [2] Artigue, M., *Epistémologie et didactique*, Recherches en Didactique des Mathématiques, 10 (2-3), 241-286. .
- [3] Artigue, M., *The role of epistemology in the analysis of teaching-learning relationships in mathematics education.*, Plenary Lecture, CMESG, Proceedings, 7-21, 1995.
- [4] Aristóteles, *Física*, Ed. Universidad Nacional Autónoma de México, 2001.
- [5] Aristóteles, *Metafísica*, Ed. Porrúa, Colección Sepan Cuántos, 1969.
- [6] Cantor, G. *Ueber die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen.* Mathematische Annalen V, 123-132. Traducido al francés en Acta Mathematica 2, 336-348, 1871.
- [7] Cauchy, A-L. *Curso de Análisis*, Colección MATHEMA, Facultad de Ciencias de la UNAM, México, 1994.
- [8] Courant, R.- John, F. *Introducción al Cálculo y al Análisis Matemático*, Ed. Limusa- Wiley, México, 1971.
- [9] Euclides, *Los Elementos*, Ed. Gredos, S.A., 1991.

- [10] Hilbert D. *Foundations of Geometry*, Second Edition, translated from the Tenth Edition, revised and enlarged by Dr. Paul Bernays. The Open Court Publishing Company, La Salle, Illinois, 1971.
- [11] Ibarra, M., Linares R., *El infinito*, Textos Científicos, BUAP, 2009,17-26.
- [12] Kline, M. El pensamiento matemático de la Antigüedad a nuestros días, Addison-Wesley, 1998.
- [13] Spivak, M. *Calculus*, Ed. Reverté, España, 1977.
- [14] Van Roostelar, B. *Bolzano's Theory of Real Numbers*, Archive for History of Exact Sciences, 2, 168-180, 1962.

Facultad de Cs. Físico Matemáticas, BUAP.

Rio Verde y San Claudio, Col. San Manuel,

Puebla, Pue. C.P. 72570.

`linares@fcm.buap.mx`.

# Sesión de Lógica Matemática

Quinta Gran Semana Nacional de la Matemática



# RESEÑA BREVE SOBRE LÓGICAS MULTIVALUDAS

IVÁN CORTÉS CORTÉS  
JOSÉ ARRAZOLA RAMÍREZ  
FCC - BUAP  
FCFM - BUAP

RESUMEN. En el presente trabajo presentamos a nivel divulgatorio algunos de los sistemas lógicos multivaluados que fueron desarrollados a principios del siglo XX, como el de Kleene, Bochvar y Lukasiewicz.

## 1. INTRODUCCIÓN

Algunos sistemas lógicos trivaluados que se presentaron a principios del siglo XX fueron realizados por

- (1) Kleene
- (2) Bochvar
- (3) Lukasiewicz

Sus principios se encuentran en los trabajos de Boole, Pierce y Vasiliev; la era moderna comienza con Post y Lukasiewicz, quienes dieron la primera descripción sistemática de los sistemas lógicos multivaluados.

**Jan Lukasiewicz.** Fue un matemático de origen polaco que centró sus intereses en la lógica matemática, además de que fue el primero que propuso una lógica no clásica, dichos intereses fueron de carácter filosófico.

**Emil Post.** Fue un matemático de origen polaco-estadounidense trabajó de igual forma en lógica matemática.

## 2. LÓGICAS DE LUKASIEWICZ

La propuesta del sistema multivaluado de Lukasiewicz se basa en rechazar el siguiente principio:

2.1. PRINCIPIO (Del tercero excluido). Sea  $P$  una proposición, o bien  $P$  es verdadera, o bien su negación  $\neg P$  lo es.

Además agrega un tercer valor de verdad, el valor de lo “*posible*”.

2.1. **Matrices trivaluadas de Lukasiewicz.** El sistema original trivaluado está basado en dos conectivos  $\rightarrow$  y  $\neg$ . Los cuales son destinados a generalizar los conectivos, implicación y negación de la lógica clásica.

$\rightarrow$	0	$\frac{1}{2}$	1	$\neg$
0	1	1	1	1
$\frac{1}{2}$	$\frac{1}{2}$	1	1	$\frac{1}{2}$
1	0	$\frac{1}{2}$	1	0

donde 0 determina lo falso, 1, lo cierto y  $\frac{1}{2}$  lo “*posible*”.

*¿Cómo determinamos una tautología?*

2.2. DEFINICIÓN. Se dice que una fórmula es una tautología trivaluada si siempre toma el valor de verdad 1, sin importar que valores de verdad tengan asignadas sus variables.

**1 es el valor designado porque juega un papel importante en la definición de tautología** Pensemos en los valores de verdad de Lukasiewicz como conjuntos de valores de verdad clásicos, i.e.

$$1 = \{V\} \quad 0 = \{F\} \quad \frac{1}{2} = \{V, F\}$$

2.3. EJEMPLO.

- Felipe Calderón fue electo presidente de México, tiene el valor de verdad  $1 = \{V\}$ , mientras que la proposición,
- “AMLO ganará las elecciones en 2012”, tiene el valor de verdad  $\frac{1}{2} = \{V, F\}$

2.2. **Otras lógicas de Lukasiewicz.** En 1922, Lukasiewicz generalizó su lógica trivaluada por su utilidad a las lógicas  $n$ -valuadas y también a un sistema infinitamente valuado.

- La matriz para los sistemas infinitamente valuados queda definida sobre los números racionales en  $[0, 1]$

$$T_n = \left\{ 0 = \frac{0}{n-1}, \frac{1}{n-1}, \frac{2}{n-1}, \dots, \frac{n-2}{n-1}, \frac{n-1}{n-1} = 1 \right\}$$

donde  $T_n$  es el conjunto de valores de verdad (Grados de verdad).

2.3. **Lógicas multivaluadas.** Lukasiewicz utilizó  $T_n$  como conjunto de valores de verdad y definió las primitivas de su lógica  $n$ -valuada así:

2.4. DEFINICIÓN.

$$\begin{aligned} \neg a &= 1 - a \\ a \wedge b &= \min(a, b) \\ a \vee b &= \max(a, b) \\ a \rightarrow b &= \min(1, 1 + b - a) \\ a \leftrightarrow b &= 1 - |a - b| \end{aligned}$$

### 3. LOS SISTEMAS MULTIVALUADOS DE POST

Emil Post, en 1921 da de manera independiente un desarrollo formal de las lógicas multivaluadas. Sus sistemas  $m$ -valuados definidos sobre el conjunto  $\{0, \dots, m-1\}$  tienen como operadores primitivos una disyunción generalizada y una negación generalizada.

$$\begin{aligned} x \vee y &= \min(x, y) \\ \neg x &= x + 1 \pmod{m} \end{aligned}$$

Lo notable en el artículo de Post:

- (1) Una prueba de completitud funcional para su sistema de conectivos
- (2) Un método general de contrucción para una axiomatización completa del sistema  $T_m^n$ , donde los valores  $0, \dots, n$  son designados para  $0 \leq n < m - 1$ .

### 4. LÓGICA DE BOCHVAR

Bochvar en 1939, propone un sistema lógico con objeto de evitar las paradojas lógicas

4.1. EJEMPLO (Paradoja de Russell). Sea  $M$  un conjunto,  $M$  es el conjunto de todos los conjuntos que no pertenecen a si mismos. i.e.

$$M = \{x : x \notin x\}$$

- Introduce un valor indeterminado,  $I$  junto con los valores clásico  $V, F$ .
- Declara como “sin sentido” a las proposiciones que involucran paradojas.

4.1. **Valores de verdad de Bochvar.** Las tablas de verdad de Bochvar para sus conectivas, son las siguientes:

$\wedge$	$V$	$I$	$F$	$\neg$
$V$	$V$	$I$	$F$	$F$
$I$	$I$	$I$	$I$	$I$
$F$	$F$	$I$	$F$	$V$

### 5. EL SISTEMA DE KLEENE

Kleene en 1938, introduce otro sistema lógico trivaluado. Sus conectivos se definen como sigue:

	$q$	$p \wedge q$			$p \vee q$			$p \rightarrow q$			$p \equiv q$			$\neg q$
	$V$	$V$	$I$	$F$	$V$	$I$	$F$	$V$	$I$	$F$	$V$	$I$	$F$	$F$
$p$	$I$	$I$	$I$	$F$	$V$	$I$	$I$	$V$	$I$	$I$	$I$	$I$	$I$	$I$
	$F$	$F$	$F$	$F$	$V$	$I$	$F$	$V$	$V$	$V$	$F$	$I$	$V$	$V$

La motivación de Kleene surge de la teoría de las funciones recursivas.

#### ¿Cómo podemos darnos una idea?

- Imaginemos una computadora que respondiera  $V$  o  $F$  a ciertas preguntas.
- Para algunas no tendrá respuesta
- En ese caso se puede pensar que la respuesta es indefinida

5.1. **Lógica difusa.** Una de las razones para el reciente renacimiento del interés en la lógica multivaluada es el crecimiento de la investigación en el área de los conjuntos difusos y la lógica difusa. Desde su inicio a mediados de los años 60 del pasado siglo, esta rama ha visto un explosivo crecimiento y ahora existen cientos de artículos al respecto, numerosos volúmenes de conferencias, y revistas completamente dedicadas al tema. En este artículo sólo tocaremos lo que respecta a preguntas acerca de lógica.

El crecimiento de este campo se le debe en su mayor parte al entusiasmo y vocación de L.A. Zadeh [1965], quien introdujo el concepto de “conjunto difuso”. Dada una colección de  $X$  elementos, un conjunto difuso sobre  $X$  es caracterizado por una función relación  $f(x)$ , la cual asocia a cada punto en  $X$  con un número real en el intervalo  $[0, 1]$ . El valor más cercano de  $f(x)$  a la unidad, es el mayor grado de relación en el conjunto. Cuando el rango de  $f$  está restringido al conjunto  $\{0, 1\}$ , tenemos la función característica de un conjunto ordinario clásico.

La intención detrás de la definición de conjunto difuso, es representar predicados en un lenguaje ordinario los cuales son vagos o carentes de un criterio definido para la relación. Por ejemplo, la clase de números mayores que 1, es sin duda, bastante

vaga. Zadeh sugiere que ésta pueda ser modelada por un conjunto difuso definido sobre toda la recta real, con una función característica que satisfaga  $f(0) = 0$ ,  $f(1) = 0$ ,  $f(5) = 0.01$   $f(100) = 0.95$   $f(500) = 1$  y así sucesivamente.

Al complemento, a la contención, la unión y la intersección, Zadeh las define como sigue. Si  $f$  es la función asociada con un conjunto difuso sobre  $X$ , entonces el complemento del conjunto tiene la función característica  $1 - f(x)$ . La contención está definida por  $f \leq g$ . i.e.,  $f(x) \leq g(x)$  para todo  $x \in X$ . La unión y la intersección corresponden al  $\max(f, g)$  y  $\min(f, g)$  respectivamente.

Zadeh propone usar estas ideas a la modelación de predicados difusos tal como 'hermoso', 'alto', 'largo', y así sucesivamente. Una aplicación filosófica es sugerida por Goguen [1969] quien intenta una solución de la paradoja clásica del hombre malo (o paradoja del montón). La paradoja se ejecuta como sigue. Estamos inclinados a admitir la verdad de los siguientes dos enunciados:

- (1) Un hombre con 20,000 cabellos sobre su cabeza no es calvo,
- (2) Si removemos un cabello de un hombre que no es calvo, entonces él sigue siendo no calvo.

Sin embargo aplicando (2), 20,000 veces *modus ponens* derivamos la absurda conclusión que un hombre con cero cabellos sobre la cabeza no es calvo. Goguen sugiere que pensemos de "calvo" como un predicado difuso. Entonces si atamos a la implicación (2) un valor de verdad ligeramente menos que 1 y adoptamos la implicación de Lukasiewicz, encontramos que aunque podemos iniciar asignando el valor de verdad 1 a (1) cada aplicación sucesiva del *modus ponens* disminuye el valor de verdad de la proposición "un hombre con 20,000- $x$  cabellos es calvo". De esta manera Goguen argumenta, la lógica difusa evita la paradoja del monte.

Una objeción inmediata que se presenta a esta línea de investigación es su naturaleza extremadamente artificial de ligar a valores numéricos precisos, proposiciones como "73 es el número más grande" o "Guernica de Picasso es hermosa". De hecho parece plausible decir que la naturaleza de los predicados difusos se opone a ligar valores numéricos precisos tanto como se opone a ligar valores de verdad clásicos precisos. La investigación de Zadeh y Goguen, emerge fácilmente de una pequeña reflexión sobre los ejemplos numéricos dados anteriormente.

Zadeh es por su puesto consciente de la naturaleza artificial de su procedimiento, y en publicaciones posteriores el introduce la idea de "valores de verdad difusos". Un valor de verdad difuso (tal como "verdadero", "muy verdadero", "no tan verdadero"), es un conjunto difuso de la recta real. Las asignaciones de valores de verdad a proposiciones ahora toman una forma tal como, "la compatibilidad de los valores numéricos 0.8 con el valor de verdad lingüístico 'verdadero' es 0.7." Sin embargo parece que la "difusión" de los valores de verdad, ha sólo empujado el problema original de regreso al principio, pues aún nos queda precisar numéricamente valores para compatibilidad con los valores de verdad difusos.

Otro problema que surge con la lógica difusa es una dificultad muy similar a las dificultades encontradas al interpretar las lógicas de Lukasiewicz. ¿Cómo vamos a interpretar las operaciones en los conjuntos difusos?. Si interpretamos la intersección y la unión como el equivalente algebraico de la conjunción y la disyunción, y el complemento como negación, entonces las cosas no parecen funcionar bien. Supongamos que nos encontramos ante un problema en reconocimiento de patrones (un caso generalmente discutido en la "literatura difusa"). Entonces un objeto  $x$ , puede ser un triángulo con el valor 0.9;  $f_{\Delta}(x) = 0.9$ . Si el complemento de  $f_{\Delta}$

“no es un triángulo”, y a la unión disyunción, entonces el  $\max(f_{\Delta}, 1 - f_{\Delta})$  podría representar, “es un triángulo o no es triángulo”, y debería ser la función constante 1; pero no es así; pues las operaciones sobre conjuntos difusos no corresponden a conectivas; pero es difícil hacer lo que se supone que representan. Puede verse que la raíz de la dificultad que aquí se nos presenta es idéntica con las dificultades encontradas en los sistemas de Lukasiewicz.

Podemos encontrar una gran variedad de artículos donde se presentan diversas aplicaciones de los conjuntos difusos en diferentes áreas como, psicología, programación de computadoras e inteligencia artificial. Tales aplicaciones fueron expuestas por Zadeh, Fu, Tanaka, y Shimira [1975].

**5.2. La lógica del significado.** Hemos visto como Bochvar extendió los valores de verdad clásicos agregando un tercer valor de verdad, el que es leído como “sin sentido”.

Sus ideas fueron extendidas por Halldén [1949] quien formalizó una versión de la lógica proposicional trivaluada la que contiene un operador de aridad 1  $S\varphi$  en suma a los conectivos proposicionales usuales.  $S\varphi$  se lee como “ $\varphi$  es una proposición significante” y toma el valor  $V$  si  $\varphi$  toma el valor  $V$  o  $F$ , en otro caso el valor  $F$ . Tomando las tablas de Bochvar para los conectivos básicos, y  $V$  como el unico valor designado, ahora tenemos fórmulas válidas como:

$$S(\varphi \wedge \varphi) \leftrightarrow S\varphi \wedge S\psi$$

Las ideas de Handell han sido muy extendidas en el trabajo de Goddard y Rouley [1973]. Estas ideas no sólo incluyen análisis sintáctico y semántico, sino además formalizaciones de cuantificado y lógicas de significado de alto orden. Su libro es la más completa y rica discusión en la literatura de la lógica de significado, y su rol en la historia de la lógica y filosofía.

## 6. RETROSPECTIVA

Hasta ahora hemos evitado la pregunta planteada con anterioridad tomando más o menos un enfoque histórico. La lógica multivaluada desde este punto de vista consiste simplemente en el sistema desarrollado por Lukasiewicz, Post, Bochvar y Kleene o sistemas íntimamente relacionados a estos.

No obstante, esta caracterización deja algo que desear. Sería mejor si diéramos una idea más analítica de lo que es la lógica multivaluada, y cual es la diferencia entra ésta y la clásica. Una definición la cual es explícitamente o implícitamente adoptada por algunos autores es que una lógica multivaluada es involucrada en cualquier momento que asignemos a las fórmulas valores de un sistema lógico valores de un álgebra la cual no son los dos elementos del álgebra Booleana (i.e. tablas de verdad clásicas). Por esta definición, los modelos Booleanos-valuados para la teoría de conjuntos cuentan como lógica multivaluada. Pero este amplio uso de la frase “lógica multivaluada”, no tiene mucho que elogiarsele. Así emerge claramente del hecho que por el teorema 2, cualquier relación de consecuencia uniforme puede ser considerada como una lógica multivaluada.

Con el fin de obtener una idea más sensata de lo que es la lógica multivaluada, regresaremos a las ideas de los pioneros para ver si podemos extraer algunas ideas esenciales de sus suposiciones.

De acuerdo a la interpretación  $I$  deseamos defender lo que es la característica de una lógica multivaluada, no es tanto el aparato formal de multiples valores como la

relación del formalismo de múltiples valores de verdad a la interpretación intuitiva. Las ideas claves que los sistemas de Lukasiewicz, Bochvar y Kleene tiene en común son las siguientes:

- (1) A los valores de verdad clásicos se les agrega uno o más valores de verdad con el significado como “posible”, “sin sentido” o “indeterminado”. Estos valores de verdad son usualmente considerados linealmente ordenados.
- (2) Las reglas para asignar valor a fórmulas complejas satisface una regla generalizada de funcionalidad veritativa; los valores asignados a una fórmula compleja es una función de los valores asignados a sus componentes.

Ahora ya no se trata de una interpretación intuitiva del formalismo; deberíamos empezar a considerar las interpretaciones de Lukasiewicz y las ideas relacionadas.

## 7. LA LÓGICA DE LA INCERTIDUMBRE

Lukasiewicz deseaba usar su lógica para describir situaciones que envolvían incertidumbre y el ‘futuro abierto’. Como hemos enfatizado en la discusión de la sección 1, esta idea parece ser definitivamente incorrecta, siempre que las conectivas de la lógica de Lukasiewicz son leídas como corresponden al lenguaje ordinario conectivas de conjunción, disyunción, negación e implicación. La lógica de la incertidumbre es simplemente no funcionalmente veritativa.

De hecho podemos establecer algo más fuerte que la anterior negación débil expuesta. Los argumentos de Ramsey, de Finetti y Savage establecen que la probabilidad subjetiva de los valores deben obedecer las reglas del cálculo de probabilidad. Más precisamente, si asumimos que la probabilidad de los valores representan cocientes, entonces las reglas del cálculo de probabilidad para probabilidad subjetiva emergen automáticamente como condiciones de consistencia para un agente racional. Ahora el cálculo de probabilidad no es funcional veritativo, porque por ejemplo, el valor de probabilidad de una conjunción no es una función de los valores de sus conjuntos, porque los conjuntos pueden o no ser estocásticamente independientes. Esta simple consideración basta para mostrar que las ideas de Lukasiewicz son erradas.

Muchas observaciones similares aplican a las interpretaciones que consideran los múltiples valores como representación de los grados de error, o los grados de precisión o vaguedad. Parecería que donde se está intentando formalizar conceptos de incertidumbre, vaguedad y así sucesivamente, las reglas del cálculo de probabilidad, proveen un modelo mucho más atractivo que el marco de trabajo de ideas provistas por la lógica multivaluada. No deseamos clasificar a la teoría de la probabilidad como lógica multivaluada porque viola el principio básico de funcionalidad-veritativa (2). Es un hecho curioso que a pesar de su polémica contra la lógica clásica, Lukasiewicz llegó a fracasar a pesar de los principios clásicos de funcionalidad-veritativa.

**7.1. “Indefinido” como valor de verdad.** La polémica en la sección anterior en contra de cualquier interpretación de la lógica multivaluada relacionada a la incertidumbre subjetiva deja sanas y salvas a las aplicaciones donde un valor intermedio está a favor de lo “indefinido” o “sin sentido”.

Como hemos notado, la idea de usar una lógica multivaluada como un fundamento para la teoría de conjuntos no parece muy prometedora. Sin embargo sigue

habiendo fuertes tablas de verdad de Kleene y sus aplicaciones en teoría de funciones recursivas y otras áreas donde una fórmula falla al ser asignada a un valor de verdad por razones prácticas o razones teóricas.

Aquí, como en el caso anterior, la pregunta crucial gira alrededor del principio general de funcionalidad veritativa. Es posible ofrecer un argumento plausible que en efecto, la lógica de lo “indeterminado” no es funcional veritativa. Supongase que  $\varphi$  falla al tener valores de verdad; evidentemente  $\neg\varphi$  debe además fallar al tener un valor. Pero  $\varphi \vee \neg\varphi$  debería ser verdadera no indefinida (contrario a las tablas de Kleene). Por otro lado,  $\varphi \vee \psi$  puede sin duda ser indefinida. Describiremos brevemente en la siguiente sección una investigación alternativa a las lógicas multivaluadas las cuales toman las ideas anteriores como básicas.

**7.2. Supervaluaciones.** El método de supervaluación fue introducido por Van Fraassen [1966] como un medio para proporcionar semánticas para lógica libre, esto es, una lógica que contenga términos de no significado. Sea  $L$  un lenguaje para lógica de predicados clásica, que contenga letras predicado, variables individuales, cuantificadores y conectivos clásicos. Para dar una interpretación específica  $I$  para  $L$ , daremos un dominio de discurso no vacío y una extensión en  $D$  para los predicados en  $L$ . Más aún para cada constante individual o nombre  $a$  en  $L$ , asignaremos a cada  $a$  una denotación en  $D$ , o dejaremos la denotación  $a$  indefinida. Para una proposición atómica  $Pa$ , ésta es verdadera bajo la interpretación si  $a$  tiene una denotación  $d(a)$  en  $D$  y  $d(a)$  es la extensión de  $P$ , es falsa si  $d(a)$  está definida y no está en la extensión de  $P$ ; de otra manera  $Pa$  no tiene valores de verdad. Así consideraremos una interpretación en la cual exista un hueco en los valores de verdad. ¿Cómo definiremos la verdad o la falsedad en esta interpretación? Un método puede ser tomar a lo ‘indefinido’ como un tercer valor de verdad, entonces usar la lógica trivaluada de Kleene. Van Fraassen, sin embargo, desea conformarse con los argumentos válidos clásicos. En consecuencia, él considera todas las posibles extensiones de la asignación de valores de verdad a las fórmulas atómicas en  $I$ . Vamos a llamar a cualquier extensión la cual llene los espacios en blanco en  $I$  arbitrariamente una extensión clásica de  $I$ . En cualquier extensión clásica podemos evaluar el valor de verdad de cualquier fórmula  $\varphi$  por la definición clásica, teniendo en cuenta que  $\forall x\psi(x)$  es verdadera en tal extensión si y sólo si  $\psi(d)$  es verdadera para cualquier  $d \in D$ . Entonces una *supervaluación* sobre  $I$  es una función que asigna  $T(F)$  exactamente a estas proposiciones asignadas a  $T(F)$  por todas las extensiones clásicas de  $I$ . Diremos que una fórmula  $\varphi$  es SL-válida si ésta es verdadera en todas las supervaluaciones.

La definición de supervaluación tiene una característica atractiva, que ésta hace SL-válida exactamente a los teoremas de la “lógica libre” clásica. Para más información acerca de la supervaluación podemos hacer referencia a la contribución de Bencivenga. Aquí nuestra preocupación principal con las supervaluaciones es tomarlas como una alternativa a las lógicas multivaluadas.

Una de las características más sorprendentes del acercamiento a la supervaluación es que ésta hace SL-válida a la ley del medio excluido, pero no a la ley de bivalencia. Por ejemplo, si el nombre “Bruno Díaz” falla al denotar una interpretación  $I$ , entonces “Bruno Díaz lee Proust” falla al tener un valor de verdad en la supervaluación sobre  $I$ , al igual que su negación, de modo que ni la proposición ni su negación son verdaderas. Por otro lado la proposición “Bruno Díaz lee Proust o no lo lee” es verdad en la supervaluación sobre  $I$ . Ésta a primera vista parece en

dirección intuitiva opuesta, ya que tenemos una disyunción verdadera con ningún disyuntor verdadero. Nuestra sorpresa es disminuida si reflexionamos sobre qué significa decir que una fórmula que contiene términos de no significado es verdadera bajo una supervaluación. Decir que  $\varphi \vee \neg\varphi$  es verdadera, es decir que sería verdadera si pretendemos que los términos de no significado en esta tengan una referencia. Tal fórmula es verdadera sólo en un camino amable, y no en un camino clásico puro. Es importante notar que si consideramos a lo 'indefinido' como un tercer valor de verdad (un movimiento que Van Fraassen enérgicamente resiste) entonces una supervaluación no es veritativa funcional. Por ejemplo, si "Dick Grayson" otro término de no significado, entonces "Dick Grayson lee Baudelaire" no tiene valor de verdad; pero es claro que si "Bruno Díaz lee Proust o Dick Grayson lee Baudelaire" no tiene valor de verdad, en contraste a cualquier ejemplo de la ley del medio excluido.

Para la mayoría de los propósitos, la teoría de supervaluaciones parece superior a los enfoques más viejos que involucran lógicas multivaluadas. El enfoque supervaluativo tiene una ventaja muy real, que nos permite todo lo clásico (lógica libre), mientras admite la posibilidad de valores de verdad en blanco. Las supervaluaciones logra estos objetivos aparentemente incompatibles, precisamente, por abandonar el principio funcional veritativo, el cual como hemos visto es la principal fuente de dificultad en la interpretación de las lógicas multivaluadas. Y lo hace sin duda, al precio de abandonar tales principios clásicos como la idea de que una disyunción es verdad sólo si uno de sus disyuntores es verdadero. De acuerdo a Quine [1953], tal abandono es un intento en extremo desesperado. Los que como Quine desean atenerse a este principio pase lo que pase, tal vez deberían encontrar una tabla de verdad de Kleene más congenial.

**7.3. Conclusión.** En una encuesta de lógica matemática y positivismo lógico, Jordan Zbigniew da la siguiente revisión notable de lógica multivaluada:

"Sin duda alguna se trata de un descubrimiento de primer orden, que eclipsa a cualquier cosa hecha en el campo de la investigación de la lógica en Polonia. (Jordan [1945], en McCall [1967], pág. 389). Este pasaje es el más llamativo, si se refleja que entre los resultados que dice que eclipsó son los siguientes: La simple teoría de tipos de Chwistek, el trabajo de Kuratowski sobre la herencia proyectiva, el trabajo de Jaskowski sobre deducción natural, los resultados de Lindenbaum y todos los anteriores, el trabajo fundamental de Tarski sobre metodología, definibilidad y la teoría de la verdad. Es muy difícil ponerse de acuerdo, ni siquiera en parte con esa exagerada estimación del trabajo de Lukasiewicz. Mientras que las ideas de Tarski han probado su productividad virtualmente en cada área de la lógica. Los sistemas multivaluados de Lukasiewicz ha dejado 'juguetes' y curiosidades de tipo lógico. Esto no es sorprendente si estamos dispuestos a conceder que parece que hay un error fundamental (funcional veritativo generalizado) en la raíz de los sistemas de Lukasiewicz."

Claro que demasiado trabajo ingenioso y atractivo ha sido hecho sobre las lógicas multivaluadas consideradas como una estructura puramente matemática. En esta luz, la lógica multivaluada es simplemente el estudio de las funciones definibles

sobre un conjunto finito. Obviamente, tal investigación tiene una importancia considerable como matemáticas combinatorias puras, las cuales de ninguna manera dependen de su valor en dudosas motivaciones filosóficas.

## REFERENCIAS

- [1] Ackerman, R., 1967, *Introduction to many valued logics*, Routledge and Kegan Paul, London.
- [2] Bellman, R. E and Zadeh, L.A., 1977, *Local and fuzzy logics*, in Dunn and Epstein [1977].
- [3] Belnap, N. D. Jr., *A useful four valued logic* in Dunn and Epstein [1977].
- [4] Bochvar, D.A., 1939, *Ob odnom trézhnacnom iscislénii i égo priménénii k analiza paradoksov klassicéskogo rassirénnoho funkcional'nogo isciléniiá (On a 3 valued logical and calculus and its applications to the analysis of contradictions)*, *Matématiceskij sbornik* 4 287-308.
- [5] Chang, C.C., 1963, *The axiom of comprehension in infinite valued logics*, *Math. Scand.* 13, 9-30.
- [6] Church A., 1939, Review of Bochvar [1939], *J. Symbolic logic* 4, 98-99.
- [7] Dunn, J. M. and Epstein (eds.): 1977, *Modern uses of many valued logic*, D.Reidel, Dordrecht and Boston.
- [8] Epstein, G., 1960, *The lattice theory of Post algebras*, *Trans. Amer Math. Soc.*, 95, 300-317.
- [9] Fenstad, J.E., 1964, *On a consistency of the axiom of comprehension in the Lukasiewicz infinite valued logic*, *Math. Scand.* 14, 65-74.



# Sesión de Probabilidad y Estadística

Quinta Gran Semana Nacional de la Matemática



# KAPLAN-MEIER Y NELSON-AALEN, DOS MÉTODOS NO PARAMÉTRICOS EMPLEADOS EN SUPERVIVENCIA

BULMARO JUÁREZ HERNÁNDEZ  
OSCAR PALMEROS ROJAS  
FRANCISCO SOLANO TAJONAR SANABRIA  
FCFM - BUAP

RESUMEN. Usualmente en el análisis de supervivencia, si se tiene un conjunto de datos de tiempo de vida y es posible ajustar a éste, mediante algún modelo paramétrico, entonces, el trabajo consiste en realizar la estimación del parámetro o parámetros del modelo empleado. Sin embargo, en muchas situaciones del mundo real, los datos teóricos o experimentales no siempre siguen un modelo paramétrico. Por esta razón, surge la necesidad de emplear métodos no paramétricos.

En este trabajo, se presentan los métodos no paramétricos de Kaplan-Meier y Nelson-Aalen, los cuales proporcionan alternativas eficaces para realizar estimaciones para la función de Supervivencia y de riesgo acumulada, respectivamente.

La importancia de éstos es proporcionar métodos descriptivos que permitan calcular probabilidades aproximadas del tiempo de vida de individuos bajo estudio. Además, pueden emplearse para intentar ajustar a algún modelo paramétrico particular, teniendo en cuenta los resultados básicos que relacionan las diferentes funciones en análisis de supervivencia y conociendo la forma de las diferentes curvas de los modelos paramétricos empleados. Sin embargo, presentan la limitante de *no poder hacer predicciones a largo o corto plazo, acerca del evento de interés*.

## 1. INTRODUCCIÓN

El análisis de supervivencia es el conjunto de técnicas que permiten analizar, estudiar y modelar la variable de interés, que en este caso, representa una variable de tipo evento-tiempo. Para llevar a cabo un análisis de tiempos de vida, es necesario considerar lo siguiente:

1. Tener bien definido el evento de interés.
2. Definir de forma apropiada el origen o inicio del estudio.
3. Definir la escala de tiempo.

Las funciones que caracterizan a la variable  $T$ , las cuales se emplean para realizar el análisis estadístico de datos de tiempos de vida, son las siguientes,

### **Función de densidad de probabilidad (f.d.p.):**

es la probabilidad no condicional de que ocurra el evento en el tiempo  $t$ .

### **Función de Supervivencia:**

es la probabilidad de que un individuo sobreviva al menos hasta el tiempo

$t$ , esta se denota por  $S(t)$ , y está dada por:

$$(1) \quad S(t) = P(T \geq t), \quad t \in \mathbb{R}.$$

**Función de Riesgo:**

es la probabilidad de que un individuo de  $t$  años experimente el evento de falla en el siguiente instante de tiempo.

La función de riesgo para la v.a.  $T$ , se define de la siguiente manera,

$$(2) \quad h(t) = \lim_{\Delta t \rightarrow 0} \frac{P[t \leq T < t + \Delta t | T \geq t]}{\Delta t}, \quad t > 0.$$

**Función de riesgo acumulada:** esta función, está definida en términos de (2) de la siguiente manera,

$$(3) \quad H(x) = \int_0^x h(u) du.$$

En análisis de supervivencia, para cada  $t \geq 0$ , las funciones enunciadas anteriormente, están relacionadas de la siguiente manera,

$$S(t) = 1 - F(t); \quad h(t) = \frac{f(t)}{S(t)},$$

$$h(t) = -\frac{d}{dt} \ln[S(t)], \quad S(t) = \exp[H(t)].$$

$$f(t) = h(t) \exp \left[ - \int_0^t h(r) dr \right].$$

Dentro de este contexto, existen varios modelos paramétricos que se han propuesto para describir datos relacionados con los tiempos de vida. Entonces, el problema básico dentro del análisis de supervivencia es dar y representar las distribuciones de tiempos de vida mediante dichos modelos estadísticos y hacer inferencia sobre los parámetros de estos modelos. Entre los modelos clásicos se encuentran, el Exponencial, el Weibull, el Log-normal, Gama, el Gama Generalizado, el de valores Extremo, etc.

Sin embargo, existen situaciones en donde los datos de tiempos de vida, no pueden ser ajustados por los modelos paramétricos mencionados anteriormente. Por esta razón, surge la necesidad de emplear métodos no paramétricos.

## 2. MÉTODOS NO PARAMÉTRICOS

En la mayoría de los problemas de inferencia estadística, se asume que la distribución de las variables aleatorias es conocida, excepto, por supuesto, por los parámetros de ésta. En la práctica sin embargo, la forma funcional de la distribución raramente es conocida. Entonces, es deseable derivar métodos que no involucren la suposición de una forma funcional para la distribución. A estos métodos se les conoce como *métodos no-paramétricos*, para estos, no se hacen suposiciones acerca de alguna forma funcional para la distribución, o no se involucran parámetros en la forma usual. La única condición que se pide a la función de distribución es que pertenezca a una familia de distribuciones absolutamente continua.

Para estos métodos, se pueden tratar casi todos los problemas que se abordan en los métodos paramétricos, por ejemplo, estimación, estadísticos de orden, pruebas de hipótesis, etc. El lector interesado en un análisis más profundo de los métodos no paramétricos en general, puede consultar [6].

En esta sección se presentan dos métodos no paramétricos empleados en el análisis de supervivencia, el de Nelson-Aalen y el de Kaplan-Meier. Realizándose una aplicación de estos, en un problema particular.

## 2.1. Estimador de Kaplan-Meier.

El estimador de Kaplan-Meier (también conocido como el **estimador producto límite**) se emplea para estimar a la función de supervivencia.

- La gráfica del estimador de Kaplan-Meier, aproxima a la función de supervivencia, mediante una serie de pasos horizontales de magnitud decreciente.
- Mientras mayor sea la muestra observada, mejor será la aproximación a la curva de supervivencia verdadera de la población.
- El valor de la función de supervivencia estimada, entre valores sucesivos distintos en la muestra, es constante.
- Una característica que posee este estimador es que puede ser usado en bases de datos que presentan censura, tales como censura tipo I, tipo II, progresiva o censura aleatoria a la derecha, ver [1], [2], [3] y [4].

En estudios médicos, se emplea para medir el porcentaje de pacientes que sobreviven después de ser sometidos a alguna operación de alto riesgo. Un economista, puede estar interesado en medir el tiempo que les toma a las personas emplearse, después de que ha perdido su empleo. Un ingeniero puede medir el tiempo transcurrido hasta que falla alguna componente de una máquina.

Para la construcción del estimador, suponga que  $X$ , es el evento de interés, y suponga que este ocurre en  $N$  momentos distintos, es decir,  $t_1 < t_2 < \dots < t_N$ . Además, se permite que en cada tiempo de ocurrencia haya  $d_i$  fallas.

Sea  $Y_i$  el número de individuos que están en riesgo en  $t_i$ . Obsérvese que  $Y_i$  cuenta el número de individuos que tienen un tiempo de estudio  $t_i$  o más, es decir, el número de individuos que viven más de  $t_i$  o experimentan el evento en  $t_i$ .

El cociente  $d_i/Y_i$ , proporciona un estimador de la probabilidad condicional de que algún individuo que ha sobrevivido hasta antes del tiempo  $t_i$ , experimente el evento en  $t_i$ .

El estimador **Kaplan-Meier**, está definido para todos los valores de  $t$ , en el rango donde se encuentran los datos, está dado por,

$$(4) \quad \hat{S}(t) = \begin{cases} 1, & \text{si, } t < t_i; \\ \prod_{t_i \leq t} \left[ 1 - \frac{d_i}{Y_i} \right], & \text{si, } t_i \leq t. \end{cases}$$

Notemos que los intervalos entre cada tiempo de ocurrencia, no necesariamente son de igual longitud. Por ejemplo, suponga que se tiene una muestra aleatoria pequeña de tamaño 15, en esta muestra, se observa que un individuo muere en el día 5, otro

individuo se pierde o abandona el estudio en el día 11 (es decir, es censurado) y pensemos que uno más muere para el día 15, entonces, para este ejemplo, se tiene que,  $t_1 = 5, t_3 = 15, Y_1 = 10, Y_2 = 8, d_1 = 1, d_2 = 1$ .

La varianza de este estimador se obtiene usando la fórmula de Greenwood, la cual está dada por,

$$\widehat{V} [\widehat{S}(t)] = \widehat{S}(t)^2 \sum_{t_i} \frac{d_i}{Y_i(Y_i - d_i)}.$$

De donde, el error estándar es,

$$\left\{ \widehat{V} [\widehat{S}(t)] \right\}^{\frac{1}{2}}.$$

Una ventaja que posee el estimador, es que proporciona una eficiente herramienta calcular la función acumulativa de riesgo  $H(\cdot)$ , esto se realiza empleando el hecho de que  $H(t) = -\ln[S(t)]$ , entonces,  $\widehat{H}(t) = -\ln[\widehat{S}(t)]$ .

Sin embargo, un estimador alternativo para la función de riesgo acumulada  $\widehat{H}(\cdot)$ , que actúa mejor que el estimador límite producto, cuando el tamaño de la muestra es pequeño, fue sugerido por primera vez por Nelson, en el contexto de rehabilitación. El estimador fue retomado por Aalen, quien dedujo el estimador empleando técnicas modernas de procesos continuos, ver [3].

**2.2. Estimador de Nelson-Aalen.** El estimador de Nelson-Aalen estima a la función acumulativa de riesgo, dicho estimador, está definido sólo hasta el tiempo de observación más grande, es decir,

$$(5) \quad \widehat{H}(t) = \begin{cases} 0, & \text{si, } t \leq t_i; \\ \sum_{t_i \leq t} \frac{d_i}{Y_i}, & \text{si, } t_i \leq t. \end{cases}$$

La varianza de este estimador, la cual se debe a Aalen, está dada por:

$$\widehat{\sigma}_H^2(t) = \sum_{t_i \leq t} \frac{d_i}{Y_i^2}, \quad t \geq 0.$$

Usando el estimador de Nelson-Aalen, podemos dar un estimador para  $S(t)$ , de la siguiente manera,

$$\widehat{S}(t) = \exp \left[ -\widehat{H}(t) \right], \quad \text{para } t > 0.$$

### 3. APLICACIÓN DE LOS ESTIMADORES DE KAPLAN-MAIER Y NELSON-AALEN

En esta sección, se realiza una aplicación de los dos estimadores mencionados anteriormente, el ejemplo se encuentra en [3].

#### PRUEBA HECHA PARA DETERMINAR EL TIEMPO DE RECAÍDA EN PACIENTES ENFERMOS CON LEUCEMIA AGUDA.

En 1963, Freireich reportó los resultados de una prueba clínica para el medicamento 6-mercaptopurine (6-MP) contra un Placebo, en 42 niños enfermos de Leucemia Aguda. La prueba fue realizada en 11 hospitales de Estados Unidos. Los pacientes fueron seleccionados de entre los que presentaban una recaída parcial o total de la

enfermedad, la cual fue inducida por la droga Prednisona (por recaída parcial o total, se entiende que la mayoría o todos los signos de la enfermedad han desaparecido de la médula ósea). La prueba se realizó haciendo pares de pacientes en los diferentes hospitales, teniendo en cuenta el estatus de recaída (completa o parcial), en cada par se escogió de manera aleatoria el tratamiento asignado, 6-MP o Placebo. Se les dio seguimiento a los pacientes hasta que su leucemia regresó (recae) o hasta el final del estudio. Los datos del experimento están dados en la siguiente tabla.

Par	E. Recaída	T.recaída Placebo	T. Recaída 6-MP
1	Parcial	1	10
2	Completa	22	7
3	Completa	3	32 <sup>+</sup>
4	Completa	12	23
5	Completa	8	22
6	Parcial	17	6
7	Completa	2	16
8	Completa	11	34 <sup>+</sup>
9	Completa	8	32 <sup>+</sup>
10	Completa	12	25 <sup>+</sup>
11	Completa	2	11 <sup>+</sup>
12	Parcial	5	20 <sup>+</sup>
13	Completa	4	19 <sup>+</sup>
14	Completa	15	6
15	Completa	8	17 <sup>+</sup>
16	Parcial	23	35 <sup>+</sup>
17	Parcial	5	6
18	Completa	11	13
19	Completa	4	9 <sup>+</sup>
20	Completa	1	6 <sup>+</sup>
21	Completa	8	10 <sup>+</sup>

**Tabla 1.** Tiempo de recaída de 6-MP, contra el Placebo, en niños enfermos de Leucemia Aguda.

Empleando los datos anteriores, se calcula el estimador Límite Producto. Los resultados se muestran en la siguiente tabla.

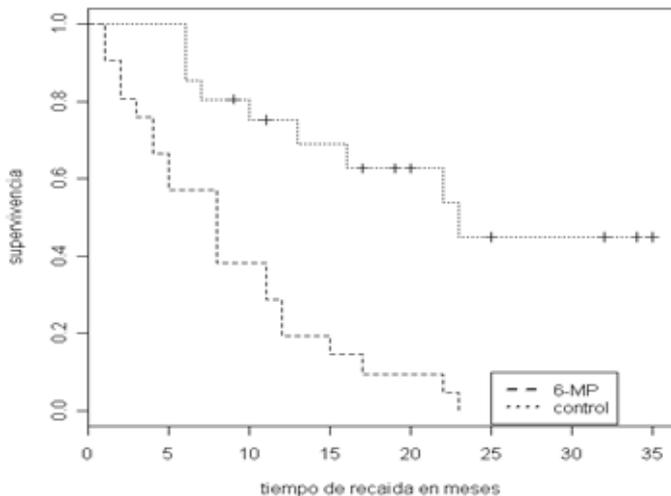
T. Estudio	$\widehat{S}(t)$	E. estándar
$0 \leq t < 6$	1.000	0.000
$6 \leq t < 7$	0.857	0.076
$7 \leq t < 10$	0.807	0.087
$10 \leq t < 13$	0.753	0.096
$13 \leq t < 16$	0.690	0.107
$16 \leq t < 22$	0.628	0.114
$22 \leq t < 23$	0.538	0.128
$23 \leq t < 35$	0.448	0.135

**Tabla 2.** Estimador límite producto y su varianza aproximada para 6-MP.

T. Estudio	$\widehat{S}(t)$	E. estándar
$0 \leq t < 1$	1.000	0.000
$1 \leq t < 2$	0.809	0.085
$2 \leq t < 3$	0.762	0.093
$3 \leq t < 4$	0.667	0.103
$4 \leq t < 5$	0.572	0.106
$5 \leq t < 8$	0.382	0.108
$8 \leq t < 11$	0.286	0.101
$11 \leq t < 12$	0.191	0.085
$12 \leq t < 15$	0.143	0.076
$15 \leq t < 17$	0.095	0.064
$17 \leq t < 22$	0.047	0.046
$22 \leq t < 23$	0.000	0.000

**Tabla 3.** Estimador límite producto y su varianza aproximada para el Placebo.

Las curvas de supervivencia para ambos tratamientos, se muestran en la siguiente figura.



**Figura 1.** Curvas de supervivencia para los dos grupos, las cruces en la curva del tratamiento 6-MP, representan el tiempo de censura.

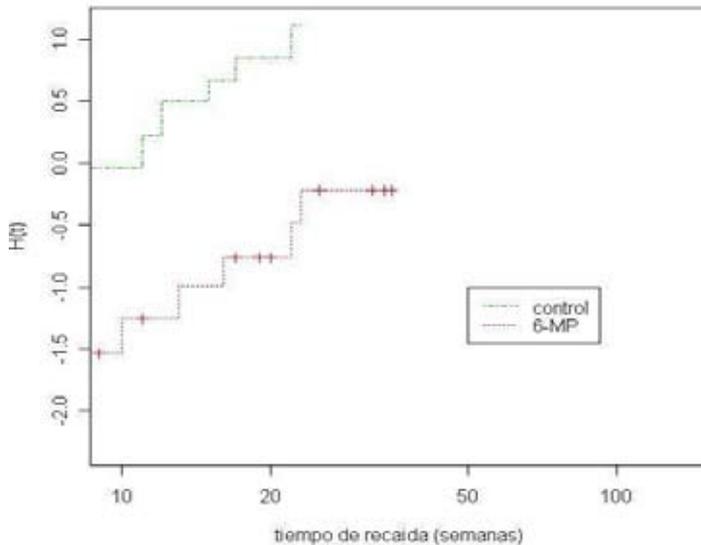
Ahora, empleando el mismo ejemplo, se calcula el estimador de Nelson-Aalen, con su respectiva gráfica, para los dos tratamientos.

T. Estudio	$\hat{H}(t)$	E. estándar
$0 \leq t < 6$	0	0
$6 \leq t < 7$	0.1428	0.0825
$7 \leq t < 10$	0.2017	0.1015
$10 \leq t < 13$	0.2683	0.1212
$13 \leq t < 16$	0.3517	0.1473
$16 \leq t < 22$	0.4426	0.1729
$22 \leq t < 23$	0.5854	0.2243
$23 \leq t < 35$	0.7521	0.2795

**Tabla 4.** Estimador de Nelson-Aalen para la función de riesgo acumulada y su varianza aproximada, para 6-MP.

T. Estudio	$\hat{H}(t)$	E. estándar
$0 \leq t < 1$	0	0
$1 \leq t < 2$	0.095	0.067
$2 \leq t < 3$	0.020	0.100
$3 \leq t < 4$	0.258	0.116
$4 \leq t < 5$	0.270	0.130
$5 \leq t < 8$	0.383	0.158
$8 \leq t < 11$	0.425	0.171
$11 \leq t < 12$	0.501	0.185
$12 \leq t < 15$	0.543	0.192
$15 \leq t < 17$	0.695	0.204
$17 \leq t < 22$	0.857	0.216
$22 \leq t < 23$	1.000	0.230

**Tabla 5.** Estimador de Nelson-Aalen para la función de riesgo acumulada y su varianza aproximada, para el Placebo.



**Figura 1.** Curvas de la función de riesgo acumulada para los dos grupos, las cruces en la curva del tratamiento 6-MP, representan el tiempo de censura.

#### 4. CONCLUSIONES

Los métodos presentados en las secciones anteriores, son los más conocidos y los más empleados dentro del análisis de supervivencia. Cada uno tiene características particulares que los pone en ventaja y desventaja frente al otro método. Sin embargo, vale la pena señalar que como son métodos descriptivos, *no se puede hacer predicciones a largo o corto plazo, acerca del evento de interés.*

No existe regla general para inclinarse por alguno, hay que tener en cuenta el experimento de interés, el tamaño de la base de datos, además del tipo de censura presente.

Entre las principales características, están,

- Proporcionar métodos descriptivos.
- Permiten calcular probabilidades aproximadas del tiempo de vida de individuos bajo estudio.
- Pueden emplearse para intentar usar algún modelo paramétrico particular.

#### REFERENCIAS

- [1] Cox D. R. and Oakes, D. Analysis of Survival Data, Chapman and Hall.
- [2] Kalbfleisch J.D. and Prentice R.L. The Statistical Analysis of Failure Time Data, New York John Wiley. 1980.
- [3] Klein P. John and Moeschberger L. Melvin, Survival Analysis: Techniques for Censored and Truncated Data, New York: 1997.
- [4] Lawless J. F. Statistical Models and Methods for Lifetime Data. Second Edition Wiley Series. 2003.
- [5] W.N. Venables, B.D. Ripley. Modern Applied Statistics with S. 2002.
- [6] Vijay K. Rohatgi, A. K. MD. Ehsanes Saleh, An introduction to Probability and Statistics. New York John Wiley.

Facultad de Ciencias Físico Matemáticas, BUAP.  
Avenida San Claudio y 18 Sur, Col. San Manuel  
Puebla, Pue. C.P. 72570. Mexico.  
bjuarez@fcfm.buap.mx  
opalmeros\_rojas@hotmail.com  
ftajonar@fcfm.buap.mx

# VALUACIÓN DE UNA OPCIÓN CALL BARRERA DOBLE TIPO A

CARLOS PALOMINO JIMÉNEZ  
FRANCISCO SOLANO TAJONAR SANABRIA  
FCFM - BUAP

RESUMEN. En la teoría de opciones se entiende que el valor de un call europeo con una barrera de tipo down-and-out depende de la barrera. En este trabajo se presenta el estudio de opciones con barreras dobles y se realiza el análisis de un call europeo con barreras de tipo A.

## 1. INTRODUCCIÓN

Es sabido que los procesos que suceden dentro de las bolsas de valores modernas son increíblemente diversos y complejos. La teoría moderna de las matemáticas financieras [[1], [4], [5], [7], [14], [13]] que pretende describir estos procesos y predecir su dirección, se finca en los logros y la experiencia de más de 300 años de la Teoría de la Probabilidad y otras ramas de las matemáticas, con ayuda de las computadoras electrónicas dotadas con un software sofisticado.

La Matemática Financiera es uno de los campos de más rápido desarrollo en el mundo moderno bancario y corporativo. La creciente sofisticación de los productos financieros ha propiciado un fuerte ímpetu para este desarrollo; ya nadie que trabaje en los mercados de valores puede sobrevivir sin el apoyo de los métodos de las matemáticas financieras.

La actividad principal de las matemáticas financieras es la construcción e investigación de modelos matemáticos de los procesos que se presentan en los mercados financieros. Su importancia yace en que estos modelos se emplean en el diseño de muchos tipos de estrategias de inversión. Para que un modelo matemático sea funcional, debe de tomar en cuenta con suficiente precisión las características más relevantes de los procesos financieros involucrados, pero a la vez debe ser suficientemente sencillo para admitir una investigación matemática a fondo y una implementación computacional efectiva. Como consecuencia ningún modelo por sí solo podrá ser adecuado para todas las posibles aplicaciones a los mercados financieros.

## 2. TIPOS DE OPCIONES

2.1. DEFINICIÓN. Las opciones pueden dividirse en dos clases amplias: opciones call y opciones put. Un call es un valor que le da a su dueño el derecho de comprar un número fijo de acciones de una cierta empresa en un precio establecido en un momento permitido hasta una fecha especificada. Un put es un contrato que le da a su tenedor el derecho de vender un número fijo de acciones de una empresa especificada en un precio establecido en un momento permitido hasta una fecha predeterminada. Puesto que el valor que puedan tener las opciones depende de otro valor, que es la acción subyacente, se le conocen como derivados financieros.

La opción básica más conocida y la que es más sencilla de analizar y de valorar, es la opción Europea. Para una opción europea, la fecha de vencimiento  $T$  y el precio de ejercicio  $K$  están dados de antemano. En particular, el tenedor no puede ejercer la opción con anterioridad a la fecha de vencimiento. Igual que para la mayoría de las opciones, tiene sus versiones put y call. Un put europeo normalmente se compra cuando el tenedor cree que el precio de la acción bajará.

Las opciones han existido en una forma u otra durante cientos de años, por ejemplo para proveer protección en negocios de la producción agrícola. Sin embargo, hacía varias décadas, eran instrumentos financieros poco conocidos, puesto que los contratos de opciones generalmente eran concertados en circunstancias improvisadas (ad hoc), haciéndolos bastante costosos.

La opción europea como ya se dijo es la opción más popular, pero existen otras que también son utilizadas muy frecuentemente. Una de estas es la llamada *opción americana* que difiere de la europea en que puede ejercerse en cualquier momento antes de la fecha de vencimiento  $T$ ; el precio de ejercicio  $K$  nuevamente tiene un valor preespecificado. Las opciones americanas están entre las opciones más populares en los mercados financieros porque el tenedor tiene la oportunidad de ejercer su derecho cuando lo desee (hasta la fecha  $T$ ) y esto puede reflejarse en el valor de la opción.

La *opción asiática* le da al tenedor el derecho de comprar (para un call) o de vender (para un put) las acciones subyacentes al precio que es el promedio del precio de la acción hasta la fecha de vencimiento especificada  $T$ . Las opciones asiáticas comparten con las europeas la característica de que no pueden ser ejercidas antes de la fecha de vencimiento.

Las opciones asiáticas son populares en algunos mercados de commodities. Esto es porque si una firma piensa comprometerse a comprar o vender una cantidad grande de mercancías en el futuro, entonces es atractivo basar la operación en algún tipo de precio promedio.

Las *opciones lookback* le dan al tenedor el derecho de vender o de comprar (según sea put o call) en un precio igual al máximo o mínimo del precio de la acción hasta la fecha pre-especificada  $T$ .

En los mercados financieros actualmente también trabajan con opciones barreras [[6],[8],[15]], las cuales se pueden ejercer si durante la vida de la opción el precio de la acción subyacente es siempre mayor (o siempre menor) que cierto valor  $X_0$  (la barrera) o alternativamente, si esta barrera se alcanza durante la vida de la opción. A diferencia de las opciones de tipo europeo, el valor del derecho contingente de una opción barrera no está dado por ninguna función que dependa del precio final de la acción solamente, pues depende de toda la historia bursátil.

Además de lo anterior, las opciones barreras son opciones que dependen de las trayectorias, es decir, su pago no sólo depende del precio del asset subyacente al expirar, sino que también depende de la historia pasada.

Las opciones barreras difieren de las opciones vanilla ya que parte del contrato de la opción es ocasionado si el precio del bien subyacente,  $S$ , alcanza una cierta barrera,  $B$ , en un tiempo anterior a expirar. El derecho a ejercer la opción puede ser abandonado en esta barrera, una barrera out, o puede darse la situación que la barrera sólo existe si el precio del asset cruza un cierto valor, una barrera in. Las Opciones barrera pueden ser un put o un call, y se clasifican en:

*Sube y sale (up-and-out)*: La opción caduca sin valor, si el precio de la acción sube hasta la barrera  $S = X_0$  (i.e la barrera es alcanzada desde abajo) antes del día del vencimiento.

*Baja y entra (down-and-in)*: La opción caduca sin valor a menos que la barrera  $S = X_0$  sea alcanzada desde arriba antes del vencimiento.

*Baja y sale (down-and-out)*: la opción caduca sin valor, si el precio de la acción cae hasta la barrera  $S = X_0$  (i.e la barrera es alcanzada desde arriba) antes del vencimiento.

*Sube y entra (up-and-in)*: la opción caduca sin valor a menos que la barrera  $S = X_0$  sea alcanzada desde abajo antes del vencimiento.

Frecuentemente se permite una rebaja, por lo que el tenedor de la opción recibe un monto especificado si la barrera es alcanzada, en el caso de barreras out, o no cruzada, en el caso de barreras in.

Una de las características por las cuales las opciones barreras son populares es por que tienen menor costo que las opciones vanilla; esto es porque las opciones vanilla suponen una vista más precisa de la dirección del mercado. Es por esto que en este trabajo nos ocuparemos de estas opciones con una barrera.

Existen mucho tipos diferentes de opciones de barreras (barrera europea, barrera americana, barrera asiática, etc.), según el tipo de pago.

En la literatura clásica de matemáticas financieras [[1], [4], [5], [7], [11], [12], [13]] se trata sobre la valuación de las opciones más comunes, las cuales son las Europeas y las Americanas. Actualmente al menos en el marco teórico, surge el estudio de la valuación de opciones barreras [15], las cuales sólo se han modelado con un activo. Es importante hacer notar que en la mayoría de los textos elementales no se toca a profundidad el tema de las opciones barreras. En [3] se deriva el precio de una opción con dos barreras de tipo knock-and-in. Por otro lado, en [14] se presenta un enfoque más unificado que permite modelar derivados financieros haciendo uso de ecuaciones diferenciales parciales, así como aquí, también se expone por qué la teoría de las ecuaciones diferenciales parciales es de las mejores aproximaciones para modelar problemas relacionados con temas financieros. Pelsser [9] obtiene una fórmula para valorar una opción con dos barreras utilizando la transformada de Laplace. En [2] se propone una solución analítica para el precio de una opción barrera discreta, se reduce el problema de valuación a una ecuación de Wiener-Hopf que puede ser resuelta analíticamente, todo lo anterior se presenta dentro del marco de Black-Scholes. En [10], se utiliza un resultado de reflexión para presentar demostraciones simples de técnicas de valuación de opciones con una barrera con cero rebaja en el modelo de Black-Scholes.

Para valorar y proteger derivados de seguridad es necesario e importante tener un buen modelo de distribución de probabilidad del producto subyacente. El modelo más famoso a tiempo continuo es el de Black-Scholes, el cual utiliza la distribución normal para ajustar los precios de las acciones subyacentes.

### 3. OPCIONES BARRERAS DOBLES

Las opciones barreras se dividen en tres grandes clases:

**Opción tipo A:** La opción se lleva a cabo sii durante el tiempo de vida de la opción  $X_1 < S_t < X_2$ . Si para algún  $t_0$  el precio de la acción alcanza a  $X_1$  ó  $X_2$  entonces la opción pierde su valor.

**Opción tipo B:** La opción se lleva a cabo sii durante el tiempo de vida de la opción  $S_t > X_1$  y  $S_t$  alcanza antes a la barrera superior  $X_2$ . Se sugiere que  $X_1 < S_0 < X_2$ .

**Opción tipo C:** La opción se lleva a cabo sii durante el tiempo de vida de la opción  $S_t > X_1$  y el precio de la acción  $S_t$  alcanza la barrera superior  $X_2$  por arriba. Se sugiere que  $S_0 > X_2$ .

#### 4. OPCIÓN CALL EUROPEO CON BARRERAS DE TIPO A

En este trabajo sólo se trata con opciones del tipo A. El problema de un call europeo con dos barreras es el siguiente:

$$\begin{cases} \frac{\partial C_A}{\partial t} + \frac{1}{2}\mu^2 S^2 \frac{\partial^2 C_A}{\partial S^2} + rS \frac{\partial C_A}{\partial S} - rC_A = 0, \\ X_1 < S < X_2, t < T, \\ C_A(S, T) = (S - K)_+, \\ C_A(X_1, t) = 0, \\ C_A(X_2, t) = 0. \end{cases}$$

Se hace el siguiente cambio de variables

$$S = Ke^x, t = T - \frac{\tau}{\frac{1}{2}\mu^2}, C_A(S, t) = Ke^{\alpha x + \beta \tau} u(x, \tau),$$

y se obtiene

$$(1) \quad \begin{cases} \frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2}, x_1 < x < x_2, \tau > 0, x_{1,2} = \ln\left(\frac{X_{1,2}}{K}\right), \\ u(x, 0) = (e^{(1-\alpha)x} - e^{-\alpha x})_+ := g(x), x_1 < x < x_2, \\ u(x_1, t) = 0, \\ u(x_2, t) = 0. \end{cases}$$

Aplicando la Transformada de Laplace para la variable  $\tau$ , se tiene

$$\Phi(x, \omega) = \int_0^\infty u(x, \tau) e^{-\omega \tau} d\tau.$$

Aplicando la transformada de Laplace a la parte izquierda de la ecuación (1) (y en las condiciones iniciales, finales y de frontera) se obtiene que,

$$\begin{aligned} \int_0^\infty \frac{\partial u}{\partial \tau}(x, \tau) e^{-\omega \tau} d\tau &= \int_0^\infty e^{-\omega \tau} du(x, \tau) \\ &= e^{-\omega \tau} u(x, \tau) \Big|_0^\infty \\ &+ \omega \int_0^\infty u(x, \tau) e^{-\omega \tau} d\tau \\ &= -u(x, 0) + \omega \Phi(x, \omega) \end{aligned}$$

Aplicando la transformada de Laplace a 1,3,4 y tomando en cuenta (2) se obtiene

$$(2) \quad \frac{\partial^2 \Phi}{\partial x^2}(x, \omega) - \omega \Phi(x, \omega) = g(x)$$

$$(3) \quad \Phi(x_1, \omega) = 0$$

$$(4) \quad \Phi(x_2, \omega) = 0.$$

5. SOLUCIÓN DEL PROBLEMA (2)-(4)

**Paso 1 : Solución fundamental**

$$\frac{\partial^2 \Phi_{x_0}}{\partial x^2}(x, \omega) - \omega \Phi_{x_0}(x, \omega) = \delta(x - x_0), x_1 < x_0 < x_2,$$

$$\Phi_{x_0}(x_1, \omega) = 0,$$

$$\Phi_{x_0}(x_2, \omega) = 0,$$

donde  $\delta(x - x_0)$  es la función delta de Dirac con soporte en el punto  $x_0$ .

De acuerdo a la fórmula de Lagrange

$$(5) \quad \Phi_{x_0}(x, \omega) = -\frac{\Phi_1(x_<, \omega) \Phi_2(x_>, \omega)}{W(\Phi_1, \Phi_2)},$$

donde,

$$x_< = \min(x, x_0), \quad x_> = \max(x, x_0),$$

$\Phi_{1,2}(x, \omega)$  son las soluciones de los siguientes problemas

$$(6) \quad \begin{cases} \frac{\partial^2 \Phi_1}{\partial x^2} - \omega \Phi_1 = 0, \\ \Phi_1(x_1, \omega) = 0 \end{cases},$$

$$(7) \quad \begin{cases} \frac{\partial^2 \Phi_2}{\partial x^2} - \omega \Phi_2 = 0, \\ \Phi_2(x_2, \omega) = 0 \end{cases},$$

y  $W(\Phi_1, \Phi_2)$  es el Wronskiano

$$W(\Phi_1, \Phi_2) = \begin{vmatrix} \Phi_1 & \Phi_2 \\ \frac{\partial \Phi_1}{\partial z} & \frac{\partial \Phi_2}{\partial z} \end{vmatrix}.$$

Algunas soluciones de las ecuaciones (6) y (7) son

$$\Phi_1(x, \omega) = \text{senh} \{(x - x_1)\sqrt{\omega}\},$$

y

$$\Phi_2(x, \omega) = \text{senh} \{(x - x_2)\sqrt{\omega}\},$$

donde,

$$\text{senhu} = (e^u - e^{-u})/2, \quad \text{coshu} = (e^u + e^{-u})/2.$$

Entonces

$$W(\Phi_1, \Phi_2) = \sqrt{\omega} \text{senh} \{(x_2 - x_1)\sqrt{\omega}\}.$$

Así, se obtiene (de acuerdo a (5)) que la solución fundamental tiene la siguiente forma

$$\Phi_{x_0}(x, \omega) = -\frac{\operatorname{senh}\{(x_{<} - x_1)\sqrt{\omega}\} \operatorname{senh}\{(x_{>} - x_2)\sqrt{\omega}\}}{\sqrt{\omega} \operatorname{senh}\{(x_2 - x_1)\sqrt{\omega}\}}.$$

**Paso 2 : Solución del problema (2)-(4)**

$$\Phi(x, \omega) = \int_{x_1}^{x_2} \Phi_x(u, \omega) g(u) du, \quad x \in (x_1, x_2).$$

Aplicando la Transformada de Laplace Inversa se obtiene la solución del problema (1)-(4)

$$(8) \quad u(x, \tau) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \Phi(x, \omega) e^{\tau\omega} d\omega,$$

donde  $c \leq 0$ .

Cambiando el orden de integración se obtiene

$$(9) \quad u(x, \tau) = \frac{1}{2\pi i} \int_{x_1}^{x_2} g(u) \int_{c-i\infty}^{c+i\infty} \Phi_x(u, \omega) e^{\tau\omega} d\omega du.$$

Calculando la integral anterior

$$I_x(u) := -\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{\operatorname{senh}\{(u_{<} - x_1)\sqrt{\omega}\} \operatorname{senh}\{(u_{>} - x_2)\sqrt{\omega}\}}{\sqrt{\omega} \operatorname{senh}\{(x_2 - x_1)\sqrt{\omega}\}} d\omega.$$

con

$$u_{<} = \min(u, x), \quad u_{>} = \max(u, x).$$

De acuerdo al teorema del residuo se tiene que

$$I_x(u) = -\sum_{j=1}^{\infty} \operatorname{Res} \Phi_x(u, \omega) e^{\tau\omega} |_{\omega=\omega_j},$$

donde  $\omega_j$  son las soluciones de la siguiente ecuación

$$\operatorname{senh}\{(x_2 - x_1)\sqrt{\omega}\} = 0.$$

Así,

$$\omega_j = -\left(\frac{\pi j}{x_2 - x_1}\right)^2, \quad j = 1, 2, \dots$$

Ahora calculemos los residuos

$$I_x(u) = -\sum_{j=1}^{\infty} \frac{\operatorname{senh}\{(u_{<} - x_1)\sqrt{\omega_j}\} \operatorname{senh}\{(u_{>} - x_2)\sqrt{\omega_j}\}}{\frac{x_2 - x_1}{2} \frac{\sqrt{\omega_j}}{\sqrt{\omega_j}} \operatorname{cosh}\{(x_2 - x_1)\sqrt{\omega_j}\}} e^{\tau\omega_j},$$

i.e.,

$$I_x(u) = \frac{2}{x_2 - x_1} \times \sum_{j=1}^{\infty} (-1)^j e^{\tau \omega_j} \sin \{(u_{<} - x_1) \beta_j\} \sin \{(u_{>} - x_2) \beta_j\},$$

con

$$\beta_j = -i\sqrt{\omega_j} = \frac{\pi j}{x_2 - x_1},$$

y

$$\begin{aligned} \cosh \{(x_2 - x_1) \sqrt{\omega_j}\} &= \cos \{(x_2 - x_1) \beta_j\} \\ &= \cos(\pi j) = (-1)^j. \end{aligned}$$

De acuerdo a (9) se obtiene

$$\begin{aligned} u(x, \tau) &= \frac{2}{x_2 - x_1} \sum_{j=1}^{\infty} (-1)^j \times \\ &e^{\tau \omega_j} \left( \int_{x_1}^{x_2} g(u) \operatorname{sen} \{(u - x_1) \beta_j\} \operatorname{sen} \{(x - x_2) \beta_j\} du \right. \\ &\left. + \int_{x_1}^{x_2} g(u) \operatorname{sen} \{(u - x_2) \beta_j\} \operatorname{sen} \{(x - x_1) \beta_j\} du \right). \end{aligned}$$

Note que

$$\operatorname{sen} \{(x - x_1) \beta_j\} = (-1)^j \operatorname{sen} \{(x - x_2) \beta_j\}.$$

De forma análoga se obtiene

$$\operatorname{sen} \{(u - x_1) \beta_j\} = (-1)^j \operatorname{sen} \{(u - x_2) \beta_j\}.$$

Así, se obtiene

$$(10) \quad u(x, \tau) = \frac{2}{x_2 - x_1} \sum_{j=1}^{\infty} e^{\tau \omega_j} \operatorname{sen} \{(x - x_2) \beta_j\} \times \int_{x_1}^{x_2} g(u) \operatorname{sen} \{(u - x_2) \beta_j\} du,$$

donde  $g(u) = (e^{(1-\alpha)u} - e^{-\alpha u})_+$ .

Las integrales de la clase

$$I_j(a) = \int_{\widehat{x}_1}^{x_2} e^{au} \operatorname{sen} \{(u - x_2) \beta_j\} du,$$

donde

$$a \in \mathbb{R}, \widehat{x}_1 = \begin{cases} x_1, & \text{si } x_1 > 0 \\ 0, & \text{si } x_1 \leq 0. \end{cases}$$

son

$$\begin{aligned}
I_j(a) &= \operatorname{Im} \left( \int_{\widetilde{x}_1}^{x_2} e^{au+i(u-x_2)\beta_j} du \right) \\
&= \frac{ae^{a\widetilde{x}_1} \sin\{\beta_j(x_2 - \widetilde{x}_1)\}}{(a^2 + \beta_j^2)} \\
&\quad - \frac{-\beta_j(e^{ax_2} - e^{a\widetilde{x}_1} \cos\{\beta_j(x_2 - \widetilde{x}_1)\})}{(a^2 + \beta_j^2)}.
\end{aligned}$$

Así, si  $x_1 > 0$  entonces  $\widetilde{x}_1 = x_1$ ,  $\sin\{\beta_j(x_2 - x_1)\} = 0$ ,  $\cos\{\beta_j(x_2 - x_1)\} = (-1)^j$ . Esto es, se tiene

$$(11) \quad I_j(a) = \begin{cases} -\beta_j \frac{(e^{ax_2} - (-1)^j e^{ax_1})}{a^2 + \beta_j^2} & \text{si } x_1 > 0, \\ \frac{a \sin\{\beta_j x_2\} - \beta_j (e^{ax_2} - \cos\{\beta_j x_2\})}{(a^2 + \beta_j^2)} & \text{si } x_1 < 0. \end{cases}$$

Ahora, se pueden escribir las soluciones del problema (1)-(4) en la forma (vea(10)-(11))

$$\begin{aligned}
u(x, \tau) &= -\frac{2}{x_2 - x_1} \times \\
&\quad \sum_{j=1}^{\infty} (I_j(1 - \alpha) - I_j(-\alpha)) e^{\tau\omega_j} \operatorname{sen}\{(x - x_2)\beta_j\}.
\end{aligned}$$

Haciendo el cambio inverso de variable,

$$x = \ln(S/K), \quad \tau = \frac{1}{2}\mu^2(T - t),$$

$$C_A(S, t) = K e^{\alpha x + \beta \tau} u(x, \tau),$$

se obtiene el precio de una *Opción Call del tipo A*:

$$\begin{aligned}
C_A(S, t) &= \frac{2K \left(\frac{S}{K}\right)}{\ln\left(\frac{X_1}{X_2}\right)} \sum_{j=1}^{\infty} (I_j(-\alpha) - I_j(1 - \alpha)) \times \\
&\quad e^{-\frac{1}{2}\mu\beta_j(T-t)} \operatorname{sen}\left(\beta_j \ln\left(\frac{X_2}{S}\right)\right)
\end{aligned}$$

con  $X_1 \leq S \leq X_2$ , donde,

$$\beta_j = \frac{\pi j}{\ln(X_2/X_1)}, \quad \alpha = -\frac{1}{2} \left( \frac{2r}{\mu} - 1 \right), \quad \beta = -\frac{1}{4} \left( \frac{2r}{\mu} + 1 \right)^2,$$

y

$$I_j(a) = \begin{cases} -\beta_j \frac{(e^{ax_2} - (-1)^j e^{ax_1})}{a^2 + \beta_j^2}, & x_1 > 0, \\ \frac{a \operatorname{sen}\{\beta_j x_2\} - \beta_j (e^{ax_2} - \cos\{\beta_j x_2\})}{a^2 + \beta_j^2}, & x_1 \leq 0. \end{cases}$$

La anterior es la ecuación de Black-Scholes con condiciones finales y de frontera,  $X_1, X_2$  son las barreras y  $(S - K)_+$  es el pago final.

## REFERENCIAS

- [1] Cox J. C. y Rubinstein M., *Option Markets*. Prentice Hall, New Jersey, 1985.
- [2] Fusai G., Abrahams D., Sgarra C., *An exact analytical solution for discrete barrier options*. Finance and Stochastics, 1-26, 2006.
- [3] German H., *Pricing and Hedging Double-Barrier Options: A probabilistic Approach*. Math. Finance 6:4, 365-378, 1996.
- [4] Karatzas I., *Lectures on the Mathematics of Finance*, 1997.
- [5] Karatzas I. y Yor M., *Methods of Mathematical Finance*, 1995.
- [6] Merton R., *Theory of Rational Option Pricing*, 1973.
- [7] Musiela M., y Rutkowski M., *Martingale Methods in Financial Modelling*, 2007.
- [8] Nelken I., *Handbook of Exotic Options*, 1995
- [9] Pelsler A. *Pricing double barrier options using Laplace transform*. Finance Stochastics, 4:1, 95-104, 2000.
- [10] Poulsen R., *Barrier Option and Their Static Hedges: Simple derivations and Extensions*. 2006
- [11] Shreve S., *Stochastics calculus for finance I (The binomial asset pricing model)*, 2004.
- [12] Shreve S., *Stochastics calculus for finance II (The binomial asset pricing model)*, 2004.
- [13] Wilmott P., Dewynne J. and Howison S., *Option Pricing: Mathematical Model and computation* 1993
- [14] Wilmott P., Dewynne J. and Howison S., *The Mathematics of financial derivatives: A Student Introduction*, 1995
- [15] Zhang P., *Exotic Options: A Guide to the Second Generation Option*, 1996.

Facultad de Ciencias Físico Matemáticas, BUAP.

Av. San Claudio y 18 Sur, Col. San Manuel, Ciudad Universitaria

C.P. 72570

Tel: (01222)2295500 Ext. 7550,7577

carlos\_cpj@hotmail.com

ftajonar@fcfm.buap.mx



# Sesión de Topología

Quinta Gran Semana Nacional de la Matemática



# PROYECCIONES ORBITALES COMO $G$ -FIBRACIONES APROXIMATIVAS

ALEXANDER BYKOV  
FCFM - BUAP

RESUMEN. Consideramos las proyecciones orbitales  $q_E : E \rightarrow E/G$  de  $G$ -espacios  $E$  que tienen un solo tipo de órbitas. Mostramos que  $q_E$  admite una  $G$ -ANR-resolución que consta de  $G$ -fibraciones si  $G$  es un grupo compacto metrizable.

## 1. INTRODUCCIÓN

El concepto de una  $G$ -fibración representa la versión equivariante del concepto de una fibración de Hurewicz, de esta manera una  $G$ -función se llama  $G$ -fibración si tiene la propiedad de la levantamiento de  $G$ -homotopías para la clase de todas las  $G$ -espacios.

En la teoría homotópica equivariante es bien conocido el siguiente resultado: si  $G$  es un grupo *compacto de Lie* y  $E$  es un  $G$ -espacio paracompacto con un solo tipo de órbitas, entonces la proyección orbital  $q_E : E \rightarrow E/G$  es una  $G$ -fibración (véase, por ejemplo, [11]). Naturalmente surge la pregunta: ¿Qué se puede decir sobre la proyección orbital  $q_E$  cuando  $G$  no es necesariamente un grupo de Lie?

En el presente trabajo intentamos dar una respuesta a esta pregunta para el caso de los grupos  $G$  compactos metrizable. Según nuestro resultado principal (el teorema 6.11), en este caso,  $q_E$  admite una  $G$ -ANR-resolución que consta de  $G$ -fibraciones. En efecto, este teorema afirma que la proyección orbital  $q_E$  puede aproximarse, de una manera suficiente buena, por medio de  $G$ -fibraciones  $q_i$ . Además, es posible elegir como  $q_i$  proyecciones orbitales de  $G$ -ANR espacios. Por lo tanto, las llamamos  *$G$ -fibraciones aproximativas* a tales proyecciones orbitales en el título de este artículo. Sin embargo, el término más adecuado podría ser a lo mejor " *$G$ -fibraciones de shape*" ya que en el caso *no equivariante* tales funciones continuas se llaman en inglés "shape fibrations". En realidad, en la demostración del resultado principal aprovechamos las ideas de la teoría de shape, es decir, consideramos las aproximaciones tanto de grupos  $G$  como de  $G$ -espacios mediante  $G$ -ANR espacios. Los fundamentos de la teoría equivariante de shape están dados en los artículos [4] y [5].

## 2. PRELIMINARES

A lo largo del presente trabajo el símbolo  $G$  designa un grupo *compacto de Hausdorff*. Todos los  $G$ -espacios que consideramos son *metrizables*.

Las nociones básicas de la teoría de  $G$ -espacios (la teoría de grupos de transformaciones) se pueden encontrar en los libros de Bredon [7], de tom Dieck [11] u otro

excelente libro de S. de Neymet [8]. Sin embargo para comodidad del lector, en lo que sigue, enunciaremos algunas definiciones y resultados importantes.

Por un  $G$ -espacio entenderemos un espacio topológico  $X$  dotado de una acción continua  $\cdot : G \times X \rightarrow X$  de  $G$  sobre  $X$ . Designamos por  $g \cdot x$  o  $gx$  la imagen de la pareja  $(g, x)$  bajo la acción. Si  $A \subseteq X$ , por  $G(A)$  denotamos la imagen de  $G \times A$  bajo la acción, esto es el conjunto  $\{ga \mid g \in G, a \in A\}$ . Si  $H$  es un subgrupo de  $G$ , el  $G$ -espacio  $X$  se puede tratar como un  $H$ -espacio.

Un subconjunto  $A$  de un  $G$ -espacio  $X$  es denominado *invariante* si  $G(A) = A$ . Notemos que si  $A \subseteq X$  es invariante, entonces la acción (sobre  $X$ ) restringida al conjunto  $G \times A$  es una acción sobre  $A$ .

Para un subgrupo  $H \subset G$ , siempre consideramos el conjunto  $G/H = \{gH \mid g \in G\}$  de clases laterales como un  $G$ -espacio con la acción por traslaciones izquierdas.

Si  $X$  y  $Y$  son  $G$ -espacios, una función continua  $f : X \rightarrow Y$  se llama *equivariante* o  $G$ -función si, para cada  $g \in G$  y toda  $x \in X$ , cumple  $f(gx) = gf(x)$ . Los homeomorfismos equivariantes son denominados  $G$ -equivalecias.

Una homotopía  $H : X \times I \rightarrow Y$ , donde  $I = [0, 1]$ , se llama *equivariante* ó  $G$ -homotopía si  $H(gx, t) = gH(x, t)$ , para cada  $(x, t) \in X \times I$  y toda  $g \in G$ . Así,  $H$  es una  $G$ -función considerando  $X \times I$  con la acción  $g(x, t) = (gx, t)$  o, equivalentemente, para cada  $t \in I$  la función  $f_t : X \rightarrow Y$ , dada por  $f_t(x) = H(x, t)$ , es equivariante.

Si  $X$  es un  $G$ -espacio y  $x \in X$ , el conjunto  $G(x) = \{gx \in X \mid g \in G\}$  se llama la  $G$ -órbita o simplemente *órbita de  $x \in X$* . Es fácil ver que se tiene una partición de  $X$  en órbitas. El conjunto de todas las órbitas, dotado con la topología cociente, se denota como  $X/G$  y es llamado *espacio orbital*. A la proyección natural  $q_X : X \rightarrow X/G$  (dada por  $q_X(x) = G(x)$ ) la llamaremos *proyección orbital*. Es continua (por la definición misma de la topología cociente) y abierta. Además, si  $G$  es compacto y  $X$  es de Hausdorff,  $q$  es cerrada y, por lo tanto, es perfecta.

Siempre podemos considerar el espacio orbital  $X/G$  como un  $G$ -espacio con la acción trivial. En este caso la proyección orbital  $q_X$  es una  $G$ -función.

Si  $f : X \rightarrow Y$  es una  $G$ -función, entonces existe una única función continua  $f/G : X/G \rightarrow Y/G$ , se llama la función *inducida por  $f$* , que hace conmutativo el diagrama

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ q_X \downarrow & & \downarrow q_Y \\ X/G & \xrightarrow{f/G} & Y/G \end{array}$$

donde  $q_X$  y  $q_Y$  son las proyecciones orbitales, y que se define como:

$$(f/G)(G(x)) = G(f(x)), \quad x \in X$$

Si  $N$  es un subgrupo normal cerrado de  $G$  y  $X$  un  $G$ -espacio, el espacio orbital  $X/N$  se puede considerar como  $G/N$ -espacio al definir la acción de  $G/N$  sobre  $X/N$  por  $(gN) * (N(x)) = N(gx)$ , donde  $gN \in G/N$  y  $N(x)$  es la  $N$ -órbita de  $x \in X$  (También  $X/N$  es un  $G$ -espacio bajo la acción  $g * N(x) = N(gx)$ ).

Sean  $H$  un subgrupo de  $G$  y  $X$  un  $H$ -espacio. El *producto torcido*  $G \times_H X$  es el espacio orbital del  $H$ -espacio  $G \times X$ , donde  $H$  actúa como sigue:

$$h \cdot (g, x) = (gh^{-1}, hx)$$

El producto torcido  $G \times_H X$  es un  $G$ -espacio con la acción

$$g'[g, x] = [g'g, x]$$

donde por  $[g, x]$  se denota el  $H$ -órbita de  $(g, x) \in G \times X$ .

Para un punto  $x$  de un  $G$ -espacio  $X$ , el subgrupo  $G_x = \{g \in G \mid gx = x\}$  se llama el *estabilizador* o el *grupo de isotropía* de  $x$ . Para un subgrupo  $H \subset G$ , el conjunto  $X^H = \{x \in X \mid H \subset G_x\}$  se llama el *conjunto de puntos  $H$ -fijos*.

Denotemos por  $(H)$  la clase conjugada del subgrupo  $H$  de  $G$ , es decir  $(H) = \{gHg^{-1} \mid g \in G\}$ . Cada clase conjugada  $(H)$  se llama *tipo de  $G$ -órbitas* por el siguiente motivo. Si  $X$  un  $G$ -espacio y  $\alpha = G(x)$  es la órbita de  $x \in X$ , entonces  $(G_x) = \{G_z \mid z \in \alpha\}$  y  $\alpha$  es  $G$ -equivalente a  $G/G_z$  para cada  $z \in \alpha$ . Vamos a designar por  $[\alpha]$  la clase conjugada de  $G_x$ ,  $[\alpha]$  se llama el *tipo de órbita de  $\alpha \in X/G$* .

En el presente trabajo, nos interesarán los  $G$ -espacios con un solo tipo de órbitas. Un  $G$ -espacio  $X$  tiene *un solo tipo de órbitas*  $(H)$  si y sólo si  $(G_x) = (H)$  para cada punto  $x \in X$ . Un  $G$ -espacio se llama *libre* (y la acción correspondiente se llama *libre*) si tiene un solo tipo de órbitas  $(\{e\})$ .

Un subconjunto invariante  $A$  de un  $G$ -espacio  $X$  se llama un *retracto equivariante* o  *$G$ -retracto* de  $X$  si existe una  $G$ -función  $r : X \rightarrow A$  tal que  $r(a) = a$ , para toda  $a \in A$ .

Un  $G$ -espacio  $X$  es denominado  *$G$ -retracto absoluto* o, en forma abreviada  *$G$ -AR*, si cada vez que se tenga un encaje  $X \hookrightarrow Z$  de  $X$  como un subespacio invariante cerrado en un  $G$ -espacio  $Z$ ,  $X$  es un  $G$ -retracto de  $Z$ .

Un  $G$ -espacio  $X$  se llama  *$G$ -retracto absoluto de vecindad* o, en forma abreviada  *$G$ -ANR*, si cada vez que se tenga un encaje  $X \hookrightarrow Z$  de  $X$  como un subespacio invariante cerrado en un  $G$ -espacio  $Z$ ,  $X$  es un  $G$ -retracto de alguna vecindad invariante  $U$  de  $X$  en  $Z$ .

Los espacios  $G$ -ANRs y  $G$ -ARs poseen la siguiente propiedad *extensorial* (véase [8], Proposición 6.14):  $Y$  es un  $G$ -ANR (resp.  $G$ -AR) si y sólo si para todo  $G$ -espacio metrizable  $X$  y cualquier subconjunto invariante cerrado  $A$  de  $X$ , cada  $G$ -función  $f : A \rightarrow Y$  admite una  $G$ -extensión  $\bar{f} : U \rightarrow Y$ , donde  $U$  es una vecindad invariante de  $A$  en  $X$  (resp. admite una  $G$ -extensión  $\bar{f} : X \rightarrow Y$ ).

Cabe mencionar dos más propiedades importantes de  $A(N)R$  que utilizaremos en este trabajo.

2.1. TEOREMA. ([3],[4]) Sea  $G$  un grupo de Hausdorff compacto.

Si  $X$  es un  $G$ -espacio metrizable, entonces existe  $M$  un  $G$ -espacio lineal normado  $G$ -AR, y una inclusión equivariante,  $i : X \hookrightarrow M$ , tal que la imagen  $i(X)$  es cerrado en  $M$

2.2. TEOREMA. ([2]) Sean  $G$  un grupo compacto de Hausdorff y  $N$  un subgrupo normal cerrado en  $G$ . Si  $X$  es un  $G$ - $A(N)R$ -espacio, entonces  $X/N$  es un  $G/N$ - $A(N)R$ -espacio. En particular,  $X/G$  es un  $A(N)R$ -espacio.

## 3. G-FIBRACIONES Y EL TEOREMA DE PALAIS

La siguiente definición nos proporciona una versión equivariante del concepto de una fibración de Hurewicz.

3.1. DEFINICIÓN. Una  $G$ -función  $p : E \rightarrow B$  se llama  $G$ -fibración si para todo  $G$ -espacio  $X$  y cada diagrama conmutativo de  $G$ -funciones

$$\begin{array}{ccc} X & \xrightarrow{h} & E \\ \downarrow \partial_0 & & \downarrow p \\ X \times I & \xrightarrow{H} & B \end{array}$$

donde  $H$  es una  $G$ -homotopía y  $\partial_0(x) = (x, 0)$  para cada  $x \in X$ , existe una  $G$ -homotopía  $\hat{H} : X \times I \rightarrow E$  tal que  $\hat{H} \circ \partial_0 = h$  y  $p \circ \hat{H} = H$ .

En relación con el concepto de  $G$ -fibración debemos recordar el famoso teorema (de levantamiento de homotopías) de Palais; para formular este teorema necesitamos las siguientes definiciones.

3.2. DEFINICIÓN. Sean  $X$  y  $Y$   $G$ -espacios.

- (1) Una función continua de los espacios orbitales  $\tilde{f} : X/G \rightarrow Y/G$  preserva estructura orbital si  $[\tilde{f}(\alpha)] = [f(\alpha)]$  para cada  $\alpha \in X/G$ ,
- (2) una  $G$ -función  $f : X \rightarrow Y$  preserva estructura orbital si  $f/G$  preserva estructura orbital.

Darémos aparte la definición análoga para  $G$ -homotopías aunque es una consecuencia de la definición anterior.

3.3. DEFINICIÓN. Sean  $X$  y  $Y$   $G$ -espacios.

- (1) Una homotopía  $\tilde{H} : X/G \times I \rightarrow Y/G$  preserva estructura orbital si  $[\tilde{H}(\alpha, t)] = [\alpha]$  para cada  $\alpha \in X/G$  y cada  $t \in I$ .
- (2) Una  $G$ -homotopía  $H : X \times I \rightarrow Y$  preserva estructura orbital si  $H/G : X/G \times I \rightarrow Y/G$  preserva estructura orbital.

3.4. OBSERVACIÓN. Sea  $f : X \rightarrow Y$  una  $G$ -función.

- (1) Si  $f$  es *isovariante*, es decir  $G_{f(x)} = G_x$  para cada  $x \in X$ , entonces  $f$  preserva estructura orbital.
- (2) Si  $f$  es una  $G$ -equivalencia, entonces  $f$  preserva estructura orbital. (Notemos que una  $G$ -equivalencia no es, en general, isovariante.)
- (3)  $f$  preserva estructura orbital si sólo si la restricción de  $f$  en cada  $G$ -órbita  $\alpha$  de  $X$  realiza una  $G$ -equivalencia de  $\alpha$  con su imagen  $f(\alpha)$ . (Recordemos que  $f(\alpha)$  es siempre una  $G$ -órbita de  $Y$  para cualquiera  $G$ -función  $f$ )

3.5. TEOREMA. (R.Palais, 1960, [9]) Sea  $G$  un grupo compacto de Lie. Sean  $X$  y  $Y$   $G$ -espacios. Si en el diagrama conmutativo de  $G$ -funciones

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ \delta_0 \downarrow & & \downarrow q_Y \\ X/G \times I & \xrightarrow{\tilde{F}} & Y/G \end{array}$$

donde  $\delta_0(x) = (G(x), 0)$ , la homotopía  $\tilde{F}$  preserva estructura orbital, entonces existe una  $G$ -homotopía  $F : X \times I \rightarrow Y$  tal que  $F \circ \partial_0 = f$  y  $F/G = \tilde{F}$ , es decir el diagrama

$$\begin{array}{ccc} X \times I & \xrightarrow{F} & Y \\ q_X \times id_I \downarrow & & \downarrow q_Y \\ X/G \times I & \xrightarrow{\tilde{F}} & Y/G \end{array}$$

es conmutativo. En particular,  $F$  preserva estructura orbital.

3.6. OBSERVACIÓN. En el teorema 3.5, como en el resto del presente trabajo, suponemos que los  $G$ -espacios son metrizables, pero, en efecto, el teorema ya es cierto bajo la condición que  $X/G$  es un espacio hereditariamente paracompacto respecto subconjuntos abiertos. (véase [7], Teorema 7.3)

3.7. DEFINICIÓN. Un diagrama conmutativo de  $G$ -funciones

$$\begin{array}{ccc} E' & \xrightarrow{f'} & E \\ p' \downarrow & & \downarrow p \\ B' & \xrightarrow{f} & B \end{array}$$

se llama *cuadrado universal* o bien *diagrama pull-back*, si satisface la siguiente *propiedad universal*: cualesquiera que sean un  $G$ -espacio  $Z$  y  $G$ -funciones  $u : Z \rightarrow B'$ ,  $v : Z \rightarrow E$  tales que  $f \circ u = p \circ v$ , existe una única  $G$ -función  $h : Z \rightarrow E'$  que hace conmutativo el diagrama

$$\begin{array}{ccccc} & & & & Z \\ & & & & \swarrow v \\ & & & & \searrow h \\ & & & & \downarrow u \\ & & & & E' \xrightarrow{f'} E \\ & & & & \downarrow p' \quad \downarrow p \\ & & & & B' \xrightarrow{f} B \end{array}$$

En este caso el  $G$ -espacio  $E'$  también se llama *producto fibrado* o bien, *pull-back* del diagrama  $B' \xrightarrow{f} B \xleftarrow{p} E$ .

3.8. OBSERVACIÓN. El pull-back  $E'$  de cada diagrama  $B' \xrightarrow{f} B \xleftarrow{p} E$  existe y es único salvo  $G$ -equivalencia. Se puede definirlo en la *forma explicita* como sigue:

$$E' = \{(b', e) \in B' \times E \mid f(b') = p(e)\}$$

$$p'((b', e)) = b', \quad f'(b', e) = e$$

Observemos que  $E'$  es un  $G$ -espacio por ser un subconjunto invariante de  $B' \times E$  con respecto a la acción diagonal:  $g \cdot (b', e) = (g \cdot b', g \cdot e)$ .

3.9. PROPOSICIÓN. Sea  $f : X \rightarrow Y$  es una  $G$ -función. Entonces el diagrama

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ q_X \downarrow & & \downarrow q_Y \\ X/G & \xrightarrow{f/G} & Y/G \end{array}$$

es un diagrama pull-back si y sólo si  $f$  preserva estructura orbital.

DEMOSTRACIÓN. El diagrama dado es pull-back si y sólo si, en el diagrama conmutativo

$$\begin{array}{ccccc} X & & & & \\ & \searrow f & & & \\ & & X' & \xrightarrow{f'} & Y \\ & \searrow h & \downarrow q' & & \downarrow q_Y \\ & & X/G & \xrightarrow{f/G} & Y \\ & \searrow q_X & & & \end{array}$$

$h : X \rightarrow X'$  es una  $G$ -equivalencia, donde  $X'$  (junto con  $q'$  y  $f'$ ) es el pull-back, en su forma explícita, de  $X/G \xrightarrow{f/G} Y/G \xleftarrow{q_Y} Y$ .

Sea  $x' \in X'$ . Puesto que  $x' = (\alpha, y)$ , donde  $\alpha \in X/G$  y  $y \in Y$  son tales que  $(f/G)(\alpha) = G(y)$ , tenemos  $G_{x'} = G_y = G_{f'(x')}$ , en otras palabras  $f'$  es isovariante.

Si  $h$  es una  $G$ -equivalencia, entonces  $f$  preserva estructura orbital como composición de los  $G$ -funciones  $h$  y  $f'$  que preservan estructura orbital.

Por otro lado, si  $f$  preserva estructura orbital, entonces  $h$  preserva estructura orbital ya que  $f'$  es isovariante. Para cada  $x \in X$ , la restricción

$$h|_{G(x)}^{G(x')} : G(x) \rightarrow G(x'),$$

donde  $x' = h(x)$ , puede ser interpretada como la  $G$ -función  $G/G_x \rightarrow G/G_{x'}$ . Como  $h$  preserva estructura orbital, los subgrupos  $G_x$  y  $G_{x'}$  son conjugados, y, por consiguiente,  $h|_{G(x)}^{G(x')}$  debe ser una  $G$ -equivalencia (véase [7], Teorema 4.2 o [8], Proposición 5.10). De ahí  $h$  es inyectiva. Por supuesto,  $h$  es sobreyectiva y, además, es perfecta ya que  $q_X$  es una función perfecta y  $q_X = q' \circ h$ . Por lo tanto,  $h$  es una  $G$ -función biyectiva y cerrada, es decir  $h$  es una  $G$ -equivalencia.  $\square$

3.10. OBSERVACIÓN. El diagrama que se construye en el teorema de Palais (el Teorema 3.5) es un diagrama pull-back.

Para probar la proposición final de esta sección necesitamos el siguiente corolario (en realidad, es equivalente) del Teorema de Palais.

3.11. PROPOSICIÓN. Sea  $Y$  un  $G$ -espacio tal que  $G$  es un grupo compacto de Lie. Si  $H : Z \times I \rightarrow Y/G$  una homotopía (usual) tal que  $[H(z, t)] = [H(z, 0)]$  para cada  $(z, t) \in Z \times I$ , entonces existe una  $G$ -homotopía  $H : X \times I \rightarrow Y$  tal que  $X/G = Z$  y el diagrama

$$\begin{array}{ccc} X \times I & \xrightarrow{H'} & Y \\ q_X \times id_I \downarrow & & \downarrow q_Y \\ Z \times I & \xrightarrow{H} & Y/G \end{array}$$

es un diagrama pull-back. En particular,  $H'$  preserva estructura orbital.

DEMOSTRACIÓN. Por supuesto, consideramos  $Z$  como un  $G$ -espacio con la acción trivial. Sea  $X$  el pull-back de  $Z \xrightarrow{h} Y/G \xleftarrow{q_Y} Y$  en su forma explícita, donde  $h(z) = H(z, 0)$ . En el correspondiente diagrama pull-back

$$\begin{array}{ccc} X & \xrightarrow{h'} & Y \\ p \downarrow & & \downarrow q_Y \\ Z & \xrightarrow{h} & Y/G \end{array}$$

la función  $p$  es sobreyectiva y abierta como restricción de la proyección  $Z \times Y \rightarrow Z$ . Además, es fácil verificar que los conjuntos  $p^{-1}(z)$ ,  $z \in Z$ , son exactamente las  $G$ -órbitas de  $X$ .

En realidad, si  $x \in X$ , entonces, puesto que  $x$  es de la forma  $(z, y)$  con  $h(z) = G(y)$ ,  $G(x) = \{(z, gy) \in Z \times Y \mid g \in G\} \subseteq p^{-1}(z)$  ya que  $p(z, y) = z$ . Por otro lado, si  $(z', y') \in p^{-1}(z)$ , entonces  $z' = z$  y  $G(y') = h(z') = h(z) = G(y)$ , es decir  $y' = gy$  para algún  $g \in G$ . Concluimos que  $(z', y') \in G(x)$ .

Por consiguiente, podemos identificar  $Z$  con el espacio orbital  $X/G$  y  $p$  con la proyección orbital  $q_X$ .

Sabemos que la función  $h'$  es isovariante y, por lo tanto,

$$[H(z, t)] = [H(z, 0)] = [h(z)] = (G_{h'(x)}) = (G_x) = [z]$$

para  $z \in Z$  y  $x \in X$  con  $p(x) = z$ . De ahí  $H$  preserva estructura orbital y el diagrama pull-back requerido existe en virtud del Teorema 3.5 y de la Proposición 3.9 □

3.12. PROPOSICIÓN. Sea  $G$  un grupo compacto de Lie. Si un  $G$ -espacio  $E$  tiene un solo tipo de órbita, entonces la proyección orbital  $q : E \rightarrow E/G$  es una  $G$ -fibración.

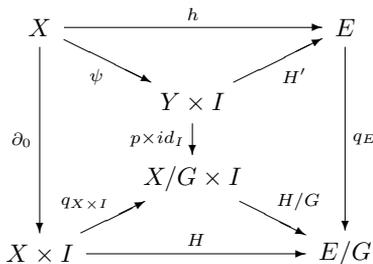
DEMOSTRACIÓN. Sea dado el diagrama conmutativo de  $G$ -funciones

$$\begin{array}{ccc} X & \xrightarrow{h} & E \\ \partial_0 \downarrow & & \downarrow q_E \\ X \times I & \xrightarrow{H} & E/G \end{array}$$

Como el espacio orbital  $E/G$  se considera como un  $G$ -espacio con la acción trivial, la homotopía  $H$  induce la función continua  $H/G : (X \times I)/G \rightarrow E/G$  tal que  $H = H/G \circ q_{X \times I}$ . Por supuesto, consideramos que  $(X \times I)/G = X/G \times I$  y  $q_{X \times I} = q_X \times id_I$ .

Como  $E$  tiene un solo tipo de órbita, por ejemplo  $(K)$ , la homotopía  $H/G : X/G \times I \rightarrow E/G$  desde luego satisface las condiciones de la Proposición 3.11 ya que  $[H/G(\alpha, t)] = (K)$  para cada  $(\alpha, t) \in X/G \times I$ .

Construimos el siguiente diagrama conmutativo



donde  $Y \times I$  es el pull-back de  $X \times I \xrightarrow{H/G} E/G \xleftarrow{q_E} E$  que existe según la Proposición 3.11 y  $\psi$  es la  $G$ -función única definida por la propiedad universal de pull-back.

De la conmutatividad del diagrama

$$(p \times id_I)(\psi(x)) = (q_X \times id_I)(\partial_0(x)) = (q_X(x), 0)$$

y de aquí  $\psi(x) \in (p \times id_i)^{-1}(X/G \times \{0\})$  para cada  $x \in X$ . Concluimos que  $\psi(x) = (\psi'(x), 0)$  para alguna  $G$ -función  $\psi' : X \rightarrow Y$ . Obviamente tiene que ser  $p(\psi'(x)) = q_X(x)$  para cada  $x \in X$ .

Finalmente, definimos  $\widehat{H} : X \times I \rightarrow Y$  como sigue:  $\widehat{H}(x, t) = H'(\psi'(x), t)$  para cada  $(x, t) \in X \times I$ . Entonces  $\widehat{H}$  es una  $G$ -función ya que  $\psi'$  y  $H'$  son  $G$ -funciones. Además,  $\widehat{H}(x, 0) = H'(\psi'(x), 0) = H'(\psi(x)) = h(x)$  y

$$q_E(\widehat{H}(x, t)) = q_E(H'(\psi'(x), t)) = (H/G)(p(\psi'(x)), t) = (H/G)(q_X(x), t) = H(x, t).$$

Así  $\widehat{H}$  es un levantamiento de la homotopía  $H$  que deseabamos encontrar. □

#### 4. CAMBIO DE GRUPOS

Cada homomorfismo de grupos  $\varphi : G' \rightarrow G$  induce el *functor de restricción*

$$\mathcal{F}_\varphi : G\text{-}\mathcal{M} \rightarrow G'\text{-}\mathcal{M}$$

de la categoría de  $G$ -espacios a la categoría de  $G'$ -espacios. Si  $X$  es un a  $G$ -espacio, entonces  $\mathcal{F}_\varphi(X)$  es el mismo espacio topológico  $X$  con la  $G'$  acción  $\cdot$  que se define por  $g \cdot x = \varphi(g)x$ . Análogamente, si  $f : X \rightarrow Y$  es una  $G$ -función,  $\mathcal{F}_\varphi(f)$  es la misma función continua  $f$  considerada como  $G'$ -función. En efecto se cumple  $f(g \cdot x) = f(\varphi(g)x) = \varphi(g)f(x) = g \cdot f(x)$

Señalamos dos casos importantes del functor de restricción:

- (1) Si  $H$  es un subgrupo de un grupo  $G$ , entonces cada  $G$ -espacio puede ser considerado como un  $H$ -espacio y cada  $G$ -función puede ser considerada como una  $H$ -función debido al functor de restricción  $\mathcal{F}_i$  inducido por la inclusión natural  $i : H \hookrightarrow G$ .

- (2) Si  $N$  es un subgrupo normal de un grupo  $G$ , entonces cada  $G/N$ -espacio puede ser considerado como un  $G$ -espacio y cada  $G/N$ -función puede ser considerada como una  $G/N$ -función debido al funtor de restricción  $\mathcal{F}_p$  inducido por la proyección natural  $p : G \twoheadrightarrow G/N$ .

4.1. OBSERVACIÓN. Claramente el caso general del funtor de restricción se reduce a los casos mencionados ya que cada homomorfismo es una composición de un epimorfismo y de un monomorfismo.

Sea  $H$  un subgrupo cerrado de un grupo compacto  $G$ . Observemos que funtor de restricción  $\mathcal{F}_i : G\text{-}\mathcal{M} \rightarrow H\text{-}\mathcal{M}$ , inducido por la inclusión natural  $i : H \hookrightarrow G$ , es un funtor *adjunto derecho* al funtor del producto torcido  $G \times_H - : H\text{-}\mathcal{M} \rightarrow G\text{-}\mathcal{M}$ . Más precisamente, si  $X$  es un  $H$ -espacio y  $Y$  es un  $G$ -espacio, entonces para cada  $H$ -función  $f : X \rightarrow Y$  existe una única  $G$ -función  $\hat{f} : G \times_H X \rightarrow Y$  tal que el diagrama

$$\begin{array}{ccc} X & \xrightarrow{i_X} & G \times_H X \\ & \searrow f & \swarrow \hat{f} \\ & & Y \end{array}$$

es conmutativo, donde la  $H$ -función  $i_X$  se define por  $i_X(x) = [e, x]$ . Obviamente, la  $G$ -función  $\hat{f}$  está dada por  $\hat{f}([g, x]) = gf(x)$ .

En particular, obtenemos la biyección natural

$$H\text{-}\mathcal{M}(X, Y) \longrightarrow G\text{-}\mathcal{M}(G \times_H X, Y), \quad f \mapsto \hat{f}$$

Este hecho usamos en la prueba de la siguiente afirmación.

4.2. PROPOSICIÓN. Sea  $H$  un subgrupo cerrado de un grupo compacto metrizable  $G$ .

- (a) Si  $E$  es un  $G$ -ANR espacio, entonces  $E$  es un  $H$ -ANR espacio.
- (b) Si  $p : E \rightarrow B$  es una  $G$ -fibración, entonces  $p$  es un  $H$ -fibración.

DEMOSTRACIÓN. Recordemos primero que, por lo menos para  $H$ -espacios  $X$  de Hausdorff, la proyección  $G \times X \rightarrow G \times_H X$ ,  $(g, x) \mapsto [g, x]$  es una función cerrada pues  $G$  es compacto. Por lo tanto la  $H$ -función  $i_X : X \rightarrow G \times_H X$  es una inclusión cerrada y, para cada  $H$ -inclusión cerrada  $s : A \hookrightarrow X$ , la  $G$ -función  $G \times_H s : G \times_H A \hookrightarrow G \times_H X$  es una inclusión.

Sea  $p : E \rightarrow B$  una  $G$ -función y sea  $s : A \rightarrow X$  una  $H$ -inclusión cerrada (es decir, una  $H$ -función que es una inclusión cerrada). En toda la demostración consideramos el siguiente diagrama conmutativo

$$\begin{array}{ccc} A & \xrightarrow{f} & E \\ & \searrow i_A & \swarrow \hat{f} \\ & & G \times_H A \\ & & \downarrow \bar{s} \\ & & G \times_H X \\ & \swarrow i_X & \searrow \hat{F} \\ X & \xrightarrow{F} & B \end{array} \quad \begin{array}{c} \\ \\ \\ \\ \\ \downarrow p \end{array}$$

donde  $f$  y  $F$  son  $H$ -funciones y  $\bar{s} = G \times_H s$ . Notemos que la existencia de un relleno

$$\bar{F} : G \times_H X \rightarrow E$$

es equivalente a la existencia de un relleno  $\tilde{F} : X \rightarrow E$  en virtud de la biyección natural mencionada más arriba.

(a). Sea  $A$  un subconjunto cerrado invariante de un  $H$ -espacio  $X$  y sea  $s$  la inclusión correspondiente. En realidad, necesitamos solo la parte superior del diagrama, por eso podemos suponer que  $B = *$ . Como  $E$  es un  $G$ -ANR espacio, según la propiedad extensorial existe una  $G$ -función  $\bar{f} : U \rightarrow E$ , donde  $U$  es una vecindad  $G$ -invariante de  $G \times_H A$  en  $G \times_H X$ , tal que  $\bar{f}|_{G \times_H A} = \hat{f}$ . Entonces  $V = i_X^{-1}(U)$  es una vecindad  $H$ -invariante de  $A$  en  $X$  y  $(\bar{f} \circ i_X)|_V : V \rightarrow E$  es una  $H$ -extensión de  $f$ . Concluimos que  $E$  es un  $H$ -ANR.

(b). Sea  $X = A \times I$  y  $s(a) = (a, 0)$ . Obviamente,  $G \times_H (A \times I)$  puede ser identificado con  $(G \times_H A) \times I$  (por medio de  $[g, (a, t)] \mapsto ([g, a], t)$  y  $\bar{s}([g, a]) = ([g, a], 0)$ ). Si  $p$  es una  $G$ -fibración, entonces existe un relleno  $\bar{F} : (G \times_H A) \times I \rightarrow E$ . En consecuencia, la  $H$ -función  $\tilde{F} : A \times I \rightarrow E$ ,  $\tilde{F} = \bar{F} \circ i_X$ , es también un relleno del diagrama, lo que significa que  $p$  es una  $H$ -fibración.  $\square$

4.3. OBSERVACIÓN. La condición de ser *metrizable* para el grupo compacto  $G$  en la Proposición 4.2 es necesaria a fin de tener  $G \times_H X$  como un  $G$ -espacio metrizable para un  $H$ -espacio  $X$  dado. Según la Proposición 3 de [2], se puede cambiar esta condición por una condición más débil: el espacio  $G/H$  debe ser metrizable.

Ahora bien, sea  $N$  un subgrupo normal de un grupo  $G$ . Si  $X$  es un  $G$ -espacio y  $Y$  es un  $G/N$ -espacio, entonces para cada  $G$ -función  $f : X \rightarrow Y$  existe una única  $G$ -función  $f' : X/N \rightarrow Y$  tal que el diagrama

$$\begin{array}{ccc} X & \xrightarrow{p_X} & X/N \\ & \searrow f & \swarrow f' \\ & & Y \end{array}$$

es conmutativo, donde la  $G$ -función  $p_X$  se define por  $p_X(x) = Nx$ . Claramente, la función  $f'$  se define por  $f'(Nx) = f(x)$ .

La existencia de la biyección natural

$$G\text{-}\mathcal{M}(X, Y) \longrightarrow G/N\text{-}\mathcal{M}(X/N, Y), \quad f \mapsto f'$$

significa que el functor de restricción  $\mathcal{F}_p : G/N\text{-}\mathcal{M} \rightarrow G\text{-}\mathcal{M}$  inducido por la proyección natural  $p : G \rightarrow G/N$  es un functor adjunto derecho al functor de  $N$ -órbitas  $-/N : G\text{-}\mathcal{M} \rightarrow G/N\text{-}\mathcal{M}$ .

De esta manera la demostración de la siguiente proposición es una repetición simple de la prueba de la Proposición 4.2.

4.4. PROPOSICIÓN. Sea  $N$  un subgrupo normal cerrado de un grupo compacto  $G$ .

- (a) Si  $E$  es un  $G/N$ -ANR-espacio, entonces  $E$  es un  $G$ -ANR espacio.
- (b) Si  $p : E \rightarrow B$  es una  $G/N$ -fibración, entonces  $p$  es una  $G$ -fibración.

5. ESPACIOS CON UN SOLO TIPO DE ÓRBITAS

En la sección 3 ya hemos probado un resultado importante, la Proposición 3.12, sobre  $G$ -espacios con un solo tipo de órbitas. La condición esencial en este resultado es que  $G$  debe ser un grupo compacto *de Lie*.

En esta sección (y en el resto del trabajo) consideramos un caso más general cuando  $G$  es un grupo compacto metrizable.

Antes del lema que será un fundamento de toda la disertación, necesitamos un resultado crucial de la teoría equivariante de retracts que concierne a los grupos compactos de Lie.

5.1. TEOREMA. (Palais,[9], Corolario 1.6.7) Sea  $H$  un subgrupo cerrado de un grupo compacto de Lie  $G$ . Entonces  $G/H$  es un  $G$ -ANR.

5.2. COROLARIO. Sea  $X$  un  $G$ -espacio tal que  $G$  es un grupo compacto de Lie. Entonces para cada  $x \in X$ , la órbita  $G(x)$  es un  $G$ -retracto de alguna vecindad abierta invariante en  $X$ .

DEMOSTRACIÓN. Es suficiente observar que  $G(x)$  es  $G$ -equivalente al  $G$ -espacio  $G/G_x$  y, por lo tanto,  $G(x)$  es  $G$ -ANR. □

5.3. OBSERVACIÓN. En realidad, el Teorema 5.1 tiene muchos corolarios relevantes. En particular, el Teorema de Rebanada (ver [7], Teorema 5.4) y el Teorema de Palais (véase Sección 3).

El siguiente concepto fue introducido por S.Antonyan en [1].

5.4. DEFINICIÓN. Un subgrupo cerrado  $H$  de un grupo compacto  $G$  se llama *grande* (en inglés: large subgroup) si existe un subgrupo normal cerrado  $N$  de  $G$  tal que  $N \subseteq H$  y  $G/N$  es un grupo de Lie.

5.5. PROPOSICIÓN. ([1]) Sea  $H$  un subgrupo cerrado de un grupo compacto  $G$ . Entonces las siguientes condiciones son equivalentes:

- (1)  $H$  es grande.
- (2)  $G/H$  es un  $G$ -ANR-espacio.
- (3)  $G/H$  es localmente contraíble.
- (4)  $G/H$  es una variedad suave.

5.6. OBSERVACIÓN. Sea  $H$  un subgrupo grande de un grupo compacto  $G$ . Si  $N(H)$  es el normalizador de  $H$  en  $G$ , entonces el grupo cociente  $N(H)/H$  es un grupo compacto de Lie.

En efecto, si  $N$  es un subgrupo normal cerrado de  $G$  tal que  $N \subseteq H$  y  $G/N$  es un grupo de Lie, entonces  $N(H)/N$  es un grupo de Lie por ser un subgrupo cerrado de  $G/N$ . La inclusión  $N \hookrightarrow H$  induce el epimorfismo continuo  $N(H)/N \rightarrow N(H)/H$ . De ello,  $N(H)/H$  es isomorfo a un grupo cociente del grupo de Lie  $N(H)/N$ , y por lo tanto es también un grupo de Lie (ya que los grupos cocientes de grupos de Lie son de Lie).

5.7. LEMA. Sea  $H$  un subgrupo grande de un grupo compacto  $G$ . Si un  $G$ -espacio  $X$  tiene un solo tipo de órbitas ( $H$ ), entonces existe una  $G$ -inclusión cerrada  $i : X \hookrightarrow U$  tal que  $U$  es  $G$ -ANR y tiene también un solo tipo de órbita ( $H$ ).

DEMOSTRACIÓN. Supongamos primero que  $G$  es un grupo de Lie y que  $X$  es un  $G$ -espacio libre. Sabemos que se puede considerar  $X$  como un subespacio cerrado

invariante de un  $G$ - $AR$  espacio  $M$ . En virtud del Corolario 5.2 cada órbita  $G(x)$  es  $G$ -retracto de una vecindad abierta invariante  $V_x$  de  $G(x)$  en  $M$ . Si  $r : V_x \rightarrow G(x)$  es una  $G$ -retracción, entonces  $G_y \subseteq G_{r(y)}$  entonces, para cada  $y \in V_x$ . Pero  $G_r(y) = \{e\}$  y por eso,  $G_y = \{e\}$ , es decir la acción de  $G$  en  $V_x$  es libre. Por lo tanto, la acción de  $G$  en  $V = \bigcup_{x \in X} V_x$  es libre y, obviamente,  $V$  es  $G$ - $ANR$  ya que es una vecindad abierta invariante de  $X$  en  $M$ .

Ahora pasamos al caso cuando  $X$  tiene un solo tipo de órbita ( $H$ ). Se tiene el  $G$ -equivalencia canónica  $X \approx G \times_N X^H$ , donde  $X^H$  es el conjunto de puntos fijos respecto a  $H$  y  $N = N(H)$  es el normalizador de  $H$  (ver, [8], Proposición 5.12). Notemos que el grupo cociente  $N/H$  actúa en  $X^H$  según la regla:  $nH \cdot x = nx$ . Además, obviamente, esta acción es libre. Como  $N/H$  es un grupo de Lie (véase la Observación 5.6), por lo anterior existe una  $N/H$ -inclusión  $X^H \hookrightarrow V$  en un  $N/H$ - $ANR$ -espacio libre  $V$  que también puede considerarse como  $N$ -inclusión del  $N$ -espacio  $X^H$  en  $N$ - $ANR$ -espacio  $V$  (véase la Proposición 4.4(a)). Esta inclusión induce la  $G$ -inclusión  $i : G \times_N X^H \hookrightarrow G \times_N V$ . Notemos que  $U = G \times_N X^H$  es un  $G$ - $ANR$ -espacio ([9], Corolario 1.7.6) y que  $U$  tiene un solo tipo de órbita ( $H$ ).  $\square$

El lema 5.7 representa uno de los dos resultados principales de esta sección, para probar el segundo resultado (el Lema 5.9) necesitamos la siguiente proposición general

5.8. PROPOSICIÓN. Sea  $Y$  un  $H$ -espacio, donde  $H$  es un subgrupo cerrado de un grupo  $G$ . Si la proyección orbital  $q_Y : Y \rightarrow Y/H$  es una  $H$ -fibración, entonces la función  $p : G \times_H Y \rightarrow Y/H$ , dada por  $p([g, y]) = q_Y(y)$ , es una  $G$ -fibración.

DEMOSTRACIÓN. La función  $p$  está bien definida pues  $[g, y] = [g', y']$  implica que  $y' = hy$  para algún  $h \in H$  y  $q_Y(y') = q_Y(hy) = q_Y(y)$ . Se tiene el siguiente diagrama conmutativo

$$\begin{array}{ccc} G \times Y & \xrightarrow{pr_Y} & Y \\ \downarrow [\ ] & & \downarrow q_Y \\ G \times_H Y & \xrightarrow{p} & Y/H \end{array}$$

donde  $[\ ]$  es la proyección orbital de la definición del producto torcido. Como la composición  $q_Y \circ pr_Y$  es continua, la función  $p$  también es continua por la propiedad de la topología cociente.

Ahora supongamos que el siguiente diagrama de  $G$ -funciones es conmutativo

$$\begin{array}{ccc} X & \xrightarrow{f} & G \times_H Y \\ \downarrow \partial_0 & & \downarrow p \\ X \times I & \xrightarrow{F} & Y/H \end{array}$$

Sea  $S = f^{-1}(i_Y(Y))$ , donde  $i_Y : Y \hookrightarrow G \times_H Y$  la  $H$ -inclusión cerrada dada por  $i_Y(y) = [e, y]$ . Ahora bien, sea  $f' : S \rightarrow Y$  la  $H$ -función definida por  $f'(s) = (i_Y)^{-1}(f(s))$ . Sin perdida de generalidad podemos asumir que  $X = G \times_H S$  en virtud de la  $G$ -equivalencia canónica  $\eta : G \times_H S \rightarrow X$ ,  $\eta([g, s]) = gs$  (véase [9], Corolario 1.7.8). En este caso,  $f = G \times_H f'$ , es decir  $f([g, s]) = [g, f'(s)]$  para cada  $[g, s] \in X$ .

Definimos  $F' : S \times I \rightarrow Y/H$  por  $F'(s, t) = F([e, s], t)$ . Claramente,  $F'$  es una  $H$ -función y  $F'(s, 0) = F([e, s], 0) = (p \circ f)([e, s]) = p([e, f'(s)]) = q_Y(f'(s))$  para cada  $s \in S$ . Como  $q_Y$  es una  $H$ -fibración, existe una  $H$ -homotopía  $\widehat{F} : S \times I \rightarrow Y$  tal que el diagrama

$$\begin{array}{ccc}
 S & \xrightarrow{f'} & Y \\
 \partial'_0 \downarrow & \nearrow \widehat{F} & \downarrow q_Y \\
 S \times I & \xrightarrow{F'} & Y/H
 \end{array}$$

es conmutativo.

Ahora podemos definir  $\widetilde{F} : X \times I \rightarrow G \times_H Y$  como sigue:

$$\widetilde{F}([g, s], t) = [g, \widehat{F}(s, t)], \quad [s, t] \in X, \quad g \in G.$$

Entonces  $p \circ \widetilde{F}([g, s], t) = p([g, \widehat{F}(s, t)]) = q_Y(\widehat{F}(s, t)) = F'(s, t) = F([e, s], t)$ . Puesto que la acción de  $G$  en  $Y/H$  es trivial, tenemos  $F([e, s], t) = gF([e, s], t) = F([g, s], t)$ . Así,  $p \circ \widetilde{F} = F$ . Además,  $\widetilde{F} \circ \partial_0([g, s]) = \widetilde{F}([g, s], 0) = [g, \widehat{F}(s, 0)] = [g, f'(s)] = f(s)$ , es decir  $\widetilde{F} \circ \partial_0 = f$ .  $\square$

5.9. LEMA. Sea  $H$  un subgrupo grande de un grupo compacto  $G$ . Si un  $G$ -espacio  $E$  tiene un solo tipo de órbitas ( $H$ ), entonces la proyección orbital  $q_E : E \rightarrow E/G$  es una  $G$ -fibración.

DEMOSTRACIÓN. Sabemos que  $E$  y  $G \times_N E^H$  son  $G$ -equivalentes, donde  $N = N(H)$  el normalizador del subgrupo  $H$  en  $G$ , por medio de la  $G$ -equivalencia canónica  $\eta : E \rightarrow G \times_{N(H)} E^H$ ,  $\eta([g, y]) = gy$ . Podemos considerar el  $N$ -espacio  $E^H$  como  $N/H$ -espacio y en este caso la proyección orbital  $E^H \rightarrow E^H/(N/H)$  es  $N/H$ -fibración ya que  $N/H$  es un grupo de Lie. Pero, en efecto,  $E/(N/H) = E/N$ , es decir  $q_{E^H} : E^H \rightarrow E^H/N$  es una  $N/H$ -fibración, lo que significa que  $q_{E^H}$  es una  $N$ -fibración. Ahora bien,  $q_E$  es una  $G$ -fibración en virtud de la Proposición 5.8.  $\square$

Sean  $H$  un subgrupo cerrado y  $N$  un subgrupo normal cerrado de un grupo compacto  $G$ . Ahora bien, sea  $\widehat{H} = \{hN \mid h \in H\} \subset G/N$ . Entonces  $\widehat{H}$  es un subgrupo cerrado de  $G/N$  por ser la imagen del subgrupo cerrado  $H$  bajo la proyección natural  $q_N : G \rightarrow G/N$  (que es un homomorfismo continuo cerrado). Notemos también que  $q_N^{-1}(\widehat{H}) = NH = HN$ . Obtenemos el siguiente diagrama conmutativo de  $G$ -equivalencias naturales

$$\begin{array}{ccc}
 (G/H)/N & \xrightarrow{\gamma} & (G/N)/\widehat{H} \\
 \searrow \alpha & & \nearrow \beta \\
 & G/(NH) &
 \end{array}$$

donde como de costumbre  $(G/H)/N$  es el espacio de  $N$ -órbitas del  $G$ -espacio  $G/H$  con la acción natural de  $N$  en  $G/H$  por las translaciones izquierdas. Obviamente, las  $G$ -equivalencias (que también pueden ser consideradas como  $G/N$ -equivalencias)  $\alpha$ ,  $\beta$  y  $\gamma$  se definen como sigue:  $\alpha(N(gH)) = gNH$ ,  $\beta(gNH) = (gN)\widehat{H}$  y  $\gamma(N(gH)) = (gN)\widehat{H}$ .

5.10. OBSERVACIÓN. Sea  $N$  un subgrupo normal cerrado de un grupo compacto  $G$  tal que  $G/N$  es un grupo de Lie. Entonces para cada subgrupo cerrado  $H$  de  $G$ , el subgrupo  $NH$  de  $G$  es grande.

En efecto, el subgrupo  $\widehat{H}$ , es grande en  $G/N$  y, por consiguiente,  $(G/N)/\widehat{H}$  es localmente contraíble según la Proposición 5.5. Como  $G/(NH)$  es  $G$ -equivalente a  $(G/N)/\widehat{H}$ , es también localmente contraíble, y el subgrupo  $NH$  es grande.

## 6. RESOLUCIONES DE LAS PROYECCIONES ORBITALES

El siguiente resultado es bien conocido (véase, por ejemplo, [10]).

6.1. TEOREMA. Si  $G$  es un grupo metrizable compacto, entonces existe una sucesión decreciente  $\{N_i\}_{i \in \mathbb{N}}$  de subgrupos normales cerrados de  $G$  tales que los grupos  $G/N_i$  son *grupos de Lie* con  $\bigcap_{i \in \mathbb{N}} N_i = \{e\}$  y, por consiguiente,

$$\varprojlim \{G/N_i, q_i^j\} = G,$$

donde  $q_i^j : G/N_j \rightarrow G/N_i$ ,  $j \geq i$ , son las proyecciones naturales.

6.2. OBSERVACIÓN. El Teorema 6.1 es, en realidad, una consecuencia fácil del siguiente resultado (probado en [7]): si  $G$  es grupo compacto de Hausdorff, entonces cada vecindad del elemento neutro  $e \in G$  contiene un subgrupo grande.

El teorema 6.1 tiene el siguiente corolario ([3], Corolario 9, véase también la prueba en [6]).

6.3. COROLARIO. Sea  $G$  un grupo metrizable compacto y sea  $\{N_i\}_{i \in \mathbb{N}}$  una sucesión decreciente de subgrupos normales cerrados en  $G$  que satisface los requerimientos del Teorema 6.1.

Si  $X$  es un  $G$ -espacio entonces

$$X = \varprojlim \{X/N_i, p_i^j\}$$

donde  $p_i^j : X/N_j \rightarrow X/N_i$ ,  $j \geq i$ , son las proyecciones naturales.

6.4. OBSERVACIÓN. En el Corolario 6.3 los espacios  $X/N_i$  son  $G/N_i$ -espacios, pero también podemos tratarlos como  $G$ -espacios (véase la Sección 4). Asimismo, consideramos las proyecciones  $p_i^j$  como  $G$ -funciones.

6.5. DEFINICIÓN. Sea  $X$  un  $G$ -espacio compacto. Una sucesión inversa de  $G$ -ANR espacios y  $G$ -funciones  $\{X_i, q_i^j\}$  se llama  *$G$ -ANR-resolución de  $X$*  si

- (1)  $X = \varprojlim \{X_i, q_i^j\}$
- (2) la familia de proyecciones naturales  $\{q_i : X \rightarrow X_i\}$  satisface la siguiente condición: para cada  $i$  y para cada vecindad abierta e invariante  $U$  de  $q_i(X)$  en  $X_i$  existe  $j \geq i$  tal que  $q_i^j(X_j) \subseteq U$ .

6.6. OBSERVACIÓN. Sea  $X$  un subconjunto compacto e invariante de un  $G$ -ANR-espacio  $M$  y sea  $\{U_i\}_{i \in \mathbb{N}}$  una colección de vecindades abiertas invariantes de  $X$  en  $M$ , tales que  $U_{i+1} \subset U_i$  para cada  $i$  y  $\bigcap_{i \in \mathbb{N}} U_i = X$ . Entonces la sucesión inversa  $\{U_i, u_i^j\}$ , donde  $u_i^j$  son las inclusiones  $U_j \hookrightarrow U_i$ , es una  $G$ -ANR-resolución si y sólo si  $\{U_i\}_{i \in \mathbb{N}}$  es una base de vecindades de  $X$  en  $M$ .

Por consiguiente, cada  $G$ -espacio compacto admite una  $G$ -ANR-resolución.

Esta observación nos permite redactar el Lema 5.7 como sigue:

6.7. PROPOSICIÓN. Sea  $H$  en subgrupo grande de un grupo compacto  $G$ . Si un  $G$ -espacio  $X$  tiene un solo tipo de órbitas  $(H)$ , entonces existe una  $G$ -ANR-resolución  $\{X_i, u_i^j\}$  de  $X$  tal que cada  $G$ -ANR-espacio  $X_i$  tiene un solo tipo de órbitas  $(H)$  y cada  $u_i^j$  es una  $G$ -inclusión  $X_j \hookrightarrow X_i$ .

6.8. TEOREMA. Sea  $H$  un subgrupo cerrado de un grupo metrizable compacto  $G$ . Si un  $G$ -espacio compacto  $E$  tiene un solo tipo de órbita  $(H)$ , entonces existe una  $G$ -ANR-resolución  $\{E_i, q_i^j\}$  tal que  $E_i$  es tiene un solo tipo de órbita  $(HN_i)$ , donde  $\{N_i\}_{i \in \mathbb{N}}$  es una sucesión decreciente de subgrupos normales cerrados de  $G$  que satisface los requerimientos del Teorema 6.1 (en particular,  $HN_i$  es un subgrupo grande de  $G$ ).

DEMOSTRACIÓN. Según el Corolario 6.3, representamos  $E$  como

$$E = \varprojlim \{E/N_i, p_i^j\}$$

donde  $p_i^j : E/N_j \rightarrow E/N_i$ ,  $j \geq i$ , son las proyecciones naturales. Como  $E/N_i$  es  $G/N_i$ -espacio y  $G/N_i$  es un grupo compacto de Lie, por el lema anterior, existe  $G/N_i$  inclusión cerrada  $u_i : E/N_i \hookrightarrow U_i$  en  $G/N_i$ -ANR espacio  $U_i$  con un solo tipo de órbita. Notemos que este tipo de órbita es  $(HN_i)$ .

En virtud del Lema 5.7, para cada  $i$ , consideramos  $E/N_i$  como un subconjunto cerrado de un  $G/N_i$ -ANR espacio  $V_i$  que tiene el mismo tipo de órbitas  $(HN_i)$ . Hallamos por inducción las vecindades abiertas  $U_i$  de  $E/N_i$  en  $V_i$  como sigue: ponemos  $U_1 = V_1$  y si  $U_i$  ya está dada, usando la propiedad *extensorial* de  $G$ -ANR espacio  $U_i$ , encontramos  $U_{i+1}$  como una vecindad invariante abierta de  $E/N_{i+1}$  en  $V_{i+1}$  en que existe una extensión equivariante  $f_i^{i+1} : U_{i+1} \rightarrow U_i$  de  $E/N_{i+1} \rightarrow E/N_i \hookrightarrow U_i$ . Sea  $\{U_i^{(j)}\}_{j \in \mathbb{N}}$  un base de vecindades invariantes abiertas y decrecientes (es decir con  $U_i^{(j+1)} \subset U_i^{(j)}$ ) de  $E/N_i$  es  $U_i$ . Sin pérdida de generalidad podemos asumir que  $(f_i^{i+1})^{-1}(U_i^{(j)}) \supset U_{i+1}^{(j)}$ . Ahora, por fin, ponemos  $E_i = U_i^{(i)}$  para cada  $i$ , y definimos  $q_i^{i+1}$  como la restricción de  $f_i^{i+1}$  en  $E_{i+1}$  con el codominio  $E_i$  (tomando en cuenta que  $f_i^{i+1}(U_{i+1}^{(i+1)}) \subset U_i^{(i+1)} \subset U_i^{(i)}$ ).

Veamos que la sucesión inversa obtenida  $\{E_i, q_i^j\}$  es una  $G$ -ANR-resolución de  $E$ . Primero, cada  $E_i$  es  $G$ -ANR-espacio como un subconjunto invariante y abierto del  $G$ -ANR  $U_i$ . Para ver que  $E = \varprojlim \{E_i, q_i^j\}$  es suficiente observar lo siguiente: si  $\{e_i\}_{i \in \mathbb{N}}$  es una sucesión de  $e_i \in E_i$  tal que  $q_i^{i+1}(e_{i+1}) = e_i$ , entonces inevitablemente  $e_i \in E/N_i$  para cada  $i$ . En efecto, si  $e_i \notin E/N_i$  entonces existe  $j > i$  tal que  $e_i \notin U_i^{(j)}$ , pero por otro lado  $f_{j-1}^j(e_j) \in U_{j-1}^{(j)}$ ,  $f_{j-2}^{j-1} f_{j-1}^j(e_j) \in U_{j-2}^{(j)}$ , etc., y de ello  $q_i^j(e_j) \in U_i^{(j)}$ , es decir  $e_i \in U_i^{(j)}$ . Contradicción. Por último, sea  $U$  es una vecindad de  $q_i(E)$  en  $E_i$ . Como  $\{U_i^{(j)}\}_{j \in \mathbb{N}}$  es una base de vecindades, existe  $j > i$  tal que  $U_i^{(j)} \subset U$  y por lo tanto  $q_i^j(E_j) = q_i^j(U_j^{(j)}) \subset U_i^{(j)} \subset U$ .  $\square$

6.9. DEFINICIÓN. Sea  $f : X \rightarrow Y$  una  $G$ -función de  $G$ -espacios compactos. Una sucesión inversa  $\{f_i, \beta_i^j\}$  que consiste de  $G$ -funciones  $f_i$  y pares  $\beta_i^j = (q_i^j, r_i^j)$  de

$G$ -funciones tales que el diagrama

$$\begin{array}{ccc} X_i & \xleftarrow{q_i^{i+1}} & X_{i+1} \\ f_i \downarrow & & \downarrow f_{i+1} \\ Y_i & \xleftarrow{r_i^{i+1}} & Y_{i+1} \end{array}$$

es conmutativo para cada  $i$ , se llama  $G$ -ANR-resolución de  $f$  si

- (1)  $\{X_i, q_i^j\}$  es  $G$ -ANR-resolución de  $X$ ,
- (2)  $\{Y_i, r_i^j\}$  es  $G$ -ANR-resolución de  $Y$ ,
- (3)  $f = \varprojlim \{f_i, \beta_i^j\}$ .

El siguiente resultado es de interés general y, por lo tanto, daremos su demostración con todos los detalles.

6.10. PROPOSICIÓN. Sea  $\{Y_i, r_i^j\}$  una  $G$ -ANR-resolución de un  $G$ -espacio compacto  $X$ . Entonces

- (i)  $\{Y_i/G, r_i^j/G\}$  una ANR-resolución de  $Y/G$ .
- (ii)  $\{q_{Y_i}, \beta_i^j\}$ , donde  $\beta_i^j = (r_i^j, r_i^j/G)$ , es una  $G$ -ANR-resolución de la proyección orbital  $q_Y : Y \rightarrow Y/G$ .

DEMOSTRACIÓN. Debemos probar que  $Y/G = \varprojlim \{Y_i/G, r_i^j/G\}$ . Para hacerlo, hay que demostrar que la colección  $\{r_i/G : Y/G \rightarrow Y_i/G\}_{i \in \mathbb{N}}$  satisface la propiedad universal del límite inverso.

Sea  $\{f_i : Z \rightarrow Y_i/G\}_{i \in \mathbb{N}}$  una colección de funciones continuas tales que para cada  $i$ ,  $f_i = r_i^{i+1}/G \circ f_{i+1}$ . Consideramos el pull-back diagrama

$$\begin{array}{ccc} X_i & \xrightarrow{f'_i} & Y_i \\ q_{X_i} \downarrow & & \downarrow q_{Y_i} \\ Z & \xrightarrow{f_i} & Y_i/G \end{array}$$

para cada  $i$ . Según la propiedad universal de pull-back, para cada  $i$ , existe una única  $G$ -función  $h_i^{i+1} : Y_{i+1} \rightarrow Y_i$  tal que  $q_{Y_i} \circ h_i^{i+1} = q_{Y_{i+1}}$  y  $f'_i \circ h_i^{i+1} = q_i^{i+1} \circ f'_{i+1}$ . En otras palabras los siguientes diagramas son conmutativos

$$\begin{array}{ccc} X_i & \xleftarrow{h_i^{i+1}} & X_{i+1} \\ q_{X_i} \searrow & & \swarrow q_{X_{i+1}} \\ & Z & \end{array} \qquad \begin{array}{ccc} X_i & \xleftarrow{h_i^{i+1}} & X_{i+1} \\ f'_i \downarrow & & \downarrow f'_{i+1} \\ Y_i & \xleftarrow{r_i^{i+1}} & Y_{i+1} \end{array}$$

Sea  $X = \varprojlim \{X_i, h_i^{i+1}\}$  con la proyecciones naturales  $h_i : X \rightarrow X_i$ , y sea  $p = \varprojlim \{q_{X_i}\}$ ,  $p : X \rightarrow Z$ . Entonces, en particular,  $p = q_{X_i} \circ h_i$  para cada  $i$ . Veamos que se puede considerar  $p$ , en efecto, como proyección orbital de  $X$ .

La función  $p$  es sobreyectiva ya que las funciones  $q_{X_i}$  y  $h_i^{i+1}$  son sobreyectivas (notemos que  $h_i^{i+1}$  manda cada órbita  $q_{X_{i+1}}^{-1}(z)$  sobre  $q_{X_i}^{-1}(z)$ ). Además  $p$  es

abierta. Para verlo recordemos que la base de la topología de  $X$  está formada por los conjuntos de la forma  $h_i^{-1}(V)$  donde  $V$  es abierto en  $X_i$  y, por eso, los conjuntos  $p(h_i^{-1}(V)) = q_{X_i}(V)$  son abiertos pues las proyecciones orbitales  $q_{X_i}$  son abiertas.

Finalmente, mostraremos que para cada  $z \in Z$  el conjunto  $p^{-1}(z)$  es una  $G$ -órbita en  $X$ . Sean  $x, x' \in p^{-1}(z)$ . Para cada  $i$ , existe  $g_i \in G$  tal que  $h_i(x') = g_i h_i(x)$  ya que  $q_{X_i}^{-1}(z)$  es una  $G$ -órbita. Como  $G$  es compacto, la sucesión  $(g_i)$  contiene una subsección que converge a algún elemento  $g \in G$ , sin pérdida de generalidad podemos asumir que la misma sucesión  $(g_i)$  converge a  $g$ . Afirmamos que  $x' = gx$ . Si suponemos que  $x' \neq gx$ , entonces para algún  $k$ ,  $h_k(x') \neq h_k(gx) = gh_k(x)$  y podemos encontrar vecindades  $U$  y  $V$  de  $h_k(x')$  y  $gh_k(x)$  respectivamente tales que  $U \cap V = \emptyset$ . Existe  $j_0$  tal que  $g_j h_k(x) \in V$  para cada  $j > j_0 > k$ , pero por otro lado,  $h_k^j(g_j h_j(x)) = h_k^j(h_j(x'))$ , es decir  $g_j h_k(x) = h_k(x')$ . Así,  $g_j h_k(x) \in U$ . Contradicción.

La colección  $\{f'_i : X_i \rightarrow Y_i\}_{i \in \mathbb{N}}$  induce la  $G$ -función  $f' : X \rightarrow Y$ , en otras palabras  $f' = \varprojlim \{f'_i\}$ . La función continua  $f = f'/G : Z \rightarrow Y/G$  inducida por  $f'$  satisface  $r_i/G \circ f = f_i$  para cada  $i$ .

Supongamos ahora que alguna otra función  $u : Z \rightarrow X/G$  satisface la misma condición:  $r_i/G \circ u = f_i$  para cada  $i$ . Sea  $z \in Z$ . Las órbitas  $\alpha = f(z)$  y  $\alpha' = u(z)$  satisfacen  $r_i(\alpha) = r_i(\alpha')$  para cada  $i$ . Si  $y, y' \in X$  son tales que  $q_Y(x) = \alpha$  y  $q_Y(y') = \alpha'$  entonces  $q_{Y_i}(r_i(y)) = q_{Y_i}(r_i(y'))$  para cada  $i$  ya que  $q_{Y_i} \circ r_i = r_i/G \circ q_Y$ . Por consiguiente, existe una sucesión  $(g_i, g_i]inG$ , tal que  $r_i(y') = g_i r_i(y)$  para cada  $i$ . Repitiendo el argumento anterior, que existe  $g \in G$ , el punto de acumulación de  $(g_i)$ , tal que  $y' = gy$ . Concluimos que  $f(z) = q_Y(y) = q_Y(y') = u(z)$  lo que significa que  $u = f$ .

Hemos probado que  $Y/G = \varprojlim \{Y_i/G, r_i^j/G\}$ .

Cada  $Y_i/G$  es ANR según el Teorema 2.2. Para terminar la demostración hace falta probar que  $\{Y_i/G, r_i^j/G\}$  satisface la segunda condición de la definición de  $G$ -ANR-resolución. Sea  $U$  una vecindad de  $(r_i/G)(Y/G)$  en  $Y_i/G$ . Como  $\{Y_i, r_i^j\}$  una  $G$ -ANR-resolución y  $r_i(Y) \subset q_{Y_i}^{-1}(U)$ , existe  $j > i$  tal que  $r_i^j(Y_j) \subset q_{Y_i}^{-1}(U)$ . Entonces  $(r_i^j/G)(Y_j/G) \subset U$ . □

El resultado principal del presente artículo es aparentemente una consecuencia sencilla del Teorema 6.8 y de la Proposición 6.10.

**6.11. TEOREMA.** Sea  $H$  un subgrupo cerrado de un grupo compacto metrizable  $G$ . Si un  $G$ -espacio compacto  $E$  tiene un solo tipo de órbita ( $H$ ), entonces la proyección orbital  $q_E : E \rightarrow E/G$  admite una  $G$ -ANR-resolución  $\{p_i, \beta_i^{i+1}\}$  tal que  $q_i$  es una  $G$ -fibración para cada  $i$ .

Más precisamente, existe una  $G$ -ANR-resolución  $\{p_i, \beta_i^{i+1}\}$  de  $q$  que consta de las proyecciones orbitales  $p_i : E_i \rightarrow E_i/G$ , donde  $E_i$  tiene un solo tipo de órbitas ( $HN_i$ ) y  $\{N_i\}_{i \in \mathbb{N}}$  es una sucesión decreciente de subgrupos normales cerrados de  $G$  que satisface los requerimientos del Teorema 6.1 (en particular,  $HN_i$  es un subgrupo grande de  $G$ ).

**DEMOSTRACIÓN.** Es suficiente aplicar el funtor  $-/G$  de la proyección orbital a la  $G$ -ANR-resolución  $\{E_i, q_i^j\}$  dada según el Teorema 6.8. De este modo, en virtud de la Proposición 6.10, obtenemos una  $G$ -ANR-resolución  $\{p_i, \beta_i^{i+1}\}$  de  $q_E$  en que  $p_i = q_{E_i}$  para cada  $i$ . Por el Lema 5.9, cada  $p_i$  es una  $G$ -fibración. □

## REFERENCIAS

- [1] S.A. Antonyan, *Existence of a slice for arbitrary compact transformation group*, Mat. Zametki 56 (1994), no.5, 3-9.
- [2] S.A. Antonyan, *Extensorial properties of orbit spaces of proper group actions*, Topology and Applications, 98 (1999) 35-46.
- [3] S.A. Antonyan, *Equivariant embeddings into  $G$ -AR's*, Glasnik Mat. 22(42) (1987) 503-533.
- [4] S.A. Antonyan, S.Mardešić, *Equivariant shape*, Fund.Math. 127 (1987) 213-224.
- [5] A. Bykov, M. Taxis, *Equivariant strong shape*, Topology and its Applications, 154 (2007), 2026-2039.
- [6] A. Bykov, M. Taxis, *Espacios fibrantes y grupos compactos en: Topología y Sistemas Dinámicos I*, BUAP, 2007, 47-60.
- [7] G.E.Bredon, *Introduction to compact transformation groups*, Academic Press, New York, 1972.
- [8] S. de Neymet, *Introducción a los grupos topológicos de transformaciones*, Aport. Mat., Serie: Textos, SMM, 2005.
- [9] R.S.Palais, *The classification of  $G$ -spaces*, Memoirs AMS, 36, 1960.
- [10] L.S. Pontrjagin, *Topological groups*, Princeton Univ. Press, 1939.
- [11] T. tom Dieck, *Transformation groups*, Walter de Gruyter, Berlin-New York, 1987.

Facultad de Ciencias Físico Matemáticas, BUAP.  
Av. San Claudio y 18 Sur, Col. San Manuel, Ciudad Universitaria  
Puebla, Pue., C.P. 72570  
abykov@fcfm.buap.mx

# COCIENTE DE PRODUCTOS SIMÉTRICOS DE UN CONTINUO

ENRIQUE CASTAÑEDA ALVARADO  
JAVIER SÁNCHEZ MARTÍNEZ  
UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO

RESUMEN. Sean  $X$  un continuo y  $n \in \mathbb{N}$ ,  $F_n(X)$  es el  $n$ -ésimo producto simétrico de  $X$ . Para  $n, m \in \mathbb{N}$  con  $m < n$ ,  $F_n(X)/F_m(X)$  denota al espacio cociente que resulta al comprimir  $F_m(X)$  a un sólo punto en  $F_n(X)$ . En éste trabajo mostraremos algunas propiedades básicas de  $F_n(X)/F_m(X)$ , así mismo daremos algunos modelos geométricos para  $F_2(X)/F_1(X)$  de algunos continuos.

## 1. INTRODUCCIÓN

Un *continuo*  $X$  es un espacio métrico, compacto, conexo y no degenerado. Dados un continuo  $X$  y  $n \in \mathbb{N}$ , denotamos por  $F_n(X)$  al espacio que consta de todos los subconjuntos no vacíos de  $X$  que tienen a lo más  $n$  puntos, con la topología de la métrica de Hausdorff (véase Proposición 2.1, [6]). Este espacio es también llamado el  *$n$ -ésimo producto simétrico de  $X$* . Se sabe que  $F_2(S^1)$  es homeomorfo a una banda de Möebius (véase pág. 877, [1]) y  $F_3(S^1)$  es homeomorfo a  $S^3$  (véase [2]), donde  $S^n$  denota la esfera unitaria de dimensión  $n$ . Los productos simétricos fueron introducidos por K. Borsuk y S. Ulam en [1]. Ellos probaron que, para  $I = [0, 1]$  y  $n = 1, 2, 3$ ,  $F_n(I)$  es homeomorfo a  $I^n$  y para  $n \geq 4$ ,  $F_n(I)$  no es homeomorfo a ningún subconjunto de  $\mathbb{R}^n$ . En [7] R. Molski prueba que  $F_2(I^2)$  es homeomorfo a la 4-celda y para  $n \geq 3$ , tanto  $F_n(I^2)$  como  $F_2(I^n)$  no son homeomorfos a algún subconjunto de  $\mathbb{R}^{2n}$ . Y en [4] E. Castañeda muestra que si  $X$  es un  $n$ -odo simple,  $F_2(X)$  es el cono sobre un continuo  $Z$  definido como la unión de una gráfica completa  $K_n$  con  $n$  arcos ajenos por pares que intersectan a  $K_n$  en exactamente uno de sus vértices.

Para  $m < n$ ,  $F_n(X)/F_m(X)$  denota al espacio cociente que resulta al identificar  $F_m(X)$  a un sólo punto en  $F_n(X)$  con la función proyección  $\rho_X^m$  (véase pág. 41, [9]), donde  $\rho_X^m(A) = \{A\}$  si  $A \notin F_m(X)$  y  $\rho_X^m(A) = F_m(X)$  en otro caso. En este artículo mostraremos algunas propiedades del espacio cociente  $F_n(X)/F_m(X)$ , así mismo mostraremos que si  $X$  es un arco o un  $n$ -odo simple entonces  $F_2(X)/F_1(X)$  es homeomorfo a  $F_2(X)$  y que si  $X$  es una curva cerrada simple  $F_2(X)/F_1(X)$  es homeomorfo al plano proyectivo real.

## 2. DEFINICIONES

Una *gráfica finita*  $G$  es un continuo que puede ser escrito como la unión de una cantidad finita de arcos, donde cada par de ellos son ajenos o se intersectan en uno o en ambos puntos terminales. A los arcos, cuya unión es  $G$ , los llamaremos *aristas de  $G$* . Los puntos terminales de dichos arcos serán llamados *vértices de  $G$* . El orden,  $o(v)$ , de un vértice  $v$  en  $G$  es definido como el número de aristas de la gráfica  $G$

que contienen a  $v$ .

Un *punto final de  $G$*  es un vértice de orden 1 y un *punto de ramificación de  $G$*  es un vértice de orden mayor o igual a 3.

Un  *$n$ -odo simple con vértice  $v$*  es la gráfica que consta de la unión de  $n$  arcos  $I_1, I_2, \dots, I_n$  tales que  $I_i \cap I_j = v$  si  $i \neq j$  y  $v$  es punto terminal de cada  $I_i$ . Un 3-odo simple es llamado únicamente *triodo simple*. Supongamos que  $\{e_1, e_2, \dots, e_n\}$  es la base canónica de  $\mathbb{R}^n$ , como espacio vectorial. Sea  $\theta$  el origen de coordenadas de  $\mathbb{R}^n$ . Dados los puntos  $p, q \in \mathbb{R}^n$ , el segmento convexo que une a  $p$  con  $q$  será denotado por  $pq$ . Entonces podemos considerar al  $n$ -odo simple,  $T_n$ , como la unión de todos los segmentos convexos  $\theta e_i$ , con  $i \in \{1, 2, \dots, n\}$ .

Dados  $X$  un continuo y  $n, m \in \mathbb{N}$ , con  $m < n$ , definimos el conjunto  $\mathfrak{D}(F_m(X))$  de la siguiente manera:

$$\mathfrak{D}(F_m(X)) = \{F_m(X)\} \cup \{\{A\} \subset F_n(X) : m < |A| \leq n\}$$

Consideremos la función  $\rho_X^m : F_n(X) \rightarrow \mathfrak{D}(F_m(X))$ , donde

$$\rho_X^m(A) = \begin{cases} F_m(X) & \text{si } |A| \leq m \\ \{A\} & \text{si } m < |A| \leq n \end{cases}$$

2.1. DEFINICIÓN. Si  $X$  es un continuo y  $m < n$ , se define el *espacio cociente*  $F_n(X)/F_m(X)$  como el conjunto  $\mathfrak{D}(F_m(X))$  dotado con la topología

$$\tau_{\mathfrak{D}} = \left\{ A \subset \mathfrak{D}(F_m(X)) : (\rho_X^m)^{-1}(A) \text{ es abierto en } F_n(X) \right\}.$$

También se usa la notación  $\rho_X^m(A) = [A]_{F_m(X)}$ ,  $\rho_X = \rho_X^1$  y  $[A] = [A]_{F_1(X)}$ .

Por otra parte, consideremos a  $D_1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ , el disco unitario en  $\mathbb{R}^2$ , y a la partición  $\mathfrak{C}$  de  $D_1$  dada como sigue:

$$\mathfrak{C} = \{\{(x, y)\} : x^2 + y^2 < 1\} \cup \{\{(x, y), (-x, -y)\} : x^2 + y^2 = 1\}$$

Sea  $h : D_1 \rightarrow \mathfrak{C}$  la función que asigna a cada punto en  $D_1$  el único elemento en la partición  $\mathfrak{C}$  que lo contiene. Se define así  $\mathbb{P}^2(\mathbb{R})$ , el *plano proyectivo real*, como el conjunto  $\mathfrak{C}$  con la topología  $\tau_{\mathfrak{C}} = \{A \subset \mathfrak{C} : h^{-1}(A) \text{ es abierto en } D_1\}$

### 3. ALGUNAS PROPIEDADES DEL ESPACIO COCIENTE $F_n(X)/F_m(X)$

En esta sección supondremos que  $X$  denota un continuo con métrica  $d_X$ , que  $n$  y  $m$  son dos números naturales con la propiedad de que  $m < n$  y que  $H$  denota la métrica de Hausdorff en  $F_n(X)$ .

Denotaremos por  $B_H^{F_n(X)}(\epsilon, A)$  a la bola abierta en  $F_n(X)$  con la métrica de Hausdorff, de radio  $\epsilon$  y con centro en  $A$ .

Resulta claro que  $\rho_X^m$  es una función continua del continuo  $F_n(X)$  en el espacio  $F_n(X)/F_m(X)$ , por lo que  $F_n(X)/F_m(X)$  es un espacio compacto y conexo. Luego, como  $F_n(X)$  es métrico y compacto, y  $F_n(X)/F_m(X)$  es espacio de Hausdorff, resulta que  $F_n(X)/F_m(X)$  es metrizable (véase págs. 217, [8]). De esto se tiene el siguiente resultado.

3.1. TEOREMA. Si  $X$  es un continuo, el espacio cociente  $F_n(X)/F_m(X)$  es un continuo.

Otra propiedad de  $F_n(X)/F_m(X)$  es la siguiente.

3.2. TEOREMA. Si  $X$  es un continuo, entonces  $F_n(X)/F_m(X)$  es la compactación por un punto de  $F_n(X) \setminus F_m(X)$ .

DEMOSTRACIÓN. Notemos que  $H = (F_n(X)/F_m(X)) \setminus \{F_m(X)\}$  es un conjunto abierto y denso en  $F_n(X)/F_m(X)$  tal que  $H \cup \{F_m(X)\}$  es compacto y que  $H$  es homeomorfo a  $F_n(X) \setminus F_m(X)$ . Por lo tanto  $F_n(X)/F_m(X)$  es la compactación unipuntual de  $F_n(X) \setminus F_m(X)$ .  $\square$

Como  $\rho_X^m$  es una función continua entre continuos, entonces  $\rho_X^m$  es una función cerrada. Veamos que  $\rho_X^m$  nunca es abierta.

3.3. LEMA. Si  $U$  es un abierto en  $F_n(X)$  tal que  $F_m(X) \cap U \neq \emptyset$  y  $F_m(X) \not\subseteq \bar{U}$ , entonces  $\rho_X^m(U)$  no es abierto en  $F_n(X)/F_m(X)$ .

DEMOSTRACIÓN. Supongamos lo contrario, es decir, que  $\rho_X^m(U)$  es un conjunto abierto en  $F_n(X)/F_m(X)$ . Entonces  $(\rho_X^m)^{-1}(\rho_X^m(U))$  es un conjunto abierto en  $F_n(X)$ , como  $F_m(X) \cap U \neq \emptyset$ , se tiene que

$$(\rho_X^m)^{-1}(\rho_X^m(U)) = U \cup F_m(X).$$

Sea  $A \in F_m(X)$  tal que  $A \not\subseteq \bar{U}$ , el cual existe pues  $F_m(X) \not\subseteq \bar{U}$ . Así que existe  $\epsilon > 0$  tal que  $B_H^{F_n(X)}(\epsilon, A) \cap U = \emptyset$ , y como  $F_m(X) \cup U$  es un conjunto abierto, entonces existe  $\delta > 0$  tal que  $B_H^{F_n(X)}(\delta, A) \subset F_m(X) \cup U$ . Al considerar  $\eta = \min\{\epsilon, \delta\} > 0$ , se cumple que

$$B_H^{F_n(X)}(\eta, A) \subset F_m(X)$$

lo cual es imposible.

Por lo tanto, podemos concluir que  $\rho_X^m(U)$  no es abierto en  $F_n(X)/F_m(X)$ .  $\square$

3.4. LEMA. Si  $X$  es un continuo entonces existe un abierto  $U$  en  $F_n(X)$  tal que  $U \cap F_m(X) \neq \emptyset$  y  $F_m(X) \not\subseteq \bar{U}$ .

DEMOSTRACIÓN. Como  $X$  es un continuo, entonces existen  $x, y \in X$  con  $x \neq y$ , así que  $d_X(x, y) > 0$ , sea  $\epsilon = \frac{d_X(x, y)}{4}$ .

Consideremos

$$U = B_H^{F_n(X)}(\epsilon, \{x\}).$$

Entonces  $U$  es un abierto en  $F_n(X)$  y  $\{x\} \in U \cap F_m(X)$ , por lo que  $U \cap F_m(X) \neq \emptyset$ .  $\{y\} \notin \bar{U}$  debido a que  $B_H^{F_n(X)}(\epsilon, \{y\}) \cap U = \emptyset$ , pues de lo contrario, existiría

$$C \in B_H^{F_n(X)}(\epsilon, \{x\}) \cap B_H^{F_n(X)}(\epsilon, \{y\}),$$

así que  $H(\{x\}, C) < \epsilon$  y  $H(\{y\}, C) < \epsilon$ , por la desigualdad del triángulo tendríamos que

$$H(\{x\}, \{y\}) \leq 2\epsilon < d_X(x, y),$$

pero  $H(\{x\}, \{y\}) = d_X(x, y)$  y esto sería una contradicción. Por lo tanto  $\{y\} \in F_m(X) \setminus \bar{U}$ , lo cual concluye la demostración.  $\square$

De los dos lemas anteriores se obtiene el siguiente resultado.

3.5. COROLARIO. Si  $X$  es un continuo. Entonces la función cociente  $\rho_X^m : F_n(X) \rightarrow F_n(X)/F_m(X)$  no es abierta.

#### 4. EL CASO PARTICULAR $F_2(X)/F_1(X)$

En esta sección construiremos homeomorfismos del espacio cociente  $F_2(X)/F_1(X)$  en espacios ya conocidos, cuando  $X$  es un arco, un  $n$ -odo ó una curva cerrada simple.

4.1. TEOREMA. Si  $X = [0, 1]$ , entonces  $F_2(X)/F_1(X)$  es homeomorfo a  $F_2(X)$ .

DEMOSTRACIÓN. En [1] se muestra que la función  $f : F_2(X) \rightarrow \mathbb{R}^2$ , dada por

$$f(\{a, b\}) = \left(\frac{a+b}{2}, |b-a|\right),$$

es un homeomorfismo sobre su imagen.  $\square$

Definamos la función  $h : F_2(X)/F_1(X) \rightarrow \mathbb{R}^2$  de la siguiente manera

$$h(\{a, b\}) = \left(|b-a| \frac{a+b}{2}, |b-a|\right).$$

Notemos que  $h$  está bien definida pues  $h(F_1(X)) = (0, 0)$ .

Mostraremos que  $f$  es inyectiva. Supongamos que  $h(\{a, b\}) = h(\{c, d\})$ . Si  $a = b$ , entonces  $c = d$ , pues  $|c-d| = 0$ , en cuyo caso  $\{a, b\} = F_1(X) = \{c, d\}$ , lo mismo ocurre si  $c = d$ . Así que supongamos que  $a \neq b$  y  $c \neq d$ .

Como  $h(\{a, b\}) = h(\{c, d\})$ , tenemos el siguiente sistema de ecuaciones:

$$\begin{aligned} |b-a| \frac{a+b}{2} &= |c-d| \frac{c+d}{2}, \\ |b-a| &= |c-d|. \end{aligned}$$

Como  $|b-a| \neq 0$ , entonces  $\frac{a+b}{2} = \frac{c+d}{2}$ , y por tanto  $a+b = c+d$ . Sin pérdida de generalidad, podemos suponer que  $a < b$  y  $c < d$ , entonces, como  $|b-a| = |c-d|$ , también se tiene que  $b-a = d-c$ . De donde  $b = d$  y  $a = c$ . Por lo tanto  $\{a, b\} = \{c, d\}$ . Con lo que concluimos que  $h$  es inyectiva.

Consideremos  $h_1 = g \circ f : F_2(X) \rightarrow \mathbb{R}^2$ , donde  $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  es la función dada por  $g(x, y) = (xy, y)$ , es claro que la función  $h_1$  es continua. Además  $h_1$  es una función que es constante en cada fibra de la función cociente  $\rho_X$ . Por Teorema 22.2, [5], la función  $H : F_2(X)/F_1(X) \rightarrow \mathbb{R}^2$ , inducida por  $h_1$ , dada por

$$H(\{a, b\}) = \left(|b-a| \frac{a+b}{2}, |b-a|\right),$$

es una función continua. Como  $H = h$ , entonces  $h$  es una función continua e inyectiva, que sale del continuo  $F_2(X)/F_1(X)$  en el espacio de Hausdorff  $\mathbb{R}^2$ , así que  $h$  es un homeomorfismo entre  $F_2(X)/F_1(X)$  y  $h(F_2(X)/F_1(X))$ . Como  $h(F_2(X)/F_1(X))$  es la región en plano  $\mathbb{R}^2$  acotada por las gráficas de las funciones reales definidas por  $f_1(z) = \frac{z^2}{2}$  y  $f_2(z) = z \frac{1-z}{2}$ , concluimos que  $F_2(X)/F_1(X)$  es una 2-celda. Por lo tanto  $F_2(X)$  es homeomorfo a  $F_2(X)/F_1(X)$ .  $\square$

4.2. TEOREMA. Sean  $n \geq 3$  y  $T_n$  un  $n$ -odo simple, entonces  $F_2(T_n)$  es homeomorfo a  $F_2(T_n)/F_1(T_n)$

DEMOSTRACIÓN. Sean  $n \in \mathbb{N}$ , con  $n \geq 3$ , y  $T_n$  un  $n$ -odo simple, supongamos que  $T_n = \bigcup_{i=1}^n \theta e_i$ , donde  $\theta$  es el origen de coordenadas de  $\mathbb{R}^n$  y  $\{e_1, \dots, e_n\}$  la base canónica de  $\mathbb{R}^n$ , como espacio vectorial. Definamos  $Z_n$  como sigue:

$$Z_n = \{A \in F_2(T_n) : e_i \in A, \text{ para algún } i \in \{1, \dots, n\}\}.$$

Por Lema 1, [3],  $F_2(T_n)$  es homeomorfo al cono sobre  $Z_n$ , y por Lema 2, [3]  $Z_n$  es la unión de una gráfica completa  $K_n$  y  $n$  arcos disjuntos entre ellos donde cada arco intersecta a  $K_n$  en exactamente uno de sus vértices.  $K_n$  tiene como vértices a  $\{\theta, e_1\}, \{\theta, e_2\}, \dots, \{\theta, e_n\}$ , la arista  $A_{i,j}$  que une a  $\{\theta, e_i\}$  y  $\{\theta, e_j\}$  está definida como

$$A_{i,j} = \{\{e_i, x\} \in Z_n : x \in \theta e_j\} \cup \{\{e_j, x\} \in Z_n : x \in \theta e_i\},$$

para cada  $i, j \in \{1, \dots, n\}$  con  $i \neq j$ .

Finalmente, los  $n$  arcos son de la forma

$$B_i = \{\{e_i, x\} \in F_2(T_n) : x \in \theta e_i\}, \text{ para cada } i \in \{1, \dots, n\}.$$

Notemos que  $F_1(T_n)$  intersecta a  $Z_n$ , sólo en los puntos terminales de los arcos  $B_i$ .

Definamos la gráfica completa  $K_n(F_2(T_n)/F_1(T_n))$  en  $F_2(T_n)/F_1(T_n)$ .

Los vértices de  $K_n(F_2(T_n)/F_1(T_n))$ , serán los puntos  $\{\{\theta, e_1\}\}, \{\{\theta, e_2\}\}, \dots, \{\{\theta, e_n\}\}$ , y la arista  $A_{i,j}(F_2(T_n)/F_1(T_n))$  que une a  $\{\{\theta, e_i\}\}$  con  $\{\{\theta, e_j\}\}$  está definida como

$$A_{i,j}(F_2(T_n)/F_1(T_n)) = \{\{\{e_i, x\}\} \in F_2(T_n)/F_1(T_n) : x \in \theta e_j\} \cup \{\{\{e_j, x\}\} \in F_2(T_n)/F_1(T_n) : x \in \theta e_i\}.$$

Es claro que  $\rho(K_n) = K_n(F_2(T_n)/F_1(T_n))$ , y como  $K_n \cap F_1(T_n) = \emptyset$ , entonces  $K_n(F_2(T_n)/F_1(T_n))$  es una gráfica completa de  $n$  vértices.

Para cada  $i \in \{1, \dots, n\}$ , sea  $B_i(F_2(T_n)/F_1(T_n)) = \{\{\{e_i, x\}\} : x \in \theta e'_i\}$ , donde  $e'_i$  es el vector en  $\mathbb{R}^n$  que se obtiene al multiplicar al elemento de la base canónica  $e_i$  por el escalar  $\frac{1}{2}$  es decir  $e'_i = \frac{1}{2}e_i$ , es claro que  $B_i(F_2(T_n)/F_1(T_n))$  es un arco en  $F_2(T_n)/F_1(T_n)$  y que cada  $B_i(F_2(T_n)/F_1(T_n))$  intersecta a  $K_n(F_2(T_n)/F_1(T_n))$  sólo en el vértice  $\{\{\theta, e_i\}\}$ . Así que

$$Z_n(F_2(T_n)/F_1(T_n)) = K_n(F_2(T_n)/F_1(T_n)) \cup \left( \bigcup_{i=1}^n B_i(F_2(T_n)/F_1(T_n)) \right)$$

es homeomorfo a  $Z_n$ .

Con lo anterior, para mostrar que  $F_2(T_n)$  es homeomorfo a  $F_2(T_n)/F_1(T_n)$ , es necesario mostrar que  $F_2(T_n)/F_1(T_n)$  es homeomorfo a *cono*  $(Z_n(F_2(T_n)/F_1(T_n)))$ .

Definamos  $h : \text{cono}Z_n(F_2(T_n)/F_1(T_n)) \rightarrow F_2(T_n)/F_1(T_n)$ , de la siguiente manera:

Cada elemento de  $\text{cono}Z_n(F_2(T_n)/F_1(T_n))$ , se puede ver como  $(\{\{e_i, p\}\}, t)$ , donde  $p \in \theta e_j$  con  $j \neq i$  o  $p \in \theta e'_i$ , y  $t \in [0, 1]$ , definamos entonces a  $h$  como:

$$h(\{\{e_i, p\}\}, t) = \begin{cases} \{\{(1-t)e_i, (1-t)p\}\}, & \text{si } p \in \theta e_j \text{ y } j \neq i, \\ \{\{2p, (1-t)e_i + 2tp\}\}, & \text{si } p \in \theta e'_i. \end{cases}$$

Notemos que si  $t = 1$ , entonces  $h(\{\{e_i, p\}\}, t) = F_1(T_n)$ , así que el vértice de  $\text{cono}Z_n(F_2(T_n)/F_1(T_n))$  es mandado en  $F_1(T_n) \in F_2(T_n)/F_1(T_n)$ . Luego se concluye que  $h$  puede reescribirse como:

$$h(\{\{e_i, p\}\}, t) = \begin{cases} \{\{(1-t)e_i, (1-t)p\}\}, & \text{si } \{\{e_i, p\}\} \in K_n(F_2(T_n)/F_1(T_n)), \\ \{\{2p, (1-t)e_i + 2tp\}\}, & \text{si } \{\{e_i, p\}\} \in B_i(F_2(T_n)/F_1(T_n)). \end{cases}$$

Para mostrar que la función  $h$  está bien definida, supongamos que  $[\{e_i, p\}] \in K_n(F_2(T_n)/F_1(T_n)) \cap B_i(F_2(T_n)/F_1(T_n))$ , entonces  $p = \theta$ , y como

$$[\{(1-t)e_i, (1-t)\theta\}] = [\{(1-t)e_i, \theta\}] = [\{2\theta, (1-t)e_i + 2t\theta\}]$$

para toda  $t \in [0, 1]$ , pues  $2\theta = \theta$  y  $(1-t)e_i + 2t\theta = (1-t)e_i$ , entonces  $h$  coincide en  $\text{cono}K_n(F_2(T_n)/F_1(T_n)) \cap \text{cono}B_i(F_2(T_n)/F_1(T_n))$ , y por tanto está bien definida.

Mostraremos que  $h$  es continua en  $\text{cono}K_n(F_2(T_n)/F_1(T_n))$  y que para cada  $i \in \{1, \dots, n\}$ ,  $h$  es continua en el cono sobre  $B_i(F_2(T_n)/F_1(T_n))$ . Dado que  $h$  coincide en el conjunto  $\text{cono}K_n(F_2(T_n)/F_1(T_n)) \cap \text{cono}B_i(F_2(T_n)/F_1(T_n))$ , para cada  $i \in \{1, \dots, n\}$ , entonces por el lema del pegamiento (ver Teorema 18.3, [8]),  $h$  será continua.

La función  $h|_{\text{cono}K_n(F_2(T_n)/F_1(T_n))}$  es la composición de la función  $f|_{\text{cono}K_n}$  dada en el Lema 1, [3], con la función cociente  $\rho_{T_n}$ . Por lo que  $h$  es continua en el cono sobre  $K_n(F_2(T_n)/F_1(T_n))$ . Podemos ver que la función  $f_1 : \text{cono}B_i \rightarrow F_2(T_n)$ , dada por  $f(\{e_i, p\}, t) = \{2tp, (1-t)e_i + 2tp\}$  es continua, en donde  $B_i = \{\{e_i, p\} : p \in \theta e'_i\}$ , así que  $\rho \circ f_1 = h|_{\text{cono}B_i(F_2(T_n)/F_1(T_n))}$  es continua.

Por lo tanto  $h$  es continua. Además en el Lema 2, [3] se demuestra que  $h$  es biyectiva. Así que  $h$  es un homeomorfismo. Por lo tanto se concluye que  $F_2(T_n)$  es homeomorfo a  $F_2(T_n)/F_1(T_n)$ .  $\square$

Ahora consideremos la curva cerrada simple.

4.3. TEOREMA. Si  $X = S^1$ , entonces  $F_2(S^1)/F_1(S^1)$  es homeomorfo al plano proyectivo real  $\mathbb{P}^2(\mathbb{R})$ .

DEMOSTRACIÓN. Sea  $\{x, y\} \in F_2(S^1)$ . Supongamos que  $x = (\sin(\theta), \cos(\theta))$  y  $y = (\sin(\phi), \cos(\phi))$ , para algunos  $\theta, \phi \in \mathbb{R}$ , con  $\theta \leq \phi$ . Definamos las relaciones  $l : F_2(X) \rightarrow \mathbb{R}$  y  $h : F_2(X) \rightarrow \mathbb{R}^2$  de la siguiente manera:

$$l(\{x, y\}) = \begin{cases} \phi - \theta & \text{si } \phi - \theta \leq \pi \\ 2\pi - (\phi - \theta) & \text{si } \phi - \theta \geq \pi \end{cases}$$

$$h(\{x, y\}) = \left( \sin\left(\frac{\theta + \phi}{2}\right), \cos\left(\frac{\theta + \phi}{2}\right) \right)$$

Notemos que  $l$  determina la longitud del arco menor o igual a  $2\pi$  formado por  $x$  y  $y$ , y que  $h$  determina al punto medio sobre  $S^1$  entre los puntos  $x$  y  $y$ . Notemos que  $h(\{x, y\})$  está bien determinado cuando  $x$  no es el antípoda de  $y$  y, en realidad,  $h$  no es una función pues cuando  $x$  es la antípoda de  $y$ ,  $h$  asigna a la pareja  $\{x, y\}$  un par de puntos sobre  $S^1$ , que de hecho son también antípodas, pero pensemos por el momento que cuando  $x$  es antípoda de  $y$ , la relación  $h$  asigna un par de puntos a la pareja  $x, y$ .

Con base en estas dos funciones podemos construir un homeomorfismo del espacio topológico  $F_2(S^1)/F_1(S^1)$  en el plano proyectivo real  $\mathbb{P}^2(\mathbb{R})$ ,  $g : F_2(S^1)/F_1(S^1) \rightarrow \mathbb{P}^2(\mathbb{R})$ , de la siguiente manera

$$g(\{x, y\}) = \frac{l(\{x, y\})}{\pi} h(\{x, y\}), \text{ para cada } \{x, y\} \in F_2(S^1)/F_1(S^1).$$

donde  $l(\{x, y\})h(\{x, y\})$  denota al producto escalar del vector  $h(\{x, y\})$  con el número real  $\frac{l(\{x, y\})}{\pi}$ .

Podemos ver que la función  $g$  está bien definida pues si  $\{x\} \in F_1(X)$ , entonces  $g(\{x\}) = (0, 0)$  y que la imagen de  $g$  en efecto es el plano proyectivo real  $\mathbb{P}^2(\mathbb{R})$ , pues para los puntos de la forma  $\{x, y\}$ , donde  $x$  no es antípoda de  $y$ , la función  $g$  asigna al punto  $[\{x, y\}]$  un único punto en el disco unitario con norma menor estricta que 1 y para el caso en que  $x$  es antípoda de  $y$ , es decir  $x = (\cos(\theta), \sin(\theta))$  y  $y = (\cos(\theta + \pi), \sin(\theta + \pi))$  para algún  $\theta \in \mathbb{R}$ , la función  $g$  asigna al punto  $[\{x, y\}]$  el par de puntos antípodas en el disco unitario

$$\left(\cos\left(\theta + \frac{\pi}{2}\right), \sin\left(\theta + \frac{\pi}{2}\right)\right) \text{ y } \left(\cos\left(\theta + \frac{3\pi}{2}\right), \sin\left(\theta + \frac{3\pi}{2}\right)\right)$$

que son el mismo punto en  $\mathbb{P}^2(\mathbb{R})$ .

Por el Teorema de Transgresión (véase Teorema 22.2, [5]), la función  $g$  es continua. Como  $F_2(S^1)/F_1(S^1)$  es compacto y  $\mathbb{P}^2(\mathbb{R})$  es un espacio de Hausdorff, para concluir que  $g$  sea homeomorfismo bastará mostrar que  $g$  es una función biyectiva.

Primero veremos por que  $g$  es inyectiva. Notemos que  $g([\{x, y\}]) = (0, 0)$  si y sólo si  $l(\{x, y\}) = 0$  y esto ocurre sólo cuando  $\{x, y\} \in F_1(S^1)$ , es decir,  $[\{x, y\}] = F_1(S^1)$ . Ahora supongamos que

$$g([\{a, b\}]) = g([\{c, d\}]) \neq (0, 0)$$

Entonces  $l(\{a, b\}) = l(\{c, d\}) \neq 0$  y  $h(\{a, b\}) = h(\{c, d\})$ , y esto ocurre si y sólo si  $\{a, b\} = \{c, d\}$  (ver pág. 877, [1]). Por lo que se concluye que  $g$  es inyectiva.

Sea  $q = \rho(\cos(\phi), \sin(\phi))$  un punto en  $\mathbb{P}^2(\mathbb{R})$ . Si  $\rho = 0$  entonces  $g(F_1(S^1)) = q$ . Supongamos que  $\rho \in (0, 1]$ . Definamos  $x, y \in S^1$  de la siguiente manera

$$x = \left(\cos\left(\phi + \frac{\rho}{2}\right), \sin\left(\phi + \frac{\rho}{2}\right)\right)$$

$$y = \left(\cos\left(\phi - \frac{\rho}{2}\right), \sin\left(\phi - \frac{\rho}{2}\right)\right)$$

Resulta que  $g([\{x, y\}]) = q$ . Así que  $g$  es sobreyectiva.

Por lo tanto  $F_2(S^1)/F_1(S^1)$  es homeomorfo a  $\mathbb{P}^2(\mathbb{R})$ . □

Dado que la banda de Möbius no es homeomorfa al plano proyectivo real tenemos el siguiente problema.

**4.4. PROBLEMA.** Caracterizar los continuos  $X$  para los cuales  $F_2(X)$  es homeomorfo a  $F_2(X)/F_1(X)$

Y de manera general.

4.5. PROBLEMA. Dados  $n, m \in \mathbb{N}$  con  $m < n$ . Caracterizar los continuos  $X$  para los cuales  $F_n(X)$  es homeomorfo a  $F_n(X)/F_m(X)$ .

#### REFERENCIAS

- [1] K. Borsuk y S. Ulam, *On symmetric products of topological spaces*, Bull. Amer. Math. Soc. 37(1931), 875 - 882.
- [2] R. Bott, *On the third symmetric product potency of  $S^1$* , Fund. Math. 39(1952), 364 - 368.
- [3] E. Castañeda, *Embedding symmetric products in Euclidean spaces*, Continuum Theory. Lecture Notes in Pure Appl. Math., 230. New York: Dekker, (2002) 67-79.
- [4] E. Castañeda, *Symmetric products as cones and products*, Topology Proc. 28 (2004), 55-67.
- [5] J. Dugundji, *Topology*, Allyn and Bacon, Inc., Boston, 1966.
- [6] A. Illanes, *Hiperespacios de Continuos*, Serie de Aportaciones Matemáticas No. 28, Sociedad Matemática Mexicana, México, (2004).
- [7] R. Molski, *On symmetric products*, Fund. Math. 44 (1957), 165 - 170.
- [8] J. M. Munkres, *Topología. 2ª edición*. Prentice Hall, España (2002).
- [9] S.B. Nadler, Jr., *Continuum Theory, An Introduction*, Marcel Dekker, New York, Basel and Hong Kong, 1992.

Facultad de Ciencias, Universidad Autónoma del Estado de México

Instituto Literario No. 100, Colonia Centro

Toluca, Estado de México, México. C.P. 50000

eca@uaemex.mx; eca1972@gmail.com

matjavier@gmail.com

# EL SEGUNDO PRODUCTO SIMÉTRICO DE DENDRITAS

ENRIQUE CASTAÑEDA ALVARADO  
LÁZARO TREJO ALLENDE  
UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO

RESUMEN. En este artículo abordamos el problema de caracterizar al segundo producto simétrico de un continuo como un cono sobre algún continuo, en particular para la clase de dendritas.

## 1. INTRODUCCIÓN

Un continuo  $X$  es un espacio métrico, compacto, conexo y no degenerado. Dado un continuo  $X$  con una métrica  $d$ , un hiperespacio de  $X$  es una familia que se construye tomando subconjuntos de  $X$  con alguna característica particular. Algunos hiperespacios de  $X$  que han sido ampliamente estudiados son:

$$\begin{aligned}2^X &= \{A \subset X : A \neq \emptyset \text{ y cerrado}\}, \\C(X) &= \{A \in 2^X : A \text{ es conexo}\}, \\&\text{y para un entero positivo } n, \\F_n(X) &= \{A \in 2^X : A \text{ tiene a lo más } n \text{ elementos}\}.\end{aligned}$$

Al conjunto  $2^X$  se le da una métrica que se define en términos de la métrica del continuo  $X$ , dicha métrica es conocida como métrica de Hausdorff, véanse [8] y [9]. De esta forma podemos considerar a  $C(X)$  y  $F_n(X)$  con la métrica heredada como subespacios de  $2^X$ . Con esta métrica los hiperespacios  $2^X$ ,  $C(X)$  y  $F_n(X)$  resultan ser también continuos. A  $F_n(X)$  se le conoce como el  $n$ -ésimo producto simétrico del continuo  $X$ . Los productos simétricos fueron introducidos por K. Borsuk y S. Ulam en [2]. Ellos probaron que, para  $I = [0, 1]$  y  $n = 1, 2, 3$ ,  $F_n(I)$  es homeomorfo a  $I^n$ . También en [2] se puede ver que  $F_2(S^1)$  es homeomorfo a la banda de Moebius, donde  $S^1$  es una curva cerrada simple. El problema de caracterizar aquellos continuos  $X$  de dimensión finita para los cuales su hiperespacio es homeomorfo a un cono o a un producto topológico ha sido ampliamente estudiado. Véanse [1], [2], [3], [4], [5], [6] y [7]. En [3] y [4] se prueba que en la clase de gráficas finitas se tiene que el segundo producto simétrico de un continuo es homeomorfo al cono sobre algún continuo, si y sólo si el continuo es un arco o un  $n$ -odo simple.

En este artículo abordamos este problema principalmente para la clase de las dendritas. Concretamente probamos los siguientes resultados cuando  $X$  es una dendrita:

- $F_2(X)$  es homeomorfo al cono de  $X$ , si y sólo si  $X$  es un arco.
- $F_2(X)$  es homeomorfo al cono de  $Y$  para algún continuo plano  $Y$ , si y sólo si  $X$  es un arco, un triodo simple o un 4-odo simple.

## 2. DEFINICIONES Y RESULTADOS PREVIOS

2.1. DEFINICIÓN. Una  $n$ -celda se define como  $B^n = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$ , para  $n = 1, 2, \dots$

2.2. DEFINICIÓN. Decimos que el espacio  $X$  es arco conexo si para cualesquiera  $x, y \in X$  existe una función continua  $\alpha : [0, 1] \rightarrow X$ , tal que  $\alpha(0) = x$  y  $\alpha(1) = y$ . La función  $\alpha$  es llamada arco que conecta a  $x$  con  $y$ .

2.3. DEFINICIÓN. Dado un continuo  $X$ . Decimos que  $A \subset X$  es un arco componente de  $X$  si  $A$  es maximal en  $X$  con respecto a la propiedad de ser arco conexo.

2.4. DEFINICIÓN. Un continuo  $X$  es localmente conexo, si para todo  $x \in X$  y cada abierto  $U$  que contiene a  $x$ , existe un abierto conexo  $V \subset U$  que también contiene a  $x$ .

2.5. DEFINICIÓN. Una gráfica finita  $X$  es un continuo que es la unión finita de arcos, tales que cada par de ellos se intersectan en un subconjunto de sus puntos extremos.

2.6. DEFINICIÓN. Una gráfica completa  $K_n$  es una gráfica finita de  $n$  vértices, que se obtiene al unir cada par de vértices por medio de un arco.

2.7. DEFINICIÓN. Un  $n$ -odo simple  $T_n$  es una gráfica finita que es la unión de  $n$  arcos  $J_1, \dots, J_n$  tal que existe un punto  $p \in T_n$  con la propiedad de que  $J_i \cap J_j = \{p\}$ , si  $i \neq j$  y  $p$  es un punto extremo de cada arco  $J_i$ .

2.8. DEFINICIÓN. Una paleta se define como la unión de una circunferencia y un arco, tal que solo un punto extremo del arco intersecta a la circunferencia.

2.9. DEFINICIÓN. Una medalla se considera como la unión de una circunferencia y dos arcos, tal que solo un punto extremo de cada arco intersecta a la circunferencia en un mismo punto.

2.10. DEFINICIÓN. El continuo figura ocho se define como la unión de dos circunferencias, tal que su intersección es un solo punto.

2.11. DEFINICIÓN. Una dendrita  $X$  es un continuo localmente conexo que no tiene curvas cerradas simples.

2.12. DEFINICIÓN. El cono de  $X$  es el cociente  $X \times [0, 1] / X \times \{1\}$  con la topología de identificación, al cual denotaremos como  $Cono(X)$ .

La prueba de los siguientes dos resultados se puede consultar en [3].

2.13. TEOREMA. Sea  $X$  un continuo localmente conexo. Entonces  $F_2(X)$  se puede encajar en  $\mathbb{R}^3$  si y sólo si  $X$  es homeomorfo a alguno de los siguientes espacios: un arco, un tríodo simple, un 4-odo simple, una circunferencia, una paleta, una medalla o un continuo figura ocho.

2.14. LEMA. Si  $X$  es un  $n$ -odo simple, entonces  $F_2(X)$  es homeomorfo al  $Cono(Z)$ , donde  $Z$  es la unión de una gráfica completa  $K_n$  y  $n$  arcos disjuntos cada uno de los cuales intersecta a  $K_n$  en exactamente uno de sus vértices.

## 3. RESULTADOS

3.1. TEOREMA. Si  $X$  es una  $n$ -celda, entonces  $F_2(X)$  es homeomorfo al  $Cono(X)$  si y sólo si  $X$  es un arco.

DEMOSTRACIÓN. Como el  $F_2(I)$  es homeomorfo a  $I^2$  y a su vez el  $Cono(I)$  es también homeomorfo a  $I^2$ , se tiene que  $F_2(I)$  es homeomorfo al  $Cono(I)$ , con lo cual se concluye la suficiencia.

Ahora para la necesidad, se tiene que la dimensión de  $F_2(X)$  es al menos el doble de la dimensión de  $X$  y como la dimensión de una  $n$ -celda es  $n$ , se tiene que la dimensión de  $F_2(X)$  es al menos  $2n$ .

Ahora dado que el  $Cono(X)$  tiene dimensión  $n + 1$ , entonces el  $Cono(X)$  no puede ser homeomorfo a  $F_2(X)$  para  $n > 1$ . Y por lo tanto el homeomorfismo se da sólo cuando  $n = 1$ .  $\square$

3.2. LEMA. Si  $X$  es un continuo arco conexo, entonces  $F_n(X)$  es arco conexo.

DEMOSTRACIÓN. Mostraremos que para cualesquiera  $A, B \in F_n(X)$  podemos encontrar un arco que conecta a  $A$  con  $B$ .

Para esto, sean  $A = \{x_1, \dots, x_m\}$  y  $B = \{y_1, \dots, y_k\}$  con  $m, k \leq n$ . Repitiendo  $x_m$  y  $y_k$  tantas veces como sea necesario, podemos suponer que tanto  $A$  como  $B$  tienen  $n$  elementos.

Y así, como  $X$  es arco conexo, para cada  $i \in \{1, \dots, n\}$  existe una función continua  $\alpha_i : [0, 1] \rightarrow X$  tal que  $\alpha_i(0) = x_i$  y  $\alpha_i(1) = y_i$ . De manera que podemos considerar a la función

$\alpha : [0, 1] \rightarrow F_n(X)$  dada por  $\alpha(t) = \{\alpha_1(t), \dots, \alpha_n(t)\}$ , la cual es un arco que une a  $A$  con  $B$  en  $F_n(X)$ .  $\square$

3.3. TEOREMA. Si  $X$  es un continuo que tiene al menos dos arco componentes, entonces  $F_2(X)$  no es homeomorfo al  $Cono(X)$ .

DEMOSTRACIÓN. Sea  $X = X_1 \cup \dots \cup X_n$  con  $n \geq 2$ . Para cualquier  $A = \{x_1, x_2\} \in F_2(X)$  se tiene que  $A \subset X_i$  para algún  $i \in \{1, 2, \dots, n\}$ , o bien  $A \cap X_i \neq \emptyset$  y  $A \cap X_k \neq \emptyset$  para  $i \neq k$ .

Esto significa que  $A \subset F_2(X_i)$  o bien que a  $A$  lo podemos identificar como un elemento del producto  $X_i \times X_k$ , para  $i, k \in \{1, 2, \dots, n\}$  con  $i \neq k$ .

Ahora por 3.2 se tiene que  $F_2(X_i)$  es arco conexo para todo  $i \in \{1, 2, \dots, n\}$ , así que  $F_2(X) = \bigcup_{i=1}^n F_2(X_i) \cup \bigcup_{i \neq k} (X_i \times X_k)$ , es decir que  $F_2(X)$  consta de al menos dos arco componentes.

Y dado que el  $Cono(X)$  tiene solo una arco componente, se tiene que  $F_2(X)$  no es homeomorfo al  $Cono(X)$ .  $\square$

3.4. TEOREMA. Si  $X$  es una dendrita, entonces  $F_2(X)$  es homeomorfo al  $Cono(X)$  si y sólo si  $X$  es homeomorfo al intervalo  $I$ .

DEMOSTRACIÓN. Como  $F_2(I)$  es homeomorfo a  $I^2$  y además el  $Cono(X)$  también es homeomorfo a  $I^2$ , se tiene que  $F_2(X)$  es homeomorfo al  $Cono(X)$ , con lo cual se tiene la suficiencia.

Ahora para obtener la necesidad, dado que  $X$  es una dendrita se tiene que  $X$  es plano, véase [10], ejemplo 10.37, p.181. Así que el  $Cono(X)$  se puede encajar en  $\mathbb{R}^3$ . Ahora como  $F_2(X)$  es homeomorfo al  $Cono(X)$  se sigue que  $F_2(X)$  se puede encajar en  $\mathbb{R}^3$ .

De manera que por 2.13 y dado que  $X$  es una dendrita se tiene que  $X$  es homeomorfo a alguno de los siguientes espacios: un arco  $I$ , un tríodo simple  $T$  o un 4-odo simple  $C$ .

Sin embargo de estos espacios, el único para el cual  $F_2(X)$  es homeomorfo al  $Cono(X)$  es el arco  $I$ , ya que por 2.14 se tiene que  $F_2(T)$  es homeomorfo al  $Cono(Z)$ , donde  $Z$  es la unión de  $K_3$  y 3 arcos disjuntos que intersectan a  $K_3$  en solo uno de sus vértices y además  $F_2(C)$  es homeomorfo al  $Cono(Z)$ , donde  $Z$  es la unión de  $K_4$  y 4 arcos disjuntos que intersectan a  $K_4$  en solo uno de sus vértices. Por lo que  $X$  es homeomorfo al intervalo  $I$ .  $\square$

3.5. TEOREMA. Si  $X$  es una dendrita, entonces  $F_2(X)$  es homeomorfo al  $Cono(Y)$  para algún continuo plano  $Y$  si y sólo si  $X$  es homeomorfo al intervalo, a un tríodo simple o a un 4-odo simple.

DEMOSTRACIÓN. Dado que  $Y$  es plano, se tiene que el  $Cono(Y)$  se puede encajar en  $\mathbb{R}^3$ . Ahora como  $F_2(X)$  es homeomorfo al  $Cono(Y)$  se obtiene que  $F_2(X)$  se puede encajar en  $\mathbb{R}^3$ . Así que aplicando 2.13 y dado que  $X$  es una dendrita, se tiene que  $X$  es homeomorfo a alguno de los siguientes espacios: un arco, un tríodo simple o un 4-odo simple, con lo cual se concluye la necesidad.

La suficiencia se obtiene a partir de 2.14.  $\square$

#### REFERENCIAS

- [1] Andablo Reyes, G. y Castañeda Alvarado E. (2009). *Una mirada a los productos simétricos*. Ciencia Ergo Sum, Vol.16, No. 2, pp.189-197.
- [2] Borsuk, K. y Ulam, S. (1931). *On symmetric products of topological space*. Bull. Amer. Math. Soc., 37, pp. 875-882.
- [3] Castañeda, E. (2002). *Embedding symmetric product in Euclidean spaces*. Illanes, A.; Macias S.; Lewis, W. (Ed), Continuum Theory (Denton, TX, 1999), Lecture notes in Pure and Applied Mathematics, 230, Marcel Dekker, New York, pp. 137-142.
- [4] Castañeda, E. (2004). *Symmetric products as cones and products*. Topology Proc. Vol. 28, No. 1, pp. 55-67.
- [5] Illanes, A. (1995). *Hyperspaces homeomorphic to cones*, Glasnik Mat., 30(50), pp. 285-294.
- [6] Illanes, A. (1997). *Hyperspaces which are products*, Topology Appl., 79, pp. 229-247.
- [7] Illanes, A. y López, M de J. (2002). *Hyperspaces homeomorphic to cones II*, Topology Appl., 126, pp. 377-391.
- [8] Illanes, A, y Nadler Jr., S. (1999). *Hyperspaces, Fundamentals and Recent Advances*. Monographs and Textbooks in Pure and Applied Mathematics, Vol. 216, Marcel Dekker, Inc. New York, N.Y.
- [9] Nadler Jr., S. (1978). *Hyperspaces of sets. Monographs and Textbooks in Pure and Applied Mathematics*, Vol. 49, Marcel Dekker, Inc. New York, N.Y.
- [10] Nadler Jr., S. (1992). *Continuum Theory, An Introduction*. Monographs and Textbooks in Pure and Applied Mathematics, Vol. 158, Marcel Dekker, Inc., New York, N.Y.

Facultad de Ciencias, Universidad Autónoma del Estado de México  
 Instituto Literario No. 100, Colonia Centro  
 Toluca, Estado de México, México. C.P. 50000  
 eca@uaemex.mx  
 lzarotrejo@yahoo.com.mx

# DENDRITAS CUYO CONJUNTO DE PUNTOS EXTREMOS ES CERRADO

DAVID HERRERA CARRASCO  
 FERNANDO MACÍAS ROMERO  
 FRANCISCO VÁZQUEZ JUÁREZ  
 FCFM - BUAP

RESUMEN. Un continuo es un espacio métrico, compacto, conexo y no vacío. Una dendrita es un continuo localmente conexo que no contiene curvas cerradas simples. En este trabajo demostramos que todo subcontinuo de una dendrita con el conjunto de puntos extremos cerrado es una dendrita con el conjunto de puntos extremos cerrado. Además, damos una caracterización de las dendritas  $G^n$ .

## 1. INTRODUCCIÓN

El material que contiene este artículo pertenece a la rama de la topología conocida como Teoría de Continuos. Dicha teoría trata del estudio de las propiedades topológicas de espacios que son métricos, compactos, conexos y no vacíos. De hecho a un espacio topológico con estas propiedades se le llama continuo.

Un espacio topológico  $X$  es localmente conexo en un punto  $x \in X$ , si para cada vecindad  $V$  de  $x$  existe un subconjunto abierto y conexo  $U$  de  $X$  tal que  $x \in U \subset V$ . El espacio topológico  $X$  es localmente conexo si es localmente conexo en cada uno de sus puntos. Una dendrita es un continuo localmente conexo que no contiene curvas cerradas simples.

Para una dendrita  $X$ , el orden de  $X$  en un punto  $p$ ,  $ord(X, p)$ , está definido como el número de componentes de  $X - \{p\}$ . Los puntos de orden 1, 2 y mayor o igual a 3 son los puntos extremos, puntos ordinarios y puntos de ramificación de  $X$ , respectivamente. Los conjuntos de puntos extremos, puntos ordinarios y puntos de ramificación de  $X$  son denotados por  $E(X)$ ,  $O(X)$  y  $R(X)$ , respectivamente. Luego,  $X = O(X) \cup R(X) \cup E(X)$ . Para cada  $n \in \mathbb{N}$ , el símbolo  $R_n(X)$  denota el conjunto de los puntos de orden  $n$ .

Para cada  $n \in \mathbb{N} - \{1, 2\}$  definimos una dendrita  $G^n$  como sigue.

Sean  $i \in \mathbb{N}$  y  $\alpha_1, \dots, \alpha_i \in \{0, 2, \dots, 2n-4\}$ , definimos  $E_{\alpha_1, \dots, \alpha_i}^n$  como el conjunto de los números  $x$  en  $[0, 1]$  tal que si  $x$  se expresa en base  $2n-3$ , entonces los primeros  $i$  dígitos son exactamente  $\alpha_1, \dots, \alpha_i$ , es decir,

$$E_{\alpha_1, \dots, \alpha_i}^n = \left[ \frac{\alpha_1}{2n-3} + \dots + \frac{\alpha_i}{(2n-3)^i}, \frac{\alpha_1}{2n-3} + \dots + \frac{\alpha_{i-1}}{(2n-3)^{i-1}} + \frac{\alpha_i+1}{(2n-3)^i} \right].$$

Si todos los dígitos  $\alpha_1, \dots, \alpha_i$  son pares, sea  $p_{\alpha_1, \dots, \alpha_i}^n$  el punto en el plano cuya primera coordenada es la mitad del intervalo  $E_{\alpha_1, \dots, \alpha_i}^n$  y la segunda coordenada es  $\frac{1}{2^i}$ . Si  $i = 0$ , entonces  $E^n = [0, 1]$  y  $p^n = (\frac{1}{2}, 1)$ . Sean  $q = (\frac{1}{2}, 2)$  y  $H^n$  definido por:

$$H^n = [q, p^n] \cup \left( \bigcup_{i \in \mathbb{N}, \alpha_1, \dots, \alpha_i \in \{0, 2, \dots, 2n-4\}} [p_{\alpha_1, \dots, \alpha_{i-1}}^n, p_{\alpha_1, \dots, \alpha_{i-1}, \alpha_i}^n] \right).$$

Dada una sucesión  $\alpha_1, \alpha_2, \dots$  la intersección  $\bigcap_{i \in \mathbb{N}} E_{\alpha_1, \dots, \alpha_i}^n$  es un conjunto de un sólo punto. Denotemos tal punto por  $e_{\alpha_1, \alpha_2, \dots}$ , sea  $p_{\alpha_1, \alpha_2, \dots} = (e_{\alpha_1, \alpha_2, \dots}, 0) \in H^n$  y notemos que

$$E(H^n) = \{q\} \cup \{p_{\alpha_1, \alpha_2, \dots} : \alpha_i \in \{0, 2, \dots, 2n-4\}\}.$$

Luego, definimos  $G^n$  como la unión de dos subcontinuos homeomorfos  $H^n$  y  $(H^n)^*$  de  $H^n$ , identificando los puntos correspondientes  $q$  y  $q^*$ .

Los resultados principales de este trabajo que encontramos originalmente enunciados en [1, Teorema 3.2 y 4.1] son los siguientes:

- (a) Se demuestra que si  $X$  es una dendrita tal que  $E(X)$  es cerrado, entonces todo subcontinuo  $Y$  de  $X$  es una dendrita y  $E(Y)$  es cerrado.
- (b) Caracterizamos las dendritas  $G^n$ .

## 2. PRELIMINARES

En esta sección enunciamos resultados sin demostración pero damos una referencia adecuada.

Si  $X$  es un espacio topológico y  $A$  un subconjunto de  $X$ , los símbolos  $\bar{A}$ ,  $Fr(A)$  y  $A'$  denotarán la cerradura de  $A$ , la frontera de  $A$  y el derivado de  $A$  en  $X$ , respectivamente.

Sean  $X$  un espacio topológico y  $p \in X$ , un subconjunto  $V$  de  $X$  es una vecindad de  $p$  si existe un abierto  $U$  en  $X$  tal que  $p \in U \subset V$ .

Sean  $X$  un espacio métrico con métrica  $d$ ,  $p \in X$  y  $\epsilon > 0$ , la bola abierta en  $X$  con centro en  $p$  y radio  $\epsilon$ , denotada por  $B_\epsilon(p)$ , es el conjunto  $B_\epsilon(p) = \{x \in X : d(p, x) < \epsilon\}$ . De hecho, todos los conceptos no definidos en este trabajo son tomados como en [2].

2.1. TEOREMA. [2, Teorema 5.9] Sean  $X$  un continuo y  $A$  un subcontinuo propio de  $X$ . Si  $C$  es una componente de  $X - A$ , entonces  $C \cup A$  es un continuo.

2.2. DEFINICIÓN. Un *árbol* es una gráfica finita sin curvas cerradas simples.

2.3. TEOREMA. [2, Teorema 9.28] Sea  $X$  un continuo. Entonces  $X$  es un árbol si y sólo si el conjunto de puntos extremos de  $X$  es finito.

2.4. TEOREMA. [5, Teorema 2.60] Sean  $X$  una dendrita y  $p, q \in X$ , entonces existe un único arco en  $X$  con extremos  $p$  y  $q$ .

Dos puntos  $x, y$  en la dendrita  $X$  están unidos por un único arco en  $X$  que denotamos por  $[x, y]$ .

2.5. DEFINICIÓN. Un continuo  $X$  es hereditariamente localmente conexo si todo subcontinuo de  $X$  es localmente conexo.

2.6. TEOREMA. [2, Corolario 10.5] Toda dendrita es hereditariamente localmente conexo.

2.7. TEOREMA. [2, Corolario 10.6] Todo subcontinuo de una dendrita es una dendrita.

2.8. DEFINICIÓN. Sean  $X$  un espacio topológico conexo y  $p \in X$ . Si  $X - \{p\}$  no es conexo, entonces  $p$  es un punto de corte de  $X$ .

2.9. TEOREMA. [6, Teorema 2.17] Sean  $X$  un continuo localmente conexo,  $p$  un punto de corte de  $X$  y  $C$  una componente de  $X - \{p\}$ . Si  $q \in E(\overline{C})$ ,  $q \neq p$ , entonces  $q \in E(X)$ .

2.10. TEOREMA. [2, Teorema 10.7] Sea  $X$  un continuo no degenerado. Entonces  $X$  es una dendrita si y sólo si cada punto de  $X$  es un punto de corte de  $X$  o un punto extremo de  $X$ .

2.11. DEFINICIÓN. Un espacio topológico  $X$  es arco conexo si para cualesquiera  $x, y \in X$  con  $x \neq y$ , existe un arco en  $X$  con extremos  $x$  y  $y$ .

2.12. TEOREMA. [2, Teorema 10.9] Cada subconjunto conexo no degenerado de una dendrita es arco conexo.

2.13. DEFINICIÓN. Si  $X$  es un espacio topológico conexo y  $p \in X$ , entonces el número de componentes de  $p$  en  $X$ , denotado por  $c(p, X)$ , es la cardinalidad del conjunto de las componentes de  $X - \{p\}$ .

2.14. TEOREMA. [2, Teorema 10.12] Sean  $X$  un continuo no degenerado y  $p \in X$ . Si  $ord(p, X)$  es finito, entonces  $c(p, X)$  es finito y, de hecho,  $c(p, X) \leq ord(p, X)$ .

2.15. TEOREMA. [2, Teorema 10.13] Sea  $X$  un continuo no degenerado. Entonces  $X$  es una dendrita si y sólo si  $c(p, X) = ord(p, X)$ , siempre que alguno de ellos sea finito.

2.16. DEFINICIÓN. Sean  $X$  un espacio métrico y  $\{A_n\}_{n=1}^{\infty}$  una sucesión de subconjuntos de  $X$ . La sucesión  $\{A_n\}_{n=1}^{\infty}$  es una sucesión nula si  $\lim_{n \rightarrow \infty} diám(A_n) = 0$ .

2.17. TEOREMA. [3, Corolario 2.2, pág. 90] En un continuo hereditariamente localmente conexo, las componentes de cualquier conjunto abierto forman una sucesión nula cuando exista una cantidad infinita de componentes.

2.18. DEFINICIÓN. Una familia  $\{A_\alpha : \alpha \in \Lambda\}$  de conjuntos en un espacio topológico  $X$  es una familia de vecindades finita, si cada punto de  $X$  tiene una vecindad  $V$  tal que  $V \cap A_\alpha \neq \emptyset$ , para a lo más una cantidad finita de índices  $\alpha$ .

2.19. TEOREMA. [4, 9.4] Sean  $X$  un espacio topológico y  $\{A_\alpha : \alpha \in \Lambda\}$  una cubierta de  $X$  tal que:

- (1) los conjuntos  $A_\alpha$  son todos abiertos o
- (2) los conjuntos  $A_\alpha$  son cerrados, y forman una familia de vecindades finita.

Para cada  $\alpha \in \Lambda$ , sea  $f_\alpha : A_\alpha \rightarrow Y$  continua y supongamos que  $f_\alpha|_{A_\alpha \cap B_\beta} = f_\beta|_{A_\alpha \cap B_\beta}$  para cada  $(\alpha, \beta) \in \Lambda \times \Lambda$ . Entonces existe una única función continua  $f : X \rightarrow Y$ , la cual es una extensión de cada  $f_\alpha$ ; es decir,  $f|_{A_\alpha} = f_\alpha$  para cada  $\alpha \in \Lambda$ .

### 3. PROPIEDADES GENERALES

En esta sección definimos una dendrita y la denotamos por  $F_\omega$ , posteriormente probamos un resultado que nos dice bajo qué condiciones una dendrita tiene un subcontinuo homeomorfo a  $F_\omega$ ; también vemos que la propiedad de ser una dendrita con el conjunto de puntos extremos cerrado es una propiedad hereditaria.

Para nuestro estudio las siguientes dendritas son importantes.

3.1. DEFINICIÓN. Definimos tres dendritas especiales en el plano. Una de ellas es la dendrita  $F_\omega = \bigcup_{n \in \mathbb{N}} [a, a_n]$ , donde  $a = (0, 0)$  y  $a_n = (\frac{1}{n}, \frac{1}{n^2})$ .

Las otras dendritas son  $W_R = [a, b_1] \cup (\bigcup_{n \in \mathbb{N}} [t_n, b_n])$  y  $W = [c, a] \cup W_R$ , donde para cada  $n \in \mathbb{N}$  consideramos los puntos  $t_n = (\frac{1}{n}, \frac{1}{n})$ ,  $b_n = (\frac{1}{n}, 0)$  y  $c = (-1, 0)$ .

En seguida un resultado que nos da una condición suficiente para que una dendrita tenga un subcontinuo homeomorfo a  $F_\omega$ .

3.2. TEOREMA. Sea  $X$  una dendrita. Si existe  $x \in X$  tal que  $\text{ord}(x, X) = \omega$ , entonces  $X$  contiene un subcontinuo homeomorfo a  $F_\omega$ .

DEMOSTRACIÓN. Sea  $x \in X$  tal que  $\text{ord}(x, X) = \omega$ . Por el Teorema 2.15 y el Teorema 2.14, tenemos que  $c(x, X) = \omega$ . Sean  $C_1, C_2, \dots$  las componentes de  $X - \{x\}$ . Por el Teorema 2.6,  $X$  es hereditariamente localmente conexo, así por el Teorema 2.17,  $\lim_{n \rightarrow \infty} \text{diám}(C_n) = 0$ . Para cada  $n \in \mathbb{N}$ , por el Teorema 2.1,  $C_n \cup \{x\}$  es un continuo. Por el Teorema 2.12,  $C_n \cup \{x\}$  es arco conexo. Para cada  $n \in \mathbb{N}$ , sean  $y_n \in C_n$  y  $[y_n, x] \subset C_n \cup \{x\}$ . Así,  $\lim_{n \rightarrow \infty} \text{diám}([y_n, x]) = 0$  y  $[y_n, x] \cap [y_m, x] = \{x\}$  si  $n \neq m$ . Para cada  $n \in \mathbb{N}$ , sea  $A_n = [y_n, x]$ . Veamos que,  $Y = \bigcup_{n \in \mathbb{N}} A_n$  es homeomorfo a  $F_\omega$ .

Para cada  $n \in \mathbb{N}$ , sea  $B_n = [a_n, a]$  ( $a_n$  y  $a$  como en la Definición 3.1). Para cada  $n \in \mathbb{N}$ , existe  $f_n : A_n \rightarrow B_n$  un homeomorfismo con  $f_n(x) = a$ . En particular para cada  $n \in \mathbb{N}$ ,  $f_n : A_n - \{x\} \rightarrow F_\omega$  es continua. Note que la colección  $\{A_n - \{x\}\}_{n=1}^\infty$  es una cubierta abierta de  $Y - \{x\}$ , entonces por el Teorema 2.19, existe una función  $f : Y - \{x\} \rightarrow F_\omega$  continua tal que  $f|_{A_n - \{x\}} = f_n$  para cada  $n \in \mathbb{N}$ .

Definimos  $g : Y \rightarrow F_\omega$  por:

$$g(y) = \begin{cases} f(y), & \text{si } y \neq x; \\ a, & \text{si } y = x. \end{cases}$$

Veamos que  $g$  es continua en  $x$ . Sea  $U$  abierto en  $F_\omega$  tal que  $a \in U$ . Como  $\lim_{n \rightarrow \infty} \text{diám}(B_n) = 0$ , existe  $n_o \in \mathbb{N}$  tal que si  $n \geq n_o$ , entonces  $B_n \subset U$ . Esto implica que:

$$(i) \bigcup_{n \geq n_o} B_n \subset U.$$

Como  $U \cap B_1, U \cap B_2, \dots, U \cap B_{n_o-1}$  son abiertos en  $B_1, B_2, \dots, B_{n_o-1}$  respectivamente tal que contienen al punto  $a$  y dado que las funciones  $f_n$  son continuas, existen  $V_1, V_2, \dots, V_{n_o-1}$  abiertos en  $A_1, A_2, \dots, A_{n_o-1}$  respectivamente que contienen al punto  $x$  tal que:

$$(ii) f_j(V_j) \subset U \cap B_j \text{ para cada } j \in \{1, \dots, n_o - 1\}.$$

Luego para cada  $j \in \{1, \dots, n_o - 1\}$  tenemos que  $V_j = A_j \cap W_j$  con  $W_j$  abierto en  $Y$ . Sea  $W = \bigcap_{j=1}^{n_o-1} W_j$ . Note que  $W$  es abierto en  $Y$  y  $x \in W$ . Veamos que  $g(W) \subset U$ .

$$g(W) = g(W \cap \bigcup_{n \in \mathbb{N}} A_n)$$

$$\begin{aligned}
&= g\left([W \cap (\bigcup_{n < n_o} A_n)] \cup [W \cap (\bigcup_{n \geq n_o} A_n)]\right) \\
&= g(W \cap (\bigcup_{n < n_o} A_n)) \cup g(W \cap (\bigcup_{n \geq n_o} A_n)).
\end{aligned}$$

Tenemos que

$$\begin{aligned}
g(W \cap (\bigcup_{n < n_o} A_n)) &= g(\bigcup_{n < n_o} (W \cap A_n)) \\
&\subset g(\bigcup_{n < n_o} (W_n \cap A_n)) = g(\bigcup_{n < n_o} V_n) \\
&= \bigcup_{n < n_o} g(V_n) = \bigcup_{n < n_o} f_n(V_n) \\
&\subset \bigcup_{n < n_o} (U \cap B_n) \subset U.
\end{aligned}$$

Y también

$$\begin{aligned}
g(W \cap (\bigcup_{n \geq n_o} A_n)) &\subset g(\bigcup_{n \geq n_o} A_n) \\
&= \bigcup_{n \geq n_o} g(A_n) = \bigcup_{n \geq n_o} f_n(A_n) \\
&= \bigcup_{n \geq n_o} B_n \subset U.
\end{aligned}$$

Así,  $g(W) \subset U$ . Por lo que  $g$  es continua en  $x$ . Luego,  $g$  es continua en  $Y$ .

Como,

$$g(Y) = g(\bigcup_{n \in \mathbb{N}} A_n) = \bigcup_{n \in \mathbb{N}} g(A_n) = \bigcup_{n \in \mathbb{N}} f_n(A_n) = \bigcup_{n \in \mathbb{N}} B_n = F_\omega.$$

Tenemos que  $g$  es suprayectiva.

Veamos que  $g$  es inyectiva. Sean  $p, q \in X$  con  $p \neq q$ . Consideremos los dos casos posibles:

- (1) Si existe  $n \in \mathbb{N}$  tal que  $p, q \in A_n$ . Dado que  $g$  es una extensión de  $f_n$ , tenemos que  $g(p) = f_n(p)$  y  $g(q) = f_n(q)$ . Como  $f_n$  es inyectiva,  $f_n(p) \neq f_n(q)$ , así  $g(p) \neq g(q)$ .
- (2) Si  $p \in A_n$  y  $q \in A_m$  con  $n \neq m$ . Dado que  $g$  es una extensión de  $f_n$  y  $f_m$ ,  $f_n(p) = g(p)$  y  $f_m(q) = g(q)$ . Se cumple que  $f_n(p) \in B_n$  y  $f_m(q) \in B_m$ ; así  $g(p) \neq g(q)$ .

Por lo tanto,  $g$  es inyectiva. Luego,  $g$  es continua y biyectiva. Dado que  $Y$  y  $F_\omega$  son continuos, concluimos que  $g$  es un homeomorfismo. Es decir,  $Y$  es un subcontinuo de  $X$  homeomorfo a  $F_\omega$ .  $\square$

Todo árbol es una dendrita, el siguiente resultado proporciona cuando una dendrita es un árbol en terminos de las dendritas  $F_\omega$  y  $W_R$ .

**3.3. TEOREMA.** [1, Teorema 4.11] Sea  $X$  una dendrita. Entonces  $X$  es un árbol si y sólo si  $X$  no contiene ningún subcontinuo homeomorfo a  $F_\omega$  ni a  $W_R$ .

**3.4. DEFINICIÓN.** Sea  $\mathfrak{D} = \{X \text{ dendrita} : E(X) \text{ es cerrado}\}$ .

Una manera sencilla para determinar si una dendrita pertenece a la familia  $\mathfrak{D}$  es la siguiente.

**3.5. TEOREMA.** Sea  $X$  una dendrita. Entonces  $X \in \mathfrak{D}$  si y sólo si  $X$  no contiene ningún subcontinuo homeomorfo a  $F_\omega$  ni a  $W$ .

3.6. TEOREMA. Sea  $X \in \mathfrak{D}$ . Si  $Y$  es un subcontinuo de  $X$ , entonces  $Y \in \mathfrak{D}$ .

DEMOSTRACIÓN. Por el Teorema 2.7,  $Y$  es una dendrita. Si  $Y$  es un árbol, por el Teorema 2.3,  $E(Y)$  es finito. Así,  $E(Y)$  es cerrado. Por lo tanto,  $Y \in \mathfrak{D}$ . Si  $Y$  no es un árbol, veamos que  $E(Y)' \subset E(Y)$ . Sea  $y \in E(Y)'$ , es decir,  $y \in Y$  y existe una sucesión  $\{y_n\}_{n=1}^{\infty}$  en  $E(Y)$  con  $y_n \neq y_m$  si  $n \neq m$  tal que  $\lim_{n \rightarrow \infty} y_n = y$ .

Si existe una subsucesión  $\{y_{n_k}\}_{k=1}^{\infty}$  de  $\{y_n\}_{n=1}^{\infty}$  en  $E(X)$ , entonces  $y \in E(X)$ , y como  $y \in Y$ , tenemos  $y \in E(Y)$ . Entonces podemos suponer que  $y_n \notin E(X)$  para cada  $n \in \mathbb{N}$ . por el Teorema 2.10, cada  $y_n$  es un punto de corte de  $X$ .

Fijemos  $n \in \mathbb{N}$ . Dado que  $y_n$  es un punto de no corte de  $Y$ ,  $Y - \{y_n\}$  es conexo. Así existe una componente  $C_n$  de  $X - \{y_n\}$  tal que  $(Y - \{y_n\}) \cap C_n = \emptyset$ .

Por el Teorema 2.9, se tiene que  $C_n \cap E(X) \neq \emptyset$ .

Para cada  $n \in \mathbb{N}$ , sea  $x_n \in C_n \cap E(X)$ . Tomemos  $m \in \mathbb{N}$  con  $m \neq n$ . Entonces  $y_n \neq y_m$ . Luego se tiene que,

$$\overline{C_n} \cap Y = (C_n \cup \{y_n\}) \cap Y = \{y_n\} \text{ y } \overline{C_m} \cap Y = (C_m \cup \{y_m\}) \cap Y = \{y_m\}.$$

Así,  $C_n \neq C_m$ . Luego,  $C_n \cap C_m = \emptyset$  si  $n \neq m$ . Así existe una sucesión  $\{x_n\}_{n=1}^{\infty}$  en  $E(X)$  con  $x_n \neq x_m$  si  $n \neq m$ . Por el Teorema 2.17, la colección  $\{C_n\}_{n=1}^{\infty} \subset X - Y$  es una sucesión nula; de aquí que la sucesión  $\{\overline{C_n} = C_n \cup \{y_n\}\}_{n=1}^{\infty}$  también es una sucesión nula ( pues el diámetro de un conjunto coincide con el diámetro de su cerradura ). Luego,  $\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} y_n = y$ , y dado que  $E(X)$  es cerrado, tenemos  $y \in E(X)$ , por lo que  $y \in E(Y)$ . Concluimos que,  $Y \in \mathfrak{D}$ .  $\square$

3.7. PROPOSICIÓN. Si  $X \in \mathfrak{D}$  y  $\{r_n\}_{n=1}^{\infty}$  es una sucesión convergente de puntos distintos de ramificación, entonces el  $\lim_{n \rightarrow \infty} r_n$  es un punto extremo. Es decir,  $R(X)' \subset E(X)$ .

3.8. COROLARIO. Si  $X \in \mathfrak{D}$ , entonces  $\overline{R(X)} \subset R(X) \cup E(X)$ .

DEMOSTRACIÓN. Sea  $x \in \overline{R(X)}$ . Dado que  $\overline{R(X)} = R(X) \cup R(X)'$  y por la Proposición 3.7 tenemos que,  $\overline{R(X)} \subset R(X) \cup E(X)$ .  $\square$

3.9. COROLARIO. Si  $X \in \mathfrak{D}$ , entonces  $R(X)$  es discreto.

DEMOSTRACIÓN. Sea  $x \in R(X)$ . Por la Proposición 3.7,  $x \notin R(X)'$ . Esto implica que existe un subconjunto  $U$  abierto en  $X$  con  $x \in U$  tal que  $(U \setminus \{x\}) \cap R(X) = \emptyset$ . Es decir,  $U \cap R(X) = \{x\}$ . Así,  $\{x\}$  es abierto en  $R(X)$ . Concluimos que,  $R(X)$  es discreto.  $\square$

3.10. PROPOSICIÓN. Sean  $X \in \mathfrak{D}$  y  $\{e_n\}_{n=1}^{\infty}$  una sucesión en  $E(X)$  convergente tal que  $e_n \neq e_m$  si  $n \neq m$  y  $\lim_{n \rightarrow \infty} e_n = e \neq e_1$ . Entonces existe una sucesión  $\{r_n\}_{n=1}^{\infty}$  en  $R(X) \cap [e, e_1]$  tal que  $\lim_{n \rightarrow \infty} r_n = e$ .

#### 4. CARACTERIZACIÓN DE $G^n$

En esta sección construimos para cada  $n \in \mathbb{N} - \{1, 2\}$ , una dendrita  $G^n$  y damos una caracterización de ella. Revisamos algunos resultados que nos serán de gran ayuda para demostrar dicha caracterización.

4.1. DEFINICIÓN. Sean  $X$  una dendrita y  $x, y \in X$ . El arco  $[x, y]$  es un *arco libre* si el conjunto  $(x, y)$  no tiene puntos de ramificación de  $X$ .

4.2. OBSERVACIÓN. Sea  $X$  una dendrita. El arco  $[x, y]$  es un arco libre si y sólo si el conjunto  $(x, y)$  es abierto en  $X$ .

4.3. DEFINICIÓN. Sea  $X$  una dendrita. El conjunto de puntos extremos aislados de  $X$  y el conjunto de puntos extremos de acumulación de  $X$ , respectivamente son:

$$E_I(X) = \{x \in E(X) : x \text{ es un punto aislado}\}.$$

$$E_A(X) = \{x \in E(X) : x \text{ es un punto de acumulación}\}.$$

4.4. OBSERVACIÓN. Sean  $X \in \mathfrak{D}$  y  $p, q \in X$ . Si  $p$  o  $q$  están en  $E_A(X)$ , entonces  $[p, q]$  no es un arco libre.

4.5. LEMA. [6, Lema 3.24] Si  $X \in \mathfrak{D}$ , entonces todo arco libre está contenido en un arco maximal en  $X$ .

4.6. LEMA. Si  $X \in \mathfrak{D}$ , entonces todo arco libre en  $X$  está contenido en un arco libre maximal en  $X$ .

DEMOSTRACIÓN. Sea  $[p, q]$  un arco libre en  $X$ . Si  $X$  es un arco, entonces  $X$  es el arco libre maximal que contiene a  $[p, q]$ .

Supongamos que  $X$  no es un arco. Por la Observación 4.4, tenemos que  $p$  y  $q$  están en  $E_I(X)$ . Si tenemos que:

- (1) Si  $p$  y  $q$  están en  $R(X)$ , entonces  $[p, q]$  es un arco libre maximal en  $X$  que contiene a  $[p, q]$ .
- (2) Si  $p \in R(X)$  (o  $q \in R(X)$ ) y  $q \in E_I(X)$  (o  $p \in E_I(X)$ ), entonces  $[p, q]$  es un arco libre maximal en  $X$  que contiene a  $[p, q]$ .

Ahora supongamos que  $q \in O(X)$ . Sean  $U$  y  $V$  las componentes de  $X \setminus \{q\}$  tal que  $[p, q] \subset V \cup \{q\}$ . En la dendrita  $(U \cup \{q\}) \cup [p, q]$  por el Lema 4.5, existe  $[p, s]$  un arco maximal que contiene a  $[p, q]$ ; además  $[p, s] = [p, q] \cup [q, s]$ .

Si  $[p, s]$  es un arco libre y  $p \in R(X)$  (o  $p \in E_I(X)$ ), entonces  $[p, s]$  es un arco libre maximal en  $X$  que contiene a  $[p, q]$ .

Si  $[p, s]$  no es un arco libre, entonces existen puntos de ramificación de  $X$  en  $[q, s]$ . Tenemos dos casos:

- (i) Si  $[q, s] \cap R(X) = \{r_1, \dots, r_m\}$  tal que  $r_1 \leq_q r_2 \leq_q \dots \leq_q r_m$ . Tomamos el arco  $[p, r_1]$ .
- (ii) Si  $[q, s] \cap R(X) = \{r_n : n \in \mathbb{N}\}$ . Entonces por la Proposición 3.7,  $\lim_{n \rightarrow \infty} r_n = s$  ( porque para toda  $x \in [q, s]$ ,  $x \in O(X) \cup R(X)$  ). Sea  $\epsilon > 0$ . Así, existe  $N \in \mathbb{N}$  tal que si  $n \geq N$ , entonces  $r_n \in B_\epsilon(s)$  y tal que  $r_1, \dots, r_{n-1} \notin B_\epsilon(s)$ . Supongamos que,  $r_1 \leq_q r_2 \leq_q \dots \leq_q r_{n-1}$ . Luego tomamos el arco  $[p, r_1]$ .

Observe que  $[p, r_1]$  es un arco libre que contiene a  $[p, q]$ .

Si  $p \in R(X)$  o  $p \in E_I(X)$ , por (1) y (2) el arco  $[p, r_1]$  es un arco libre maximal en  $X$  que contiene a  $[p, q]$ .

Si  $p \in O(X)$ . Procediendo como en el caso cuando  $q \in O(X)$ , obtenemos un punto  $r_2 \in E_I(X) \cup R(X)$  tal que  $[q, r_2]$  es un arco libre que contiene a  $[p, q]$ . Así  $[r_1, r_2]$  es un arco libre maximal en  $X$  que contiene a  $[p, q]$ .  $\square$

4.7. LEMA. Si  $X \in \mathfrak{D}$  y  $A$  es un arco libre maximal en  $X$ , entonces  $E(A) \subset E(X) \cup R(X)$ .

DEMOSTRACIÓN. Sea  $A$  un arco libre maximal en  $X$  con puntos extremos  $p$  y  $q$ . Supongamos que  $p \in O(X)$ . Existe una bola abierta  $B_\epsilon(p)$  tal que,

$$B_\epsilon(p) \subset O(X).$$

En caso contrario, como  $X = O(X) \cup E(X) \cup R(X)$  tendríamos una sucesión  $\{e_n\}_{n=1}^\infty \subset E(X)$  tal que  $\lim_{n \rightarrow \infty} e_n = p$ , así  $p \in E(X)$  lo cual contradice a que

$p \in O(X)$ ; o bien existiría una sucesión  $\{r_n\}_{n=1}^\infty \subset R(X)$  tal que  $\lim_{n \rightarrow \infty} r_n = p$  lo que contradice a la Proposición 3.7.

Dado que  $ord(p, X) = 2$ , existen  $U$  y  $V$  componentes de  $X - \{p\}$  tales que  $X - \{p\} = U \cup V$ . Supongamos que  $[p, q] \subset U \cup \{p\}$ . Sea  $t \in V \cap B_\varepsilon(p)$ . Note que el arco  $[t, q] = [t, p] \cup [p, q]$  y que es un arco libre, así  $[p, q] \subset [t, q]$  lo cual contradice a que  $[p, q]$  es un arco libre maximal en  $X$ . Concluimos que,  $p \in E(X) \cup R(X)$ .  $\square$

4.8. LEMA. [6, Lema 3.27] Sean  $X \in \mathfrak{D}$ ,  $x \in X$  y  $n \in \mathbb{N} - \{1, 2\}$ . Si  $ord(x, X) = n$ , entonces existen exactamente  $n$  arcos libres maximales en  $X$  cuyo punto extremo en común es  $x$ .

A continuación para cada  $n \in \mathbb{N} - \{1, 2\}$  definimos una dendrita  $G^n$  como sigue.

Sean  $i \in \mathbb{N}$  y  $\alpha_1, \dots, \alpha_i \in \{0, 2, \dots, 2n - 4\}$ , definimos  $E^n_{\alpha_1, \dots, \alpha_i}$  como el conjunto de los números  $x$  en  $[0, 1]$  tal que  $x$  se expresa en base  $2n - 3$ , entonces los primeros  $i$  dígitos son exactamente  $\alpha_1, \dots, \alpha_i$ , es decir,

$$E^n_{\alpha_1, \dots, \alpha_i} = \left[ \frac{\alpha_1}{2n-3} + \dots + \frac{\alpha_i}{(2n-3)^i}, \frac{\alpha_1}{2n-3} + \dots + \frac{\alpha_{i-1}}{(2n-3)^{i-1}} + \frac{\alpha_i+1}{(2n-3)^i} \right].$$

Si todos los dígitos  $\alpha_1, \dots, \alpha_i$  son pares, sea  $p^n_{\alpha_1, \dots, \alpha_i}$  el punto en el plano cuya primera coordenada es la mitad del intervalo  $E^n_{\alpha_1, \dots, \alpha_i}$  y la segunda coordenada es  $\frac{1}{2^i}$ . Si  $i = 0$ , entonces  $E^n = [0, 1]$  y  $p^n = (\frac{1}{2}, 1)$ . Sean  $q = (\frac{1}{2}, 2)$  y  $H^n$  definido por:

$$H^n = [q, p^n] \cup \overline{\bigcup_{i \in \mathbb{N}, \alpha_1, \dots, \alpha_i \in \{0, 2, \dots, 2n-4\}} [p^n_{\alpha_1, \dots, \alpha_{i-1}}, p^n_{\alpha_1, \dots, \alpha_{i-1}, \alpha_i}]}$$

Dada una sucesión  $\alpha_1, \alpha_2 \dots$  la intersección  $\bigcap_{i \in \mathbb{N}} E^n_{\alpha_1, \dots, \alpha_i}$  es un conjunto de un sólo punto. Denotemos el punto por  $e_{\alpha_1, \alpha_2, \dots}$ , sea  $p_{\alpha_1, \alpha_2, \dots} = (e_{\alpha_1, \alpha_2, \dots}, 0) \in H^n$ , y notemos que

$$E(H^n) = \{q\} \cup \{p_{\alpha_1, \alpha_2, \dots} : \alpha_i \in \{0, 2, \dots, 2n - 4\}\}.$$

Luego, definimos  $G^n$  como la unión de dos subcontinuos homeomorfos  $H^n$  y  $(H^n)^*$  de  $H^n$  con los puntos correspondientes  $q$  y  $q^*$  identificados.

En la siguiente observación proporcionamos otra forma de escribir a la dendrita  $H^n$ , la cual usaremos en el Teorema 4.12.

4.9. OBSERVACIÓN. Si  $T_m = [q, p^n] \cup \left( \bigcup \{ [p^n_{\alpha_1, \dots, \alpha_{i-1}}, p^n_{\alpha_1, \dots, \alpha_{i-1}, \alpha_i}] : i \leq m \text{ y } \alpha_1, \dots, \alpha_i \in \{0, 2, \dots, 2n - 4\} \} \right)$  para cada  $m \in \mathbb{N}$ , entonces  $\bigcup_{m \in \mathbb{N}} T_m = H^n$ .

En seguida enunciamos algunas propiedades de la dendrita  $H^n$ .

4.10. LEMA. [6, Lema 3.28] Sea  $n \in \mathbb{N} - \{1, 2\}$  fijo. Para cada  $x \in H^n$ , el siguiente conjunto

$$V(x) = \{y \in H^n : x <_q y\} \text{ es abierto en } H^n.$$

4.11. LEMA. [6, Lema 3.29] Sea  $p_{\alpha_1, \alpha_2, \dots}$  un punto extremo de  $H^n$ , entonces una base local de vecindades de  $p_{\alpha_1, \alpha_2, \dots}$  es la siguiente  $\{V(p_{\alpha_1, \dots, \alpha_k})\}_{k=1}^\infty$ .

Para concluir el artículo presentamos una caracterización de la dendrita  $G^n$ .

4.12. TEOREMA. Sean  $X$  una dendrita y  $n \in \mathbb{N} - \{2, 3\}$  fijo. Entonces  $X$  es homeomorfo a  $G^n$  si y sólo si el conjunto  $E(X)$  es homeomorfo al conjunto de Cantor y  $R(X) = R_n(X)$ .

DEMOSTRACIÓN. Sea  $X$  una dendrita homeomorfa a  $G^n$ . Claramente,  $R(X) = R_n(X)$ . Como  $A = E(H^n) - \{q\}$  es homeomorfo al conjunto de Cantor, también lo es  $B = E(H^n) - \{q\}$ , y como  $E(G^n) = A \cup B$ , y  $A$  y  $B$  son conjuntos separados  $E(G^n)$  es homeomorfo al conjunto de Cantor. Así,  $E(X)$  es homeomorfo al conjunto de Cantor.

Construiremos un homeomorfismo  $h : G^n \rightarrow X$ . Sea  $x \in R(X)$ . Por el Corolario 3.9, existe un abierto  $U$  de  $X$  tal que  $x \in U$  y  $U \cap R(X) = \{x\}$ . Sea  $V$  un abierto y conexo tal que  $x \in V \subset \bar{V} \subset U$ . Nótese que  $\bar{V}$  es una dendrita no degenerada. Sea  $A$  un arco en  $\bar{V} - \{x\}$ . Se tiene que  $A \cap R(X) = \emptyset$ . En particular,  $A$  es un arco libre. Tomemos  $a \in A$  y hagamos  $h(q) = a$ . Por el Lema 4.6, existe un arco libre maximal  $[b, b']$  tal que  $A \subset [b, b']$ . Por el Lema 4.7,  $b$  no es un punto ordinario de  $X$ . Supongamos que  $b$  es un punto extremo de  $X$ . Entonces, ya que  $E(X)$  es homeomorfo al conjunto de Cantor, existe una sucesión  $\{e_n\}_{n=1}^\infty \subset E(X) - \{b\}$  tal que  $\lim_{n \rightarrow \infty} e_n = b$ . Por la Proposición 3.10,  $[b, b']$  contiene una sucesión de puntos de ramificación que convergen a  $b$ . En particular,  $[b, b']$  no es un arco libre. Esto es una contradicción. Luego  $b$  no es un punto extremo de  $X$ . Por lo tanto,  $b \in R(X)$ .

En general, cualquier punto extremo de cualquier arco libre maximal es un punto de ramificación de  $X$ . Hagamos  $h(p^n) = b$  y sea  $h : [q, p^n] \rightarrow [a, b]$  un homeomorfismo.

Procederemos por inducción sobre el conjunto de índices. Para cada  $m \in \mathbb{N}$ , definimos  $H_m^n$  de forma recurrente como sigue.

$$H_1^n = [q, p^n] \cup \left( \bigcup_{\alpha_1 \in \{0, 2, \dots, 2n-4\}} [p^n, p_{\alpha_1}^n] \right),$$

$$H_2^n = H_1^n \cup \left( \bigcup_{\alpha_1, \alpha_2 \in \{0, 2, \dots, 2n-4\}} [p_{\alpha_1}^n, p_{\alpha_2}^n] \right),$$

$$H_m^n = H_{m-1}^n \cup \left( \bigcup_{\alpha_1, \dots, \alpha_{m-1}, \alpha_m \in \{0, 2, \dots, 2n-4\}} [p_{\alpha_1, \dots, \alpha_{m-1}}^n, p_{\alpha_1, \dots, \alpha_{m-1}, \alpha_m}^n] \right).$$

Observe que  $H_m^n \subset H^n$  y  $H_m^n \subset H_{m+1}^n$ . Veamos que  $H^n = \overline{\bigcup_{m \in \mathbb{N}} H_m^n}$ . Para ello, probemos por inducción que  $H_m^n = T_m$ . Si  $m = 1$ , es claro de la definición que  $H_1^n = T_1$ .

Supongamos que,  $H_k^n = T_k$ . Veamos que se cumple para  $k+1$ . Tenemos que,

$$H_{k+1}^n = H_k^n \cup \left( \bigcup_{\alpha_1, \dots, \alpha_{k+1} \in \{0, 2, \dots, 2n-4\}} [p_{\alpha_1, \dots, \alpha_k}^n, p_{\alpha_1, \dots, \alpha_{k+1}}^n] \right) =$$

$$T_k \cup \left( \bigcup_{\alpha_1, \dots, \alpha_{k+1} \in \{0, 2, \dots, 2n-4\}} [p_{\alpha_1, \dots, \alpha_k}^n, p_{\alpha_1, \dots, \alpha_{k+1}}^n] \right) =$$

$$[q, p^n] \cup \left( \bigcup \{ [p_{\alpha_1, \dots, \alpha_{i-1}}^n, p_{\alpha_1, \dots, \alpha_{i-1}, \alpha_i}^n] : i \leq k \text{ y } \alpha_1, \dots, \alpha_i \in \{0, 2, \dots, 2n-4\} \} \right) \cup$$

$$\left( \bigcup_{\alpha_1, \dots, \alpha_{k+1} \in \{0, 2, \dots, 2n-4\}} [p_{\alpha_1, \dots, \alpha_k}^n, p_{\alpha_1, \dots, \alpha_{k+1}}^n] \right) =$$

$$[q, p^n] \cup \left( \bigcup \{ [p_{\alpha_1, \dots, \alpha_{i-1}}^n, p_{\alpha_1, \dots, \alpha_{i-1}, \alpha_i}^n] : i \leq k+1 \text{ y } \alpha_1, \dots, \alpha_i \in \{0, 2, \dots, 2n-4\} \} \right) =$$

$$T_{k+1}.$$

Así,  $H_m^n = T_m$  para cada  $m \in \mathbb{N}$ ; luego  $\bigcup_{m \in \mathbb{N}} H_m^n = \bigcup_{m \in \mathbb{N}} T_m$ .

Por  $\bigcup_{m \in \mathbb{N}} H_m^n = \bigcup_{m \in \mathbb{N}} T_m$  y la Observación 4.9 se tiene que  $\bigcup_{m \in \mathbb{N}} H_m^n = H^n$ ; luego  $H^n = \overline{\bigcup_{m \in \mathbb{N}} H_m^n}$ .

Supongamos que  $h$  está definida en  $H_m^n$ . Ahora definiremos  $h$  en  $H_{m+1}^n$ . Sean  $\alpha_1, \dots, \alpha_m$ . Por construcción,  $h(p_{\alpha_1, \dots, \alpha_m}^n) \in R(X)$ , así  $\text{ord}(h(p_{\alpha_1, \dots, \alpha_m}^n)) = n$ . Así existen  $n-1$  arcos libres maximales que no contienen a  $h(p_{\alpha_1, \dots, \alpha_{m-1}}^n)$ . Denotemos

por  $b_{\alpha_1, \dots, \alpha_m, 0}, b_{\alpha_1, \dots, \alpha_m, 2}, \dots, b_{\alpha_1, \dots, \alpha_m, 2n-4}$  los otros puntos extremos de dichos arcos, sea  $h(p_{\alpha_1, \dots, \alpha_{m+1}}^n) = b_{\alpha_1, \dots, \alpha_{m+1}}$  para cada  $\alpha_{m+1} \in \{0, 2, \dots, 2n-4\}$  y sea  $h : [p_{\alpha_1, \dots, \alpha_m}^n, p_{\alpha_1, \dots, \alpha_{m+1}}^n] \rightarrow [b_{\alpha_1, \dots, \alpha_m}, b_{\alpha_1, \dots, \alpha_{m+1}}]$  un homeomorfismo. De esta forma hemos definido  $h$  en  $\bigcup_{m \in \mathbb{N}} H_m^n$ .

Veamos que  $h|_{\bigcup_{m \in \mathbb{N}} H_m^n}$  es continua e inyectiva. Como  $h : [q, p^n] \rightarrow [a, b]$ ,  $h : [p_{\alpha_1, \dots, \alpha_{m-1}}^n, p_{\alpha_1, \dots, \alpha_m}^n] \rightarrow [b_{\alpha_1, \dots, \alpha_{m-1}}, b_{\alpha_1, \dots, \alpha_m}]$  para  $m \in \mathbb{N}$  son homeomorfismos, en particular son continuas, y además la familia

$$[q, p^n], \{[p_{\alpha_1, \dots, \alpha_{i-1}}^n, p_{\alpha_1, \dots, \alpha_{i-1}, \alpha_i}^n] : i \in \mathbb{N} \text{ y } \alpha_1, \dots, \alpha_i \in \{0, 2, \dots, 2n-4\}\}$$

es una familia de vecindades finita.

Luego por el Teorema 2.19 tenemos que  $h|_{\bigcup_{m \in \mathbb{N}} H_m^n}$  es continua.

Ahora mostremos que  $h|_{\bigcup_{m \in \mathbb{N}} H_m^n}$  es inyectiva. Sean  $x, y \in \bigcup_{m \in \mathbb{N}} H_m^n$  con  $x \neq y$ .

Como  $\{H_m^n\}_{m \in \mathbb{N}}$  es una sucesión decreciente existe  $m_0 \in \mathbb{N}$  tal que  $x, y \in H_{m_0}^n$ . Si  $x$  y  $y$  están en el mismo arco  $[p_{\alpha_1, \dots, \alpha_{i-1}}^n, p_{\alpha_1, \dots, \alpha_i}^n]$  con  $i \leq m_0$  y  $\alpha_1, \dots, \alpha_i \in \{0, 2, \dots, 2n-4\}$ , entonces como  $h|_{[p_{\alpha_1, \dots, \alpha_{i-1}}^n, p_{\alpha_1, \dots, \alpha_i}^n]}$  es un homeomorfismo, se tiene que  $h|_{[p_{\alpha_1, \dots, \alpha_{i-1}}^n, p_{\alpha_1, \dots, \alpha_i}^n]}(x) \neq h|_{[p_{\alpha_1, \dots, \alpha_{i-1}}^n, p_{\alpha_1, \dots, \alpha_i}^n]}(y)$ , así  $h|_{\bigcup_{m \in \mathbb{N}} H_m^n}(x) \neq h|_{\bigcup_{m \in \mathbb{N}} H_m^n}(y)$ .

Ahora bien, si  $x$  y  $y$  están en arcos distintos de la forma  $[p_{\alpha_1, \dots, \alpha_{i-1}}^n, p_{\alpha_1, \dots, \alpha_i}^n]$ . Por construcción de  $h$ ,  $h|_{\bigcup_{m \in \mathbb{N}} H_m^n}(x)$  y  $h|_{\bigcup_{m \in \mathbb{N}} H_m^n}(y)$  están en arcos distintos, así  $h|_{\bigcup_{m \in \mathbb{N}} H_m^n}(x) \neq h|_{\bigcup_{m \in \mathbb{N}} H_m^n}(y)$ . En ambos casos se tiene que,  $h|_{\bigcup_{m \in \mathbb{N}} H_m^n}$  es inyectiva.

Veamos que  $E(H^n) = (H^n - \bigcup_{m \in \mathbb{N}} H_m^n) \cup \{q\}$ . Como  $(\bigcup_{m \in \mathbb{N}} H_m^n) - \{q\} \subset O(H^n) \cup R(H^n)$ , tenemos que  $E(H^n) \subset (H^n - \bigcup_{m \in \mathbb{N}} H_m^n) \cup \{q\}$ .

Si  $x \in H^n \setminus \bigcup_{m \in \mathbb{N}} H_m^n$  con  $x \neq q$ , entonces existe una sucesión  $\{x_k\}_{k=1}^\infty \subset \bigcup_{m \in \mathbb{N}} H_m^n$  tal que  $\lim_{k \rightarrow \infty} x_k = x$  y  $x \notin \bigcup_{m \in \mathbb{N}} H_m^n$ ; podemos suponer que  $x_i \neq x_j$ , si  $i \neq j$ . Note que no puede ocurrir que para cada  $k \in \mathbb{N}$  y algún  $m \in \mathbb{N}$ ,  $x_k \in H_m^n$ . Si fuera el caso tendríamos que  $x \in H_m^n$ , lo cual no es posible. Luego podemos elegir una subsucesión  $\{x_{k_m}\}_{m=1}^\infty$  de la sucesión  $\{x_k\}_{k=1}^\infty$  tal que para cada  $m \in \mathbb{N}$ ,  $x_{k_m} \in H_{k_m}^n - H_{k_m-1}^n$ . Sea  $\epsilon > 0$ . Dado que  $\lim_{m \rightarrow \infty} \text{diám}([p_{\alpha_1, \dots, \alpha_m}^n, p_{\alpha_1, \dots, \alpha_{m+1}}^n]) = 0$ , existe  $N \in \mathbb{N}$  tal que si  $m \geq N$ ,  $\text{diám}([p_{\alpha_1, \dots, \alpha_m}^n, p_{\alpha_1, \dots, \alpha_{m+1}}^n]) < \frac{\epsilon}{2}$  con  $\alpha_1, \dots, \alpha_{m+1} \in \{0, 2, \dots, 2n-4\}$ . Puesto que  $x_{k_m} \in H_{k_m}^n - H_{k_m-1}^n$ , existe  $[p_{\alpha_1, \dots, \alpha_m}^n, p_{\alpha_1, \dots, \alpha_{m+1}}^n] \subset H_{k_m}^n - H_{k_m-1}^n$  tal que  $x_{k_m} \in [p_{\alpha_1, \dots, \alpha_m}^n, p_{\alpha_1, \dots, \alpha_{m+1}}^n]$ ; así  $d(p_{\alpha_1, \dots, \alpha_m}^n, x_{k_m}) < \frac{\epsilon}{2}$ . Como  $\lim_{m \rightarrow \infty} x_{k_m} = x$ , podemos suponer que si  $m \geq N$ ,  $d(x_{k_m}, x) < \frac{\epsilon}{2}$ .

Luego si  $m \geq N$ ,  $d(p_{\alpha_1, \dots, \alpha_m}^n, x) \leq d(p_{\alpha_1, \dots, \alpha_m}^n, x_{k_m}) + d(x_{k_m}, x) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$ . Así  $\lim_{m \rightarrow \infty} p_{\alpha_1, \dots, \alpha_m}^n = x$ ; luego por la Proposición,  $x \in E(H^n)$ . Por lo que,

$$E(H^n) = (H^n - \bigcup_{m \in \mathbb{N}} H_m^n) \cup \{q\}.$$

Veamos que la sucesión  $\{b_{\alpha_1, \dots, \alpha_k}\}_{k=1}^\infty$  es convergente. Como  $\{b_{\alpha_1, \dots, \alpha_k}\}_{k=1}^\infty \subset \overline{R(X)}$  y  $\overline{R(X)}$  es compacto, existe una subsucesión  $\{b_{\alpha_{k_1}, \dots, \alpha_{k_m}}\}_{m=1}^\infty$  tal que

$$\lim_{m \rightarrow \infty} b_{\alpha_{k_1}, \dots, \alpha_{k_m}} = t \text{ para alg\u00fan } t \in X.$$

Sea  $\epsilon > 0$ . As\u00ed, existe  $N \in \mathbb{N}$  tal que si  $m \geq N$ ,  $d(b_{\alpha_{k_1}, \dots, \alpha_{k_m}}, t) < \epsilon$ . Tomemos terminos de la sucesi\u00f3n  $\{b_{\alpha_1, \dots, \alpha_k}\}_{k=1}^\infty$  tales que  $b_{\alpha_{k_1}, \dots, \alpha_{k_m}} \leq_a b_{\alpha_1, \dots, \alpha_k}$  con  $k_m \leq k$ . Como la sucesi\u00f3n  $\{b_{\alpha_1, \dots, \alpha_k}\}_{k=1}^\infty$  es creciente, existe  $b_{\alpha_{k_1}, \dots, \alpha_{k_s}}$  tal que  $b_{\alpha_1, \dots, \alpha_k} \leq_a b_{\alpha_{k_1}, \dots, \alpha_{k_s}}$  con  $k \leq k_s$ . Luego existen arcos  $[b_{\alpha_{k_1}, \dots, \alpha_{k_m}}, b_{\alpha_1, \dots, \alpha_k}]$ ,  $[b_{\alpha_1, \dots, \alpha_k}, b_{\alpha_{k_1}, \dots, \alpha_{k_s}}]$  y  $[b_{\alpha_{k_1}, \dots, \alpha_{k_m}}, t]$ . Luego  $b_{\alpha_1, \dots, \alpha_k} \in [b_{\alpha_{k_1}, \dots, \alpha_{k_m}}, t]$ . Dado que  $X$  es suave en  $t$ , se tiene que  $\lim_{m \rightarrow \infty} [b_{\alpha_{k_1}, \dots, \alpha_{k_m}}, t] = \{t\}$ ; as\u00ed podemos suponer que si  $m \geq N$ , entonces  $H_d([b_{\alpha_{k_1}, \dots, \alpha_{k_m}}, t], \{t\}) < \epsilon$ . Es decir,  $[b_{\alpha_{k_1}, \dots, \alpha_{k_m}}, t] \subset B_\epsilon(t)$  pues  $N(t, \epsilon) = B_\epsilon(t)$ . Por lo tanto,  $d(b_{\alpha_1, \dots, \alpha_k}, t) < \epsilon$ . Es decir, la sucesi\u00f3n  $\{b_{\alpha_1, \dots, \alpha_k}\}_{k=1}^\infty$  es convergente. Definimos  $h(p_{\alpha_1, \alpha_2, \dots}) = \lim_{k \rightarrow \infty} b_{\alpha_1, \dots, \alpha_k}$ . Por la Proposici\u00f3n 3.7  $h(p_{\alpha_1, \alpha_2, \dots})$  es un punto extremo de  $X$ .

Veamos que la funci\u00f3n  $h$  es inyectiva. Sean  $x, y \in H^n$  con  $x \neq y$ . Para ello, consideramos los siguientes casos.

- (1) Si  $x, y \in \bigcup_{m \in \mathbb{N}} H_m^n$ . En este caso, unos p\u00e1rrafos arriba ya se mostr\u00f3 que  $h(x) \neq h(y)$ .
- (2) Si  $x \in \bigcup_{m \in \mathbb{N}} H_m^n$  y  $y \in E(X)$ . Luego  $x \in [p_{\alpha_1, \dots, \alpha_{i-1}}, p_{\alpha_1, \dots, \alpha_i}^n]$  para alg\u00fan  $i \in \mathbb{N}$ ; as\u00ed  $h(x) \in [b_{\alpha_1, \dots, \alpha_{i-1}}, p_{\alpha_1, \dots, \alpha_i}^n]$ . Por lo que,  $h(x) \notin E(X)$ . Y dado que  $h(y) \in E(X)$ , concluimos que  $h(x) \neq h(y)$ .
- (3) Si  $x, y \in E(H^n)$ . Sean  $\{\alpha_i\}_{i=1}^\infty$  y  $\{\beta_i\}_{i=1}^\infty$  las sucesiones que definen los puntos  $x$  y  $y$  respectivamente. Es decir,

$$\begin{aligned} x &= p_{\alpha_1, \dots, \alpha_i, \dots} \\ y &= p_{\beta_1, \dots, \beta_i, \dots} \end{aligned}$$

Como  $x \neq y$ , existe  $i + 1 \in \mathbb{N}$  tal que  $\alpha_{i+1} \neq \beta_{i+1}$ . Como  $H^n$  no contiene curvas cerradas simples,  $[p_{\alpha_1, \dots, \alpha_i}, x] \cap [p_{\alpha_1, \dots, \alpha_i}, y] = \{p_{\alpha_1, \dots, \alpha_i}\}$ .

Dado que,

$$\begin{aligned} h(p_{\alpha_1, \dots, \alpha_i, \alpha_{i+1}, \dots}) &= h(x) = \lim_{i \rightarrow \infty} b_{\alpha_1, \dots, \alpha_{i+1}} \\ h(p_{\alpha_1, \dots, \alpha_i, \beta_{i+1}, \dots}) &= h(y) = \lim_{i \rightarrow \infty} b_{\alpha_1, \dots, \beta_{i+1}} \end{aligned}$$

$$\text{y } h(p_{\alpha_1, \dots, \alpha_i}) = b_{\alpha_1, \dots, \alpha_i}.$$

Luego como  $X$  es una dendrita,  $[b_{\alpha_1, \dots, \alpha_i}, h(x)] \cap [b_{\alpha_1, \dots, \alpha_i}, h(y)] = \{b_{\alpha_1, \dots, \alpha_i}\}$ . Por lo que,  $h(x) \neq h(y)$ . Concluimos que,  $h$  es inyectiva.

Veamos que  $h$  es continua. Como  $h$  es continua en  $h|_{\bigcup_{m \in \mathbb{N}} H_m^n}$ , s\u00f3lo falta ver que  $h$  sea continua en  $E(H^n)$ . Usando el Lema 4.11 y  $h^{-1}(U(b_{\alpha_1, \dots, \alpha_k})) = V(p_{\alpha_1, \dots, \alpha_k})$ , obtenemos la continuidad de  $h$  en  $E(H^n)$ . As\u00ed  $h$  es continua en  $H^n$ . An\u00e1logamente podemos definir  $h$  en  $(H^n)^*$ . Luego por el Teorema 2.19, tenemos que la funci\u00f3n  $h : G^n \rightarrow X$  es continua.

Sea  $x \in X$ . Dado que  $[a, x] \subset X$  y  $ord(a, X) = 2$ , tenemos que  $[a, x] \subset h(H^n)$  o  $[a, x] \subset h((H^n)^*)$ . As\u00ed,  $h : G^n \rightarrow X$  es suprayectiva. Por lo tanto,  $h$  es continua y biyectiva. Luego, dado que  $G^n$  y  $X$  son continuos concluimos que la funci\u00f3n  $h : G^n \rightarrow X$  es un homeomorfismo.  $\square$

### REFERENCIAS

- [1] Daniel Ar\u00e9valo, Wlodzimierz J. Charatonik, Patricia Pellicer Covarrubias y Likin Sim\u00f3n, *Dendrites with a closed set of end points*, Topology Appl. **115** (2001), no. 1, 1-17.
- [2] S. B. Nadler, Jr., *Continuum Theory: An introduction*, Monographs and Textbooks in Pure and Applied Math., Vol. 158, Marcel Dekker, Inc., New York, 1992.

- [3] G. T. Whyburn, *Analytic Topology*, Amer. Math. Soc. Colloq. Publ., vol. 28, Providence, RI, 1942, reprinted with corrections 1971.
- [4] James Dugundji, *Topology*, 2nd ed., BCS Associates, Moscow, Idaho, USA, 1998.
- [5] Luis Alberto, *Clases de Continuos Localmente Conexos*, Tesis de Licenciatura, FCFM, BUAP, 2009.
- [6] Francisco Vázquez, *Propiedades de las Dendritas Cuyo Conjunto de Puntos Extremos es Cerrado*, Tesis de Maestría, por presentar, FCFM, BUAP.

Facultad de Ciencias Físico Matemáticas, BUAP.

Av. San Claudio y 18 Sur, Col. San Manuel, Ciudad Universitaria

Puebla, Pue., C.P. 72570

dherrer@fcfm.buap.mx

fmacias@fcfm.buap.mx

paco2013@hotmail.com

# ALGUNAS PROPIEDADES TOPOLÓGICAS Y LA FUNCIÓN $T$ DE JONES

SERGIO MACÍAS  
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

RESUMEN. Un *continuo* es un espacio métrico, compacto y conexo. Dados un espacio métrico y compacto  $X$  y un subconjunto  $A$  de  $X$ , definimos  $T(A) = \{x \in X \mid \text{para todo subcontinuo } W \text{ de } X, \text{ tal que } x \in \text{Int}_X(W), \text{ se tiene que } W \cap A \neq \emptyset\}$ . Presentaremos los conceptos: casi conexo en pequeño, conexo en pequeño, aposindesis, semiconexidad local y conexidad local; en términos de la función  $T$ .

## 1. INTRODUCCIÓN

La función  $T$  originalmente fue definida en 1948 por el Profesor F. Burton Jones [7] como una herramienta en el estudio de los continuos aposindéticos (“apo” quiere decir “lejos”, “sin” significa “junto” y “deo” quiere decir “envolver”). Posteriormente, el Profesor Jones utilizó esta función para estudiar a los continuos homogéneos y probó un teorema de descomposición [8] que ha sido muy importante para el entendimiento de dichos continuos. Desde entonces muchas propiedades relacionadas con esta función han sido estudiadas. Por ejemplo, en [1] se dan propiedades que satisfacen los continuos para los cuales la función  $T$  es continua. En [11], [13] y [15] se presentan varias clases de continuos descomponibles, no localmente conexos y de dimensión uno para las cuales la función  $T$  es continua. En [14] se demuestra un teorema de descomposición para cierta clase de continuos para los cuales la función  $T$  es continua. En [13] se caracteriza a la clase de los continuos homogéneos para los cuales la función  $T$  es continua. Como algunos ejemplos de las aplicaciones que puede tener la función  $T$ , mencionaremos los siguientes: Se ha utilizado a la función  $T$  para estudiar la contractibilidad de continuos [2]; también se ha hecho uso de la función  $T$  para estudiar a los continuos que pueden ser mandados de manera continua y suprayectiva sobre sus conos [3] y para estudiar a los productos simétricos de continuos [10].

Las propiedades topológicas que veremos son: la casi conexidad en pequeño, la conexidad en pequeño, la aposindesis, la semiconexidad local y la conexidad local. El concepto de casi conexidad en pequeño fue introducido por los Profesores H. S. Davis y P. H. Doyle, para estudiar continuos invertibles [4]. Un hecho importante es que si un continuo es casi conexo en pequeño en cada uno de sus puntos y la función  $T$  es continua entonces tal continuo es localmente conexo [14, 3.2.20]. Además, si un continuo homogéneo es casi conexo en pequeño en alguno de sus puntos entonces ese continuo es localmente conexo [16, 3.5]. Desconocemos el origen del concepto de conexidad en pequeño, lo que sí sabemos es que en algunas partes de la literatura se le da el nombre de conexidad local, como es el caso del libro del Profesor K. Kuratowski [9]. El concepto de aposindesis, como ya mencionamos, fue introducido por el Profesor F. B. Jones como una generalización de la conexidad en pequeño

y como una manera alternativa a la semiconexidad local para el estudio de los continuos en el plano [8]. El Profesor G. T. Whyburn introdujo el concepto de semiconexidad local como una generalización de la conexidad local y también lo usó para estudiar continuos en el plano [18].

## 2. PROPIEDADES TOPOLÓGICAS

Empezaremos dando un poco de notación.

2.1. NOTACIÓN. Si  $X$  es un espacio métrico y  $A$  es un subconjunto de  $X$  entonces el interior, la cerradura y la frontera de  $A$ , con respecto a  $X$ , serán denotados por  $Int_X(A)$ ,  $Cl_X(A)$  y  $Fr_X(A)$ , respectivamente. Si  $\varepsilon > 0$  entonces la bola abierta de radio  $\varepsilon$  y centro en  $A$  se denota como:  $V_\varepsilon(A)$ .

2.2. NOTACIÓN. Dado un espacio métrico  $X$ ,  $P(X)$ , denota al conjunto potencia de  $X$ ; i.e.:

$$P(X) = \{A \mid A \subset X\}.$$

2.3. OBSERVACIÓN. Sean  $X$  un espacio métrico y  $A \in P(X)$ . Recordemos que:

$$\begin{aligned} Cl_X(A) &= \{x \in X \mid \text{para todo subconjunto abierto } U \text{ de } X \\ &\quad \text{tal que } x \in U, \text{ se tiene que } U \cap A \neq \emptyset\} = \\ &= X \setminus \{x \in X \mid \text{existe un subconjunto abierto } U \text{ de } X \text{ tal que } x \in U \subset X \setminus A\}. \end{aligned}$$

2.4. DEFINICIÓN. Un *continuo* es un espacio métrico, compacto y conexo. Un *subcontinuo* es un continuo, con la topología relativa, el cual está contenido en un espacio.

La definición de la función  $T$  de Jones se expresa de manera parecida a la cerradura de un conjunto (Observación 2.3):

2.5. DEFINICIÓN. Sea  $X$  un espacio métrico y compacto. Definimos

$$T: P(X) \rightarrow P(X)$$

como

$$\begin{aligned} T(A) &= \{x \in X \mid \text{para todo subcontinuo } W \text{ de } X, \text{ tal que } x \in Int_X(W), \\ &\quad \text{se tiene que } W \cap A \neq \emptyset\} = \\ &= X \setminus \{x \in X \mid \text{existe un subcontinuo } W \text{ de } X \text{ tal que} \\ &\quad x \in Int_X(W) \subset W \subset X \setminus A\} \end{aligned}$$

para cada  $A \in P(X)$ . A esta función se le conoce como la *función  $T$  de Jones*.

2.6. OBSERVACIÓN. Es claro de la definición de la función  $T$  que  $A \subset T(A)$  para todo  $A$ .

2.7. LEMA. Si  $X$  es un espacio métrico y compacto y  $A$  y  $B \in P(X)$  entonces se cumple lo siguiente:

- (1)  $T(A)$  es cerrado en  $X$ .
- (2) Si  $A \subset B$  entonces  $T(A) \subset T(B)$ .
- (3)  $T(A) \cup T(B) \subset T(A \cup B)$ .

DEMOSTRACIÓN. Veamos que se cumple (1). Sea  $x \in X \setminus T(A)$ . Entonces existe un subcontinuo  $W$  de  $X$  tal que  $x \in \text{Int}_X(W) \subset W \subset X \setminus A$ . Observemos que, de hecho,  $\text{Int}_X(W) \subset X \setminus T(A)$ . Por tanto,  $X \setminus T(A)$  es abierto y  $T(A)$  es cerrado.

Ahora probaremos (2). Sea  $x \in X \setminus T(B)$ . Entonces existe un subcontinuo  $W$  de  $X$  tal que  $x \in \text{Int}_X(W) \subset W \subset X \setminus B$ . Como  $A \subset B$ ,  $X \setminus B \subset X \setminus A$ . De esta forma, concluimos que  $x \in \text{Int}_X(W) \subset W \subset X \setminus A$ . Por consiguiente,  $x \in X \setminus T(A)$ . Por tanto,  $T(A) \subset T(B)$ .

Notemos que (3) es una consecuencia inmediata de (2).  $\square$

2.8. OBSERVACIÓN. Dado un espacio métrico y compacto  $X$ , denotamos por  $2^X$  a la familia de los subconjuntos cerrados y no vacíos de  $X$ . Notemos que, por la parte (1) del Lema 2.7, el codominio de la función  $T$  es  $2^X \cup \{\emptyset\}$ .

2.9. OBSERVACIÓN. En general, no es cierto que se dé la igualdad en la parte (3) del Lema 2.7. Por ejemplo, en la suspensión sobre el conjunto de Cantor no se cumple la igualdad. De hecho, es una pregunta abierta el caracterizar a los espacios, en particular la clase de los continuos, para los cuales se cumple la mencionada igualdad.

2.10. EJEMPLO. Si  $X$  es el conjunto de Cantor entonces  $T(\emptyset) = X$ , ya que los únicos subcontinuos de  $X$  son puntos y todos ellos tienen interior vacío. Así, por la parte (2) del Lema 2.7, se tiene que  $T(A) = X$  para toda  $A \in P(X)$ .

2.11. EJEMPLO. Si  $X = \{0\} \cup \{\frac{1}{n}\}_{n=1}^{\infty}$  entonces  $T(\emptyset) = \{0\}$ , puesto que  $\{0\}$  es el único subcontinuo de  $X$  que no tiene interior. De esta forma, por la parte (2) del Lema 2.7, obtenemos que  $T(A) = \{0\} \cup A$  para toda  $A \in P(X)$ .

2.12. OBSERVACIÓN. Notemos que si a  $2^X$  le definimos la métrica de Hausdorff [12, 1.8.3] entonces podemos preguntarnos: ¿Cuándo es  $T|_{2^X}: 2^X \rightarrow 2^X$  continua? En los Ejemplos 2.10 y 2.11,  $T|_{2^X}$  es continua. Para más información sobre la continuidad de la función  $T$  se pueden consultar los artículos, [1] [11],[12] [13], [14], [15] y [5].

El siguiente resultado nos da una caracterización de los espacios métricos y compactos en los cuales la imagen del conjunto vacío es el conjunto vacío.

2.13. TEOREMA. Si  $X$  es un espacio métrico y compacto entonces  $T(\emptyset) = \emptyset$  si y sólo si  $X$  tiene un número finito de componentes.

DEMOSTRACIÓN. Supongamos que  $X$  tiene un número finito de componentes. Sean  $x \in X$  y  $C$  la componente de  $X$  que tiene a  $x$ . Como  $X$  tiene un número finito de componentes,  $C$  es tanto abierto como cerrado. De esta forma, todo punto de  $X$  está contenido en el interior de un subcontinuo propio de  $X$ . Por tanto,  $T(\emptyset) = \emptyset$ .

Ahora, supongamos que  $T(\emptyset) = \emptyset$ . Entonces para cada  $x \in X$ , existe un subcontinuo  $W_x$  de  $X$  tal que  $x \in \text{Int}_X(W_x)$ . Observemos que la familia  $\{\text{Int}_X(W_x) \mid x \in X\}$  forma un cubierta abierta de  $X$ . Como  $X$  es compacto, existen  $x_1, \dots, x_n \in X$  tales que  $X \subset \bigcup_{j=1}^n \text{Int}_X(W_{x_j}) \subset \bigcup_{j=1}^n W_{x_j}$ . De donde resulta que  $X$  tiene un número finito de componentes.  $\square$

Como una consecuencia inmediata del Teorema 2.13 tenemos:

2.14. COROLARIO. Si  $X$  es un continuo entonces  $T(\emptyset) = \emptyset$ .

Una demostración del siguiente teorema se puede encontrar en [17, 5.4]. Este resultado es conocido como: *El Teorema del Borde en la Frontera*.

2.15. TEOREMA. Sean  $X$  un continuo y  $U$  un subconjunto abierto y propio de  $X$ . Si  $K$  es una componente de  $Cl_X(U)$  entonces  $K \cap Fr_X(U) \neq \emptyset$ .

El siguiente teorema nos dice que la imagen de continuos bajo la función  $T$  es conexa.

2.16. TEOREMA. Sea  $X$  un continuo. Si  $W$  es un subcontinuo de  $X$  entonces  $T(W)$  es un subcontinuo de  $X$ .

DEMOSTRACIÓN. Ya sabemos que  $T(W)$  es cerrado, por la parte (1) del Lema 2.7. Supongamos que  $T(W)$  no es conexo. Entonces existen dos subconjuntos cerrados y no vacíos  $A$  y  $B$  de  $X$  tales que  $T(W) = A \cup B$ . Como  $W$  es conexo, supondremos que  $W \subset A$ . Como  $X$  es un espacio métrico, existe un subconjunto abierto  $U$  de  $X$  tal que  $A \subset U$  y  $Cl_X(U) \cap B = \emptyset$ . De aquí se obtiene que  $Fr_X(U) \cap T(W) = \emptyset$ . Entonces para cada  $z \in Fr_X(U)$ , existe un subcontinuo  $K_z$  de  $X$  tal que  $z \in Int_X(K_z) \subset K_z \subset X \setminus W$ . Como  $Fr_X(U)$  es compacta, existen  $z_1, \dots, z_n \in Fr_X(U)$  tales que  $Fr_X(U) \subset \bigcup_{j=1}^n Int_X(K_{z_j}) \subset \bigcup_{j=1}^n K_{z_j}$ . Sea  $V = U \setminus \bigcup_{j=1}^n K_{z_j}$ . Entonces  $V$  es un abierto de  $X$ . Sea  $Y = X \setminus V = (X \setminus U) \cup \left( \bigcup_{j=1}^n K_{z_j} \right)$ . Por el Teorema 2.15,  $Y$  tiene un número finito de componentes. Notemos que  $B \subset X \setminus Cl_X(U) \subset X \setminus U \subset Y$ . En consecuencia,  $B \subset Int_X(Y)$ . Sean  $b \in B$  y  $C$  la componente de  $Y$  que tiene a  $b$ . Entonces  $b \in Int_X(C)$  y  $C \cap W = \emptyset$ . De donde,  $b \in X \setminus T(W)$ , lo cual es una contradicción. Por tanto,  $T(W)$  es conexo.  $\square$

En lo que resta del trabajo, presentaremos, en términos de la función  $T$ , algunas de las propiedades estudiadas en topología, cuando el espacio en cuestión es un continuo.

2.17. DEFINICIÓN. Sean  $X$  un continuo y  $p \in X$ . Decimos que  $X$  es *casi conexo en pequeño en  $p$*  si para cada abierto  $U$  de  $X$  que tenga a  $p$ , se tiene que existe un subcontinuo  $W$  de  $X$  tal que  $Int_X(W) \neq \emptyset$  y  $W \subset U$ . El continuo  $X$  es *casi conexo en pequeño* si lo es en cada uno de sus puntos.

2.18. EJEMPLO. Si  $X$  es el cono sobre la sucesión armónica  $\{0\} \cup \{\frac{1}{n}\}_{n=1}^{\infty}$  entonces  $X$  es casi conexo en pequeño [12, 1.7.5].

La forma de expresar la casi conexidad en pequeño en un punto en términos de la función  $T$  está dada en el siguiente resultado:

2.19. TEOREMA. Si  $X$  es un continuo y  $p \in X$  entonces  $X$  es casi conexo en pequeño en  $p$  si y sólo si para cada  $A \in P(X)$  tal que  $p \in Int_X(T(A))$ , se tiene que  $p \in Cl_X(A)$ .

DEMOSTRACIÓN. Supongamos que  $X$  es casi conexo en pequeño en  $p$ . Sea  $A \in P(X)$  tal que  $p \in Int_X(T(A))$ . Entonces existe un número natural  $N$  tal que  $\mathcal{V}_{\frac{1}{n}}(p) \subset Int_X(T(A))$  para toda  $n \geq N$ . Como  $X$  es casi conexo en pequeño en  $p$ , para cada  $n \geq N$ , existe un subcontinuo  $W_n$  de  $X$  tal que  $Int_X(W_n) \neq \emptyset$  y  $W_n \subset \mathcal{V}_{\frac{1}{n}}(p)$ . De aquí se sigue que  $W_n \cap A \neq \emptyset$  para toda  $n \geq N$ . Sea  $x_n \in W_n \cap A$ . Observemos que, por construcción, la sucesión  $\{x_n\}_{n=N}^{\infty}$  converge a  $p$ . Por tanto,  $p \in Cl_X(A)$ .

Ahora, supongamos que para toda  $A \in P(X)$  tal que  $p \in Int_X(T(A))$ , se tiene que  $p \in Cl_X(A)$ . Sea  $U$  un abierto de  $X$  tal que  $p \in U$ . Sea  $V$  un abierto de  $X$  tal que  $p \in V \subset Cl_X(V) \subset U$ . Veremos que  $Cl_X(V)$  tiene alguna componente con interior distinto del vacío. Supongamos que esto no es cierto. Sea  $A = Fr_X(V)$ .

Entonces  $A$  es un subconjunto cerrado de  $X$  y  $p \in X \setminus A$ . Mostraremos que  $V \subset T(A)$ . Supongamos que hay un punto  $x \in V \setminus T(A)$ . Entonces existe un subcontinuo  $W$  de  $X$  tal que  $x \in \text{Int}_X(W) \subset W \subset X \setminus A$ . Como todas las componentes de  $Cl_X(V)$  tienen interior vacío y  $W$  es un subcontinuo con interior no vacío, resulta que  $W \cap X \setminus V \neq \emptyset$ . Lo que implica que  $W \cap Fr_X(W) \neq \emptyset$ ; i.e.,  $W \cap A \neq \emptyset$ , lo cual es una contradicción. Por tanto,  $V \subset T(A)$ . De donde,  $p \in \text{Int}_X(T(A))$ . Por hipótesis, lo anterior implica que  $p \in Cl_X(A) = A$ . Esto es una contradicción. Por tanto,  $Cl_X(V)$  tiene una componente con interior no vacío y  $X$  es casi conexo en pequeño en  $p$ .  $\square$

La versión global de la casi conexidad en pequeño usando la función  $T$  se encuentra a continuación:

**2.20. TEOREMA.** Un continuo es casi conexo en pequeño si y sólo si para toda  $F \in 2^X$ ,  $\text{Int}_X(F) = \text{Int}_X(T(F))$ .

**DEMOSTRACIÓN.** Supongamos que  $X$  es casi conexo en pequeño. Sea  $F \in 2^X$ . Como  $F \subset T(F)$  (Observación 2.6), se tiene que  $\text{Int}_X(F) \subset \text{Int}_X(T(F))$ . Sea  $x \in \text{Int}_X(T(F))$ . Por el Teorema 2.19,  $x \in F$ . De donde,  $\text{Int}_X(T(F)) \subset F$ . Así,  $\text{Int}_X(T(F)) \subset \text{Int}_X(F)$ . Por tanto,  $\text{Int}_X(F) = \text{Int}_X(T(F))$ .

Supongamos ahora que para cada  $F \in 2^X$ ,  $\text{Int}_X(F) = \text{Int}_X(T(F))$ . Sean  $x \in X$  y  $U$  un subconjunto abierto de  $X$  tal que  $x \in U$ . Como  $\text{Int}_X(X \setminus U) \cap U = \emptyset$ , por hipótesis, resulta que  $\text{Int}_X(T(X \setminus U)) \cap U = \emptyset$ . De donde, existe  $y \in U \setminus T(X \setminus U)$ . Entonces existe un subcontinuo  $W$  de  $X$  tal que  $y \in \text{Int}_X(W) \subset W \subset U$ . Por tanto,  $X$  es casi conexo en pequeño en  $x$ . Como  $x$  fue un punto arbitrario de  $X$ ,  $X$  es casi conexo en pequeño.  $\square$

**2.21. DEFINICIÓN.** Sean  $X$  un continuo  $p \in X$ . Decimos que  $X$  es *conexo en pequeño en  $p$*  si para todo subconjunto abierto  $U$  de  $X$  tal que  $p \in U$ , existe un subcontinuo  $W$  de  $X$  tal que  $p \in \text{Int}_X(W) \subset W \subset U$ .

**2.22. OBSERVACIÓN.** Es probable que el lector conozca alguna otra manera de definir conexidad en pequeño en un punto. En [12, 1.7.9] hay varias definiciones equivalentes. Observemos que no se definió un continuo conexo en pequeño como un continuo que es conexo en pequeño en cada uno de sus puntos. Esto es porque, un continuo es conexo en pequeño en cada uno de sus puntos si y sólo si el continuo es localmente conexo en cada uno de sus puntos (Definición 2.32) [12, 1.7.12].

**2.23. EJEMPLO.** Consideremos la sucesión armónica  $\{0\} \cup \{\frac{1}{n}\}$ . Entre  $\frac{1}{n}$  y  $\frac{1}{n+1}$  coloquemos una copia del espacio del Ejemplo 2.18 de tal forma que el vértice coincida con el punto  $\frac{1}{n}$  y la “barra límite” coincida con el intervalo  $[\frac{1}{n}, \frac{1}{n+1}]$ . Notemos que  $X$  es casi conexo en pequeño en 0 [12, 1.7.7].

Ahora describiremos la conexidad en pequeño en un punto en términos de la función  $T$ .

**2.24. TEOREMA.** Si  $X$  es un continuo y  $p \in X$  entonces  $X$  es conexo en pequeño en  $p$  si y sólo si para toda  $A \in P(X)$  tal que  $p \in T(A)$  se tiene que  $p \in Cl_X(A)$ .

**DEMOSTRACIÓN.** Supongamos que  $X$  es conexo en pequeño en  $p$ . Sea  $A \in P(X)$  y supongamos que  $p \in X \setminus Cl_X(A)$ . Como  $X \setminus Cl_X(A)$  es abierto y  $X$  es conexo en pequeño en  $p$ , existe un subcontinuo  $W$  de  $X$  tal que  $p \in \text{Int}_X(W) \subset W \subset X \setminus Cl_X(A) \subset X \setminus A$ . De donde,  $p \in X \setminus T(A)$ .

Ahora, supongamos que para toda  $A \in P(X)$  tal que  $p \in T(A)$  se tiene que  $p \in Cl_X(A)$ . Sea  $U$  un subconjunto abierto de  $X$  tal que  $p \in U$ . Notemos que  $X \setminus U$  es cerrado y que  $p \notin X \setminus U$ . De aquí se sigue, por hipótesis, que  $p \notin T(X \setminus U)$ . Esto implica que existe un continuo  $W$  tal que  $p \in Int_X(W) \subset W \subset U$ . Por tanto,  $X$  es conexo en pequeño en  $p$ .  $\square$

2.25. DEFINICIÓN. Sean  $X$  un continuo y  $p$  y  $q \in X$ . Decimos que  $X$  es *aposindético en  $p$  con respecto a  $q$*  si existe un subcontinuo  $W$  de  $X$  tal que  $p \in Int_X(W) \subset W \subset X \setminus \{q\}$ . El continuo  $X$  es *aposindético en  $p$*  si es aposindético en  $p$  con respecto a cualquier punto de  $X \setminus \{p\}$ . Finalmente,  $X$  es *aposindético* si lo es en cada uno de sus puntos.

Como una consecuencia inmediata de la Definición 2.25 tenemos la manera de poner a la aposindesis puntual en términos de la función  $T$ :

2.26. TEOREMA. Sean  $X$  un continuo y  $p$  y  $q \in X$ . Entonces  $X$  es aposindético en  $p$  con respecto a  $q$  si y sólo si  $p \in X \setminus T(\{q\})$ .

La versión global de la aposindesis usando la función  $T$  se expresa como sigue:

2.27. TEOREMA. Un continuo  $X$  es aposindético si y sólo si  $T(\{p\}) = \{p\}$  para toda  $p \in X$ .

DEMOSTRACIÓN. Supongamos que  $X$  es aposindético y sea  $p \in X$ . Tomemos  $q \in X \setminus \{p\}$ . Por el Teorema 2.26, se tiene que  $q \in X \setminus T(\{p\})$ . Por tanto,  $T(\{p\}) = \{p\}$ .

Supongamos ahora que  $T(\{x\}) = \{x\}$  para toda  $x \in X$ . Sean  $p$  y  $q$  dos puntos distintos de  $X$ . Como  $T(\{q\}) = \{q\}$ ,  $p \in X \setminus T(\{q\})$ . Por el Teorema 2.26,  $X$  es aposindético en  $p$  con respecto a  $q$ . Como  $p$  y  $q$  fueron dos puntos arbitrarios de  $X$ ,  $X$  es aposindético.  $\square$

2.28. DEFINICIÓN. Sean  $X$  un continuo y  $p \in X$ . Decimos que  $X$  es *semilocalmente conexo en  $p$*  si para todo abierto  $U$  de  $X$  que tiene a  $p$ , existe un subconjunto abierto  $V$  de  $X$  tal que  $p \in V \subset U$  y  $X \setminus V$  tiene un cantidad finita de componentes. Decimos que  $X$  es *semilocalmente conexo* si es semilocalmente conexo en cada uno de sus puntos.

2.29. EJEMPLO. La suspensión del conjunto de Cantor es un espacio semilocalmente conexo, el cual no es localmente conexo.

La forma de expresar a la semiconexidad local en un punto utilizando a la función  $T$  se encuentra a continuación:

2.30. TEOREMA. Sean  $X$  un continuo y  $p \in X$ . Entonces  $X$  es semilocalmente conexo en  $p$  si y sólo si  $T(\{p\}) = \{p\}$ .

DEMOSTRACIÓN. Supongamos que  $X$  es semilocalmente conexo en  $p$ . Sea  $q \in X \setminus \{p\}$ . Tomemos un subconjunto abierto  $U$  de  $X$  tal que  $p \in U$  y  $q \in X \setminus Cl_X(U)$ . Como  $X$  es semilocalmente conexo en  $p$ , existe un subconjunto abierto  $V$  de  $X$  tal que  $p \in V \subset U$  y  $X \setminus V$  tiene una cantidad finita de componentes. Sea  $K$  la componente de  $X \setminus V$  que tiene a  $q$ . Entonces  $q \in Int_X(K)$  [12, 1.6.2]. Esto implica que  $q \in X \setminus T(\{p\})$ . Por tanto,  $T(\{p\}) = \{p\}$ .

Ahora, supongamos que  $T(\{p\}) = \{p\}$ . Sea  $U$  un subconjunto abierto de  $X$  tal que  $p \in U$ . Como  $T(\{p\}) = \{p\}$ , para cada  $q \in X \setminus U$ , existe un subcontinuo

$W_q$  de  $X$  tal que  $q \in \text{Int}_X(W_q) \subset W_q \subset X \setminus \{p\}$ . Como  $X \setminus U$  es compacto, existen  $q_1, \dots, q_n \in X \setminus U$  tales que  $X \setminus U \subset \bigcup_{j=1}^n \text{Int}_X(W_{q_j}) \subset \bigcup_{j=1}^n W_{q_j}$ . Sea  $V = X \setminus \bigcup_{j=1}^n W_{q_j}$ . Entonces  $V$  es un subconjunto abierto de  $X$  tal que  $p \in V \subset U$  y  $X \setminus V$  tiene un número finito de componentes. Por tanto,  $X$  es semilocalmente conexo en  $p$ .  $\square$

Como consecuencia de los Teoremas 2.27 y 2.30 tenemos:

2.31. COROLARIO. Un continuo  $X$  es aposindético si y sólo si  $X$  es semilocalmente conexo.

2.32. DEFINICIÓN. Sean  $X$  un continuo y  $p \in X$ . Decimos que  $X$  es *localmente conexo en  $p$*  si para todo abierto  $U$  de  $X$ , existe un abierto y conexo  $V$  de  $X$  tal que  $p \in V \subset U$ . El continuo  $X$  es *localmente conexo* si lo es en cada uno de sus puntos.

La función  $T$  caracteriza a los continuos localmente conexos siendo la identidad en la familia de subconjuntos cerrados y no vacíos de dichos continuos.

2.33. TEOREMA. Un continuo  $X$  es localmente conexo si sólo si  $T(A) = A$  para toda  $A \in 2^X$ .

DEMOSTRACIÓN. Supongamos que  $X$  es localmente conexo. Sean  $A \in 2^X$  y  $p \in X \setminus A$ . Tomemos un subconjunto abierto  $U$  de  $X$  tal que  $p \in U \subset \text{Cl}_X(U) \subset X \setminus A$ . Como  $X$  es localmente conexo, existe un subconjunto abierto y conexo  $V$  de  $X$  tal que  $p \in V \subset U$ . Entonces  $\text{Cl}_X(V)$  es un subcontinuo de  $X$  tal que  $p \in \text{Int}_X(\text{Cl}_X(V)) \subset \text{Cl}_X(V) \subset X \setminus A$ . Lo que implica que  $p \in X \setminus T(A)$ . De donde,  $T(A) \subset A$ . Por la Observación 2.6,  $A \subset T(A)$ . Por tanto,  $T(A) = A$ .

Finalmente, supongamos que  $T(A) = A$  para toda  $A \in 2^X$ . Sean  $p \in X$  y  $U$  un subconjunto abierto de  $X$  tal que  $p \in U$ . Como  $U$  es abierto,  $X \setminus U$  es cerrado. Por hipótesis, se tiene que  $T(X \setminus U) = X \setminus U$ . De donde, existe un subcontinuo  $W$  de  $X$  tal que  $p \in \text{Int}_X(W) \subset W \subset U$ . Lo que implica que  $X$  es conexo en pequeño en  $p$ . Como  $p$  fue un punto arbitrario de  $X$ ,  $X$  es conexo en pequeño en cada uno de sus puntos. Por tanto,  $X$  es localmente conexo [12, 1.7.12].  $\square$

AGRADECIMIENTOS: El autor le agradece al Dr. Raúl Escobedo la invitación a participar en la *Quinta Gran Semana de las Matemáticas* de la Benemérita Universidad Autónoma de Puebla.

## REFERENCIAS

- [1] D. P. Bellamy, *Continua for which the set function  $T$  is continuous*, Trans. Amer. Math. Soc., 151 (1970) 581–587.
- [2] D. P. Bellamy and J. J. Charatonik, *The set function  $T$  and contractibility of continua*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys., 25 (1977), 47–49.
- [3] D. P. Bellamy and C. L. Hagopian, *Mapping continua onto their cones*, Colloq. Math., 41 (1979), 53–56.
- [4] H. S. Davis y P. H. Doyle, *Invertible continua*, Portugal. Math., 26 (1967), 487–491.
- [5] L. Fernández y S. Macías, *Set functions  $T$  and  $K$  and irreducible continua*, manuscrito.
- [6] F. B. Jones, *Aposyndetic continua and certain boundary problems*, Amer. J. Math., 53 (1941), 545–553.
- [7] F. B. Jones, *Concerning nonaposyndetic continua*, Amer. J. Math., 70 (1948), 403–413.
- [8] F. B. Jones, *On a Certain type of homogeneous plane continua*, Proc. Amer. Math. Soc., 6 (1955), 735–740.

- [9] K. Kuratowski, *Topology*, Vol. II, Academic Press, New York, N. Y., 1968.
- [10] S. Macías, *Aposyndetic properties of symmetric products of continua*, *Topology Proc.*, 22 (1997), 281–296.
- [11] S. Macías, *A class of one-dimensional, nonlocally connected continua for which the set function  $T$  is continuous*, *Houston J. Math.*, 32 (2006), 161–165.
- [12] S. Macías, *Topics on Continua*, Pure and Applied Mathematics Series, Vol. 275, Chapman & Hall/CRC, Taylor & Francis Group, Boca Raton, London, New York, Singapore, 2005.
- [13] S. Macías, *Homogeneous continua for which the set function  $T$  is continuous*, *Topology Appl.*, 153 (2006), 3397–3401.
- [14] S. Macías, *A decomposition theorem for a class of continua for which the set function  $T$  is continuous*, *Colloq. Math.*, 109 (2007), 163–170.
- [15] S. Macías, *On continuously irreducible continua*, *Topology Appl.*, 156 (2009), 2357–2363.
- [16] S. Macías and S. B. Nadler, Jr., *On hereditarily decomposable homogeneous continua*, *Topology Proc.*, 34 (2009), 131–145.
- [17] S. B. Nadler, Jr., *Continuum Theory: An Introduction*, Monographs and Textbooks in Pure and Applied Math., Vol. 158, Marcel Dekker, New York, Basel, Hong Kong, 1992.
- [18] G. T. Whyburn, *Semi-locally-connected sets*, *Amer. J. Math.*, 61 (1939), 733–749.

Instituto de Matemáticas, Universidad Nacional Autónoma de México  
Circuito Exterior, Ciudad Universitaria  
México D. F., C. P. 04510  
`macias@servidor.unam.mx`

# TODO ESPACIO MÉTRICO ES PERFECTO

ARMANDO MARTÍNEZ GARCÍA  
FCFM - BUAP

RESUMEN. El objetivo de este artículo es dar una demostración alterna del hecho conocido de que todo espacio métrico  $(X, d)$  es un espacio perfecto, es decir, todo conjunto abierto  $W$  es un conjunto  $F_\sigma$ .

## 1. INTRODUCCIÓN

La definición de espacio métrico aparece por primera vez en la tesis doctoral Frechet [3]. La propiedad de que un espacio sea perfectamente normal lo ve Urysohn en [4] y fue estudiada por Alexandroff y Urysohn en [1] en la clase de los espacios compactos. La clase de espacios perfectamente normales fue definida por Čech en [2]. El criterio que se aplica para demostrar que un espacio métrico es perfecto es el dado por Vedenisoff en [5] y [6]. La demostración que aquí presentamos será constructiva.

Para cada conjunto abierto  $W$  de  $(X, d)$ , construiremos una familia numerable de subconjuntos  $A(W, n)$ , tal que las siguientes condiciones se satisfacen:

- 1)  $W = \bigcup_{n \in \mathbb{N}} A(W, n)$ .
- 2)  $A(W, n) \subset A(W, n + 1)$ .
- 3)  $A(W, n)$  es un subconjunto cerrado de  $(X, d)$  para todo  $n \in \mathbb{N}$ .

## 2. PRELIMINARES

En esta sección daremos las definiciones y resultados necesarios para construir la familia de subconjuntos  $A(W, n)$ , que satisfagan las condiciones deseadas.

2.1. DEFINICIÓN. Un espacio métrico es una pareja  $(X, d)$ , consistente de un conjunto  $X$  y una función  $d$  definida en el conjunto  $X \times X$ , que toma sus valores en  $\mathbb{R}$  y satisface las siguientes condiciones:

- 1)  $d(x, y) = 0$  si y sólo si  $x = y$ .
- 2)  $d(x, y) \geq 0$  para todo  $x, y \in X$ .
- 3)  $d(x, y) = d(y, x)$  para todo  $x, y \in X$ .
- 4)  $d(x, y) \leq d(x, z) + d(z, y)$  para todo  $x, y, z \in X$ .

2.2. DEFINICIÓN. Sea  $(X, d)$  un espacio métrico.

1) Dado  $x \in X$  y  $r > 0$  la bola abierta con centro en  $x$  y radio  $r$ , la cual denotaremos como  $B(x, r)$ , es el conjunto:

$$B(x, r) = \{y \in X : d(x, y) < r\}.$$

2)  $U \subset X$  es un conjunto abierto de  $X$ , si para cada  $x \in U$ , existe  $r > 0$  tal que  $B(x, r) \subset U$ .

2.3. DEFINICIÓN. Sean  $(X, d)$  un espacio métrico y  $A \subset X$ .

- 1)  $A$  es un conjunto cerrado de  $X$ , si su complemento es un conjunto abierto.
- 2)  $A$  es un conjunto  $F_\sigma$ , si es unión numerable de conjuntos cerrados.

2.4. DEFINICIÓN. Sean  $(X, d)$  un espacio métrico y  $A \subset X$ . La cerradura de  $A$  en  $X$  es la intersección de todos los subconjuntos cerrados de  $X$  que contienen a  $A$  y se denotará como  $cl_X A$ .

Se sigue de la definición de cerradura de  $A$ , que  $cl_X A$  es cerrado de donde tenemos el siguiente resultado.

2.5. PROPOSICIÓN. Sean  $(X, d)$  un espacio métrico y  $A \subset X$ . Entonces  $A$  es cerrado si y sólo si  $A = cl_X A$ .

2.6. PROPOSICIÓN. Sean  $(X, d)$  un espacio métrico y  $A \subset X$ . Entonces para cada  $x \in cl_X A$  existe una sucesión  $\{x_n\}_{n \in \mathbb{N}} \subset A$  tal que,  $\{x_n\}_{n \in \mathbb{N}}$  converge a  $x$ .

DEMOSTRACIÓN. Como  $x \in cl_X A$ , para cada  $n \in \mathbb{N}$  elijamos  $x_n \in B(x, 1/n) \cap A$ . Es claro que la sucesión  $\{x_n\}_{n \in \mathbb{N}}$  así elegida converge a  $x$ .  $\square$

2.7. DEFINICIÓN. Sea  $(X, d)$  un espacio métrico.  $(X, d)$  es perfecto si todo conjunto abierto de  $X$  es unión numerable de conjuntos cerrados.

### 3. CONSTRUCCIÓN

En esta sección construiremos para cada conjunto abierto  $W$  del espacio métrico  $(X, d)$  una familia numerable de subconjuntos  $A(W, n)$ , tal que las siguientes condiciones se satisfacen:

- 1)  $W = \bigcup_{n \in \mathbb{N}} A(W, n)$ .
- 2)  $A(W, n) \subset A(W, n + 1)$ .
- 3)  $A(W, n)$  es un subconjunto cerrado de  $(X, d)$  para todo  $n \in \mathbb{N}$ .

Dicha construcción se hará mediante una serie de lemas.

Sea  $(X, d)$  un espacio métrico y  $W$  un conjunto abierto. Definamos el conjunto

$$A(W, n) = \{x \in W : B(x, 1/n) \subset W\}$$

como  $W$  es un conjunto abierto para cada  $x \in W$  existe  $n_0 \in \mathbb{N}$  tal que

$$B(x, 1/n_0) \subset W \text{ por lo tanto } A(W, n_0) \neq \emptyset.$$

3.1. LEMA.  $W = \bigcup_{n \in \mathbb{N}} A(W, n)$ .

DEMOSTRACIÓN. Es claro que para cada  $n \in \mathbb{N}$ ,  $A(W, n) \subset W$  y si  $x \in W$  existe  $n_x \in \mathbb{N}$  tal que  $B(x, 1/n_x) \subset W$  de donde se sigue que

$$W \subset \bigcup_{x \in W} A(W, n_x) \subset \bigcup_{n \in \mathbb{N}} A(W, n).$$

Por lo tanto;

$$W = \bigcup_{n \in \mathbb{N}} A(W, n). \quad \square$$

3.2. LEMA.  $A(W, n) \subset A(W, n + 1)$ .

DEMOSTRACIÓN. Para cada  $x \in A(W, n)$  se tiene que  $B(x, 1/n) \subset W$  y como  $B(x, 1/n + 1) \subset B(x, 1/n)$ , se sigue que;

$$A(W, n) \subset A(W, n + 1). \quad \square$$

3.3. LEMA. Si  $x \in cl_X A(W, n)$  entonces  $x \in W$ .

DEMOSTRACIÓN. Como  $x \in cl_X A(W, n)$  se sigue de la Proposición 2.6 que existe una sucesión  $\{x_i\}_{i \in \mathbb{N}} \subset A(W, n)$  tal que  $\{x_i\}_{i \in \mathbb{N}}$  converge a  $x$ .

Para  $\epsilon = 1/2n$

existe  $N \in \mathbb{N}$  tal que  $x_i \in B(x, 1/2n)$  para  $i > N$

lo cual implica que

$$d(x, x_i) < 1/2n \text{ para } i > N$$

de donde tenemos que

$$d(x, x_i) < 1/n \text{ para } i > N$$

por lo tanto

$$x \in B(x_i, 1/n).$$

Como  $x_i \in A(W, n)$  se tiene que  $B(x_i, 1/n) \subset W$  y como  $x \in B(x_i, 1/n)$  se sigue que;

$$x \in W.$$

□

3.4. LEMA.  $cl_X A(W, n) = A(W, n)$

DEMOSTRACIÓN. Sea  $x \in cl_X A(W, n)$ . Veremos que  $x \in A(W, n)$ , para lo cual será suficiente ver que  $B(x, 1/n) \subset W$ .

Como  $x \in cl_X A(W, n)$  se sigue del Lema 3.3 que  $x \in W$ .

Sea  $y \in B(x, 1/n)$ .

Si  $y = x$  entonces  $y \in W$  ya que  $x \in W$ .

Si  $y \neq x$  entonces  $0 < d(y, x) < 1/n$ , sea  $r = 1/n - d(y, x)$ .

Como  $x \in cl_X A(W, n)$  entonces  $B(x, r) \cap A(W, n) \neq \emptyset$ , sea  $z \in B(x, r) \cap A(W, n)$ , entonces  $z \in B(x, r)$  y  $z \in A(W, n)$ .

Como  $z \in B(x, r)$ ,  $d(z, x) < r$  de donde  $d(z, x) < 1/n - d(x, y)$  por lo tanto,

$$d(z, y) < d(z, x) + d(x, y) < 1/n,$$

es decir,

$$y \in B(z, 1/n).$$

Y como  $B(z, 1/n) \subset W$  tenemos que

$$y \in W.$$

□

3.5. TEOREMA. Sea  $(X, d)$  espacio métrico. Entonces  $(X, d)$  es perfecto.

DEMOSTRACIÓN. Se sigue de los Lemas 3.1 y 3.4.

□

3.6. COROLARIO.  $\mathbb{R}$  con la métrica usual es perfecto.

DEMOSTRACIÓN. Se sigue del Teorema 3.5.

□

A  $\mathbb{R}$  con la métrica usual lo denotaremos como  $X_u$

3.7. EJEMPLO. Consideremos a  $\mathbb{R}$  con la topología de Sorgenfrey, la cual tiene como base la familia de intervalos  $[a, b)$  con  $a, b \in \mathbb{R}$  y  $a < b$ .

A  $\mathbb{R}$  con la topología de Sorgenfrey lo denotaremos como  $X_s$ . Afirmamos que este es un espacio perfecto.

Como para  $x \in (a, b)$  tenemos que  $x \in [x, b) \subset (a, b)$ , se sigue que la topología de  $X_s$  es más fina que la topología de  $X_u$ .

Sea  $U$  abierto en  $X_s$ , y  $V$  el interior de  $U$  en  $X_u$ . Entonces  $V$  es unión numerable de intervalos  $(a, b)$ , cada uno de los cuales es un conjunto  $F_\sigma$  en  $X_u$ , por lo tanto es un conjunto  $F_\sigma$  en  $X_s$ . Ahora si  $A = U - V$ , como  $|A| \leq \omega$ ,  $A$  es un conjunto  $F_\sigma$  de  $X_s$ ; por lo tanto,  $U$  es un conjunto  $F_\sigma$ .

3.8. OBSERVACIÓN. No todo espacio primero numerable es perfecto.

Para ver esto consideremos el conjunto  $X = I \times I$  con la topología inducida por el orden lexicográfico en él y denotemos este espacio como  $X_I$ .

$X_I$  es primero numerable y compacto. En particular,  $X_I$  es un espacio Lindelöf. Por lo tanto, si  $X_I$  fuese perfecto éste sería hereditariamente Lindelöf lo cual es falso ya que el subespacio  $Y = \{(x, 1/2) \in X_I : x \in I\}$  tiene la topología discreta.

3.9. PREGUNTA. ¿Se puede probar que  $X_s$  es perfecto construyendo una familia numerable de conjuntos  $A(W, n)$  que satisfaga las mismas condiciones de los conjuntos construidos anteriormente?

#### REFERENCIAS

- [1] Alexandroff, P., Urysohn, P., *Memoire sur les espaces topologiques compacts*, Verh. Akad. Wetensch. Amsterdam 14(1929).
- [2] Čech, E., *Sur la dimension des espaces parfaitement normaux*, Bull. Intern. Acad. Tchèque Sci. 33(1932), 38-55.
- [3] Fréchet, M., *Sur quelques points du calcul fonctionnel*, Rend. del Circ. Mat. di Palermo 22(1906), 1-74.
- [4] Urysohn, P., *Über die Mächtigkeit der zusammenhängenden Mengen*, Math. Ann. 94(1925), 262-295.
- [5] Vedenisoff, N., *Sur les fonctions continues dans les espaces topologiques*, Fund. Math. 27(1936), 234-238.
- [6] Vedenisoff, N., *Généralisation de quelques théorèmes sur la dimension*, Comp. Math 7(1940), 194-200.

Facultad de Ciencias Físico Matemáticas, BUAP.

Av. San Claudio y 18 Sur, Col. San Manuel, Ciudad Universitaria

Puebla, Pue., C.P. 72570

maga@cfm.buap.mx

# EXTENSIONES SIMPLES, CONEXIDAD Y SEPARACIÓN

MARCELINO TEXIS TEXIS  
FACULTAD DE CIENCIAS DE LA ELECTRÓNICA - BUAP

RESUMEN. Para construir una topología más fina que otra, uno de los métodos que se usa es el de *extensión simple*, que fue introducido por Levine en [4]. En el presente trabajo probaremos que bajo ciertas condiciones las propiedades de ser normal, completamente regular y conexo se preservan bajo extensiones simples.

## 1. INTRODUCCIÓN Y PRELIMINARES

Sean  $X$  un conjunto (no vacío),  $\tau$  y  $\sigma$  dos topologías sobre  $X$ . Se dice que  $\sigma$  es una extensión de  $\tau$  si ocurre que  $\tau \subset \sigma$ . Uno tiene que en el caso en que  $\sigma$  es una extensión de  $\tau$ , la función identidad  $id : (X, \sigma) \rightarrow (X, \tau)$  es una función continua. Y de este hecho se sigue que si  $\mathcal{P}$  es una propiedad que se preserva bajo imágenes continuas, entonces podemos garantizar que si el espacio  $(X, \sigma)$  tiene la propiedad  $\mathcal{P}$ , el espacio  $(X, \tau)$  también tendrá la propiedad  $\mathcal{P}$ . Tal es el caso de  $\mathcal{P} = \textit{compacidad}$  o  $\mathcal{P} = \textit{conexidad}$ , por ejemplo. Esto resuelve, parcialmente, el problema siguiente.

1.1. PROBLEMA. Suponga que  $\mathcal{P}$  es una propiedad y que  $\sigma$  es una extensión de  $\tau$ . Suponga además que  $(X, \sigma)$  tiene la propiedad  $\mathcal{P}$ . ¿Bajo qué condiciones ocurre que  $(X, \tau)$  cumple  $\mathcal{P}$ ?

En el mismo contexto, el problema anterior nos lleva naturalmente a su recíproco:

1.2. PROBLEMA. Suponga que  $\mathcal{P}$  es una propiedad y que  $\sigma$  es una extensión de  $\tau$ . ¿Bajo qué condiciones se verifica que al tener  $(X, \tau)$  la propiedad  $\mathcal{P}$ , ocurre que  $(X, \sigma)$  cumple  $\mathcal{P}$ ?

Nos ocuparemos del Problema 1.2 para tres tipos de propiedades, a saber: *normal*, *completamente regular* y *conexo*. Vale la pena comentar que el primer trabajo sobre extensiones simples es debido a Norman Levine [4], mismo que sirvió de base y guía para nuestro trabajo [2].

Si  $X$  es un espacio topológico y  $p \in X$ , diremos que  $U$  es una vecindad de  $p$ , si  $U$  es un conjunto abierto en  $X$  y  $p \in U$ . Si  $X$  es un espacio topológico y  $A \subseteq X$ , denotaremos por  $\textit{int}(A)$  al interior de  $A$  y, por comodidad, usamos  $\bar{A}$  y  $\textit{cl}_X(A)$ , para denotar la *clausura* (también llamada *cerradura*) de  $A$ . En caso de trabajar con varios espacios a la vez, usamos sólo la notación  $\textit{cl}_Y(A)$ , para denotar la clausura de  $A$  respecto del espacio  $Y$ . Además, si  $D \subseteq X$  es tal que  $\bar{D} = X$ , diremos que  $D$  es un subconjunto denso de  $X$ . Decimos que  $X$  es conexo, si no existe una separación de  $X$ , es decir, no existe un par  $U, V$  de abiertos disjuntos no triviales de  $X$  cuya unión es  $X$ . Un espacio topológico  $X$  es:

- (1)  $T_0$ , si dados dos puntos distintos  $x, y \in X$  existe un conjunto abierto que contiene a uno de esos puntos pero no al otro.
- (2)  $T_1$ , si dados dos puntos distintos  $x, y \in X$  existen dos abiertos  $U$  y  $V$  tales que  $x \in U, y \notin U, y \in V$  y  $x \notin V$ .
- (3)  $T_2$  o Hausdorff, si dados dos puntos distintos  $x, y \in X$  existen vecindades  $U_x$  y  $U_y$  de  $x$  y  $y$ , respectivamente, tales que  $U_x \cap U_y = \emptyset$ .
- (4) Regular, si para todo subconjunto cerrado  $F$  y un punto  $x \in X \setminus F$ , existen conjuntos abiertos ajenos  $U_1$  y  $U_2$  tales que  $F \subseteq U_1$  y  $x \in U_2$ .
- (5)  $T_3$ , si  $X$  es regular y  $T_1$ .
- (6) Normal, si para cualesquiera conjuntos cerrados  $F_1$  y  $F_2$ , con  $F_1 \cap F_2 = \emptyset$ , existen conjuntos abiertos ajenos  $U_1$  y  $U_2$  tales que  $F_1 \subseteq U_1, F_2 \subseteq U_2$ .
- (7)  $T_4$ , si  $X$  es normal y  $T_1$ .
- (8) Completamente regular, si para todo  $x \in X$  y para todo  $F \subset X$  cerrado con  $x \notin F$ , existe  $f : X \rightarrow [0, 1]$  continua tal que  $f(x) = 0$  y  $f(y) = 1$  para todo  $y \in F$ .
- (9) Tychonoff, si es completamente regular y  $T_1$ .

Algunas veces escribimos  $X \in T_i$ , en lugar de decir  $X$  es un espacio  $T_i$ , para  $i \in \{1, \dots, 4\}$ .

La extensión simple es uno de los métodos para construir una topología más fina que otra. Este concepto fue introducido por Levine en [4]. A continuación daremos el concepto de extensión y presentamos algunos resultados básicos sobre extensiones simples. El lector interesado en más información sobre el tema puede consultar los trabajos [2] o [4].

1.3. DEFINICIÓN. Sean  $\tau_1$  y  $\tau_2$  topologías sobre un conjunto  $X$ , decimos que:

- (1)  $\tau_2$  es una extensión de  $\tau_1$  si  $\tau_1 \subset \tau_2$ .
- (2)  $\tau_2$  es una extensión simple de  $\tau_1$  si existe  $A \subset X$  tal que  $\tau_1 \cup \{A\}$  es una subbase para  $\tau_2$ .

Es claro que toda extensión simple de  $\tau_1$  es una extensión de  $\tau_1$ . El lema siguiente describe explícitamente cómo son los abiertos de una extensión simple.

1.4. LEMA. Sean  $\tau_1$  y  $\tau_2$  topologías sobre un conjunto  $X$ . Entonces  $\tau_2$  es una extensión simple de  $\tau_1$  si y sólo si existe  $A \subset X$  tal que  $\tau_2 = \{U \cup (V \cap A) : U, V \in \tau_1\}$

Usaremos el símbolo  $\tau(A)$  para denotar la extensión simple de  $\tau$  generada por  $A$ . A continuación veamos qué relación tienen el interior y la cerradura en la topología original, con el interior y la cerradura tomadas en la topología  $\tau(A)$ .

1.5. TEOREMA. Sea  $(X, \tau)$  un espacio topológico,  $A, B \subset X, \tau^* = \tau(A)$ . Entonces

- (1)  $int_{\tau^*} B = int_{\tau} B \cup int_{\tau|_A}(B \cap A)$
- (2)  $cl_{\tau^*} B = cl_{\tau} B \cap ((X \setminus A) \cup (A \cap cl_{\tau}(B \cap A)))$
- (3)  $(A, \tau|_A) = (A, \tau^*|_A)$
- (4)  $(X \setminus A, \tau|(X \setminus A)) = (X \setminus A, \tau^*|(X \setminus A))$
- (5)  $cl_{\tau}(B \cap A) = cl_{\tau^*}(B \cap A)$
- (6)  $A$  es cerrado en  $\tau^*$  si y sólo si  $A$  es cerrado en  $\tau$
- (7) Si  $F$  es cerrado en  $\tau$  o en  $\tau^*$ , entonces  $F \cap A$  es cerrado en  $(A, \tau|_A)$  y  $F \cap (X \setminus A)$  es cerrado en  $(X \setminus A, \tau(X \setminus A))$

## 2. AXIOMAS DE SEPARACIÓN Y CONEXIDAD

2.1. TEOREMA. Si  $(X, \tau)$  es  $T_i$  entonces  $(X, \tau(A))$  es  $T_i$  para cada  $A \subset X$ ,  $i = 0, 1, 2$ .

DEMOSTRACIÓN.

Se sigue del hecho de que la topología  $\tau(A)$  es más fina que  $\tau$ ; en otras palabras todo abierto en  $\tau$  también es abierto en  $\tau(A)$ ; es decir  $\tau \subset \tau(A)$ .  $\square$

2.2. TEOREMA. Si  $(X, \tau)$  es completamente regular,  $A \notin \tau$  y  $X \setminus A \in \tau$ , entonces  $(X, \tau(A))$  es completamente regular.

DEMOSTRACIÓN. Sean  $B$  un conjunto cerrado en  $X$  con la topología  $\tau(A)$  y  $x \in X$  tales que  $x \notin B$ . Así, existen elementos  $U$  y  $V$  en  $\tau$  tales que  $x \in (X \setminus B) = U \cup (V \cap A)$ . Tenemos dos casos:

(1)  $x \in U$ . Entonces  $x \notin (X \setminus U)$  y  $(X \setminus U)$  es un conjunto cerrado de  $X$  con la topología  $\tau$ . Como  $(X, \tau)$  es completamente regular, existe una función continua  $f : X \rightarrow [0, 1]$  relativa a  $\tau$  tal que  $f(x) = 0$  y  $f(X \setminus U) = 1$ . Sabemos que  $\tau \subset \tau(A)$ , y por lo tanto  $f : X \rightarrow [0, 1]$  es continua relativa a  $\tau(A)$ .

Ahora como  $B = X \setminus (X \setminus B) \subset (X \setminus U)$ . Así,  $f(B) \subset f(X \setminus U) = 1$ . Entonces  $f : X \rightarrow [0, 1]$  es continua relativa a  $\tau(A)$  tal que  $f(x) = 0$  y  $f(B) = 1$ . Por lo tanto,  $(X, \tau(A))$  es completamente regular.

(2)  $x \in (V \cap A)$ . Sea  $W = A \setminus (V \cap A)$ , se tiene que  $W$  es cerrado en  $(A, \tau|_A)$  y  $x \notin W$ . Pero  $(A, \tau|_A)$  es completamente regular, entonces existe una función continua  $f : A \rightarrow [0, 1]$  relativa a  $\tau|_A$  tal que  $f(x) = 0$  y  $f(W) = 1$ . Ahora como  $(A, \tau|_A) = (A, \tau(A)|_A)$ , entonces  $f : A \rightarrow [0, 1]$  es continua relativa a  $\tau(A)|_A$ .

Ahora definimos una función  $f^* : X \rightarrow [0, 1]$  de la forma siguiente  $f^*(x) = f(x)$ , si  $x \in A$  y  $f^*(x) = 1$ , si  $x \in (X \setminus A)$ . Como  $f^*|(X \setminus A) = 1$  es continua relativa a  $\tau$ , entonces  $f^*|(X \setminus A) = 1$  es continua relativa a  $\tau(A)$ .

Como  $A, X \setminus A \in \tau(A)$ ,  $X = A \cup (X \setminus A)$  y  $f^*|_A, f^*|(X \setminus A)$  son continuas relativas a  $\tau(A)$ , entonces  $f^* : X \rightarrow [0, 1]$  es continua relativa a  $\tau(A)$ .

Por otro lado

$$f^*(B) = f^*(X \setminus (U \cup (V \cap A))) \subseteq f^*(X \setminus (V \cap A)) =$$

$$f^*((X \setminus A) \cup (A \setminus (V \cap A))) = f^*(X \setminus A) \cup f^*(W) = 1,$$

entonces  $f^*(B) = 1$ . Pero también se tiene que  $f^*(x) = f(x) = 0$ . Así, existe una función,  $f^* : X \rightarrow [0, 1]$  tal que  $f^*$  es continua relativa a  $\tau(A)$ ,  $f^*(B) = 1$  y  $f^*(x) = 0$ . Por lo tanto  $(X, \tau(A))$  es completamente regular.  $\square$

2.3. TEOREMA. Si  $(X, \tau)$  es un espacio de Tychonoff y  $X \setminus A \in \tau$ , entonces  $(X, \tau(A))$  es un espacio de Tychonoff.

DEMOSTRACIÓN. La prueba se sigue del teorema anterior y a causa de que si  $(X, \tau)$  es  $T_1$ , entonces  $(X, \tau(A))$  es  $T_1$ .  $\square$

2.4. TEOREMA. Sean  $(X, \tau)$  un espacio topológico  $T_4$  y  $A \subset X$  con  $A \notin \tau$  y  $X \setminus A \in \tau$ .  $(X, \tau(A))$  es  $T_4$  si y sólo si  $(X \setminus A, \tau|(X \setminus A))$  es  $T_4$ .

DEMOSTRACIÓN. (Necesidad). Supongamos que  $(X, \tau(A))$  es  $T_4$ . Como  $A \in \tau(A)$ , entonces  $X \setminus A$  es cerrado en  $X$  con la topología  $\tau(A)$ . Así,  $(X \setminus A, \tau(A)|(X \setminus A))$  es  $T_4$ . Sabemos que  $(X \setminus A, \tau|(X \setminus A)) = (X \setminus A, \tau(A)|(X \setminus A))$ . Así que  $(X \setminus A, \tau|(X \setminus A))$  es  $T_4$ .

(Suficiencia). Sean  $F$  y  $G$  dos subconjuntos cerrados y disjuntos en  $(X, \tau(A))$ , entonces  $F \cap A$  y  $G \cap A$  son cerrados y disjuntos en  $(A, \tau(A)|A)$ , pero como  $(A, \tau(A)|A) = (A, \tau|A)$ . Entonces  $F \cap A$  y  $G \cap A$  son cerrados y disjuntos en  $(A, \tau|A)$ .

Ahora, como  $F \cap A$  y  $G \cap A$  son cerrados y disjuntos en  $(X, \tau)$  y  $(X, \tau)$  es normal, entonces existen conjuntos abiertos y disjuntos  $U$  y  $V$  en  $(X, \tau)$ , tales que  $F \cap A \subset U$  y  $G \cap A \subset V$ .

Por otro lado  $F \cap (X \setminus A)$  y  $G \cap (X \setminus A)$  son cerrados y disjuntos en  $(X \setminus A, \tau(A)|(X \setminus A)) = (X \setminus A, \tau|(X \setminus A))$ , sabemos que  $(X \setminus A, \tau|(X \setminus A))$  es normal, entonces existen  $U'$  y  $V'$  abiertos y disjuntos en  $(X \setminus A, \tau|(X \setminus A))$  tales que  $F \cap (X \setminus A) \subset U'$  y  $G \cap (X \setminus A) \subset V'$ .

Sabemos que  $X \setminus A \in \tau(A)$ , entonces  $U'$  y  $V'$  son abiertos en  $(X, \tau(A))$ . Sean  $U^* = U' \cup (U \cap A)$  y  $V^* = V' \cup (V \cap A)$ . Es claro que  $U^*$  y  $V^*$  son abiertos y disjuntos en  $(X, \tau(A))$ . De aquí  $F = (F \cap A) \cup (F \cap (X \setminus A)) \subset (U \cap A) \cup U' = U^*$  y  $G = (G \cap A) \cup (G \cap (X \setminus A)) \subset (V \cap A) \cup V' = V^*$ . Así,  $U^*$  y  $V^*$  son conjuntos abiertos y disjuntos en  $(X, \tau(A))$  tales que  $F \subset U^*$  y  $G \subset V^*$ . Entonces  $(X, \tau(A))$  es normal. También se tiene que  $(X, \tau(A))$  es  $T_1$ , ya que  $(X, \tau)$  lo es. Por lo tanto  $(X, \tau(A))$  es  $T_4$   $\square$

2.5. TEOREMA. Sean  $(X, \tau)$  un espacio topológico y  $A \subset X$  con  $A \notin \tau$ . Si  $(A, \tau|A)$  es conexo y  $A$  es denso en  $(X, \tau)$ , entonces  $(X, \tau(A))$  es conexo

DEMOSTRACIÓN. Supongamos que  $(X, \tau(A))$  no es conexo, entonces existe una separación  $F \cup (F' \cap A)$  y  $G \cup (G' \cap A)$  en  $(X, \tau(A))$ . Es fácil probar que  $F \cup F' \neq \emptyset$  y  $G \cup G' \neq \emptyset$ . Ahora, como  $A$  es denso en  $(X, \tau)$ , entonces  $(F \cup F') \cap A \neq \emptyset$  y  $(G \cup G') \cap A \neq \emptyset$ .

Por otro lado  $((F \cup F') \cap A) \cap ((G \cup G') \cap A) = ((F \cap A) \cup (F' \cap A)) \cap ((G \cap A) \cup (G' \cap A)) \subset (F \cup (F' \cap A)) \cap (G \cup (G' \cap A)) = \emptyset$ . Así,  $((F \cup F') \cap A)$  y  $((G \cup G') \cap A)$  son disjuntos.

Ahora  $(F \cup F') \cup (G \cup G') = (F \cup (F' \cap A)) \cup (G \cup (G' \cap A)) \cup (F' \cup G') = X \cup (F' \cup G') = X$ , de aquí se tiene que  $X = (F \cup F') \cup (G \cup G')$ , entonces  $A = X \cap A = ((F \cup F') \cap A) \cup ((G \cup G') \cap A)$ . Por lo tanto  $((F \cup F') \cap A)$  y  $((G \cup G') \cap A)$  es una separación de  $(A, \tau|A)$  lo cual es una contradicción ya que  $(A, \tau|A)$  es conexo.  $\square$

#### REFERENCIAS

- [1] Engelking R., *General Topology*, Heldermann Verlag, Berlin 1989.
- [2] Hernández Rendón A., Sánchez Tesis F., Tesis Tesis M., Ramírez Páramo A., *Extensiones simples*, En revisión para su publicación en Topología y Sistemas Dinámicos II, Textos Científicos, BUAP.
- [3] Hrbacek K., Jech T., *Introduction to Set Theory*, Marcel Dekker, Inc., Second Edition 1978.
- [4] Levine N., *Simple extensions of topologies*, Amer. Math. 71 (1964), 22-25.

Facultad de Ciencias de la Electrónica, BUAP.

Av. San Claudio y 18 Sur, Col. San Manuel, Ciudad Universitaria

C.P. 72570

mtesis@ece.buap.mx

*Memorias de la Quinta Gran Semana Nacional de la Matemática*, edición a cargo de  
Juan Angoa Amador, Fernando Macías Romero y Jacobo Oliveros Oliveros  
se terminó de imprimir en el mes de marzo de 2010 en los talleres de  
El Errante Editor, S.A. de C.V., sito en Privada Emiliano Zapata 5947,  
Col. San Baltasar Campeche, Puebla, Pue.

El tiro consta de 300 ejemplares.





