Mixba´al

Revista Metropolitana de Matemáticas

ISSN en trámite

AÑO 1 VOL. 1 JUNIO 2010

Casa abierta al tiempo

ISSN En trámite
(Edición impresa)

Contacto

Posgrado en Matemáticas, Departamento de Matemáticas
Universidad Autónoma Metropolitana, Iztapalapa
Tel:(01) 55 5804 4657 Ext. 201 Fax: (01) 55 5804 4660
e-mail: mixbaal2009@gmail.com, e-mail: mixb@xanum.uam.mx
Web Revista: http://repos.izt.uam.mx/

Presentación

Mixba'al no es sólo una revista más de investigación en matemáticas. Se concibió con el propósito de ser específicamente una revista para los alumnos de posgrado de matemáticas en México, hecha por ellos mismos. Como el lector puede observar, la mayoría de los autores de los artículos en este primer número, son alumnos del Posgrado en Matemáticas de la Unidad Iztapalapa de la Universidad Autónoma Metropolitana y se ha hecho un esfuerzo para seleccionar aquellos que pueden ser leídos y entendidos por alumnos del último año de licenciatura o de una maestría en matemáticas. Entre los artículos de este número, se encontrarán tanto anuncios de investigación original como exposición de resultados importantes conocidos en varias ramas de la matemática básica y aplicada. La intención es continuar con este formato y la revista invita contribuciones en español de esta índole de la comunidad matemática nacional e internacional. Inicialmente se publicará un número al año en el mes de junio.

Toda comunicación relevante a la revista debe ser dirigida al Comité Editorial, al correo electrónico: mixbaal2009@gmail.com.mx.

A los autores

Mixba'al es una publicación del Posgrado en Matemáticas de la Universidad Autónoma Metropolitana, Unidad Iztapalapa, está dirigida a la comunidad académica matemática mexicana, formada tanto por estudiantes como por profesores de matemáticas. Esta revista surge de la necesidad de crear un espacio que permita al alumno desarrollar sus habilidades para escribir matemáticas. Busca fortalecer el Posgrado en Matemáticas y, en la medida de lo posible, la última etapa de la Licenciatura en Matemáticas. **Es importante señalar que la revista no pretende ser una revista internacional de investigación. El publicar un artículo en la revista no limita que el autor pueda enviar el artículo, o una variante del mismo, a otra revista .** Los artículos de exposición pueden ser trabajos que presenten de manera original algún tema de las matemáticas; por ejemplo, demostraciones nuevas de resultados conocidos, artículos panorámicos sobre un área de investigación, la presentación de una visión distinta de algún tema vinculado con la docencia, notas de cursos avanzados, aplicaciones de las matemáticas, historia y filosofía de las matemáticas y, aspectos lúdicos de la matemática, entre otros.

La revista Mixba'al está a cargo de un comité editorial y los coordinadores de los dos programas de Posgrado en Matemáticas. Los trabajos deben ser presentados en español, en casos excepcionales podrán aceptarse artículos en inglés. Todos los trabajos serán sometidos a revisión. El comité editorial decidirá sobre la aceptación de los artículos sometidos. Los artículos de investigación no necesariamente deben ser trabajos originales sobre investigaciones terminadas e incluir demostraciones. Dichos artículos pueden ser parte de memorias de congesos, simposios, coloquios, talleres u otros.

Cada volumen está a cargo del comité editorial de la revista. Dicho comité tiene la responsabilidad de garantizar su calidad, así como de su presentación de acuerdo con los lineamientos de formato y tipografía de la revista y de la corrección del lenguaje (ortografía, estilo, etcétera).

La versión preliminar de los trabajos sometidos a la revista deberá enviarse en formato pdf. Puesto que la presentación final de los trabajos se hará en Latex2ε, aquellos autores, cuyos trabajos sean

aceptados, deberan enviarlos escritos en este sistema de preparación de textos, con el formato y macros que proporcionará la revista para su publicación final. Las fotografías o gráficas que acompañen al texto deberán ser enviadas, por separado, en formato pdf y deberán tener la calidad y resolución suficientes para una buena reproducción impresa. Se recomienda que la extensión de los trabajos no exceda de 20 páginas.

Laura Hidalgo Solís
Coordinadora

4

.

Mixba'al Rev. Met. de Mat.
Nº 1, Vol. I, Junio 2010. 5 – 9

In memoriam: Peter Seibert (1927-2009)

Luis Aguirre Castillo

Peter Seibert Kopp, nació el 13 de abril de 1927 en Múnich, Baviera, Alemania, y murió en México el 13 de agosto de 2009.

Fué alumno de Constantin Carathéodory durante su formación matemática preuniversitaria, durante su formación universitaria lo fué de E. Hopf, H. Tietze y A. Pfluger, entre otros. Realizó su tesis doctoral bajo la dirección de R. König, alumno de D. Hilbert, con O. Perron, como "segundo referente". El 4 de agosto de 1950 recibió el grado de Doctor en Ciencias por la Universidad de Múnich y en julio de 2000 el Doctorado de Oro por la misma universidad.

Tuvo sus primeras posiciones académicas y de investigación en Alemania, en las universidades de Giessen, Wuerzburg y en el Centro de Investigación de la Universidad de Wuerzburg.

Inicialmente Peter se especializó en la teoría de funciones de variable compleja, con énfasis en las superficies de Riemann. Entre otros resultados, resolvió un problema propuesto por S. Mazurkiewicz, dando el primer ejemplo de una superficie de Riemann con frontera bidimensional [1].

En 1958 reorientó su investigación hacia la teoría de estabilidad en sistemas dinámicos, después de que en febrero de ese mismo año se entrevistó en París con Solomon Lefchetz, quién lo invitó a trabajar en el Research Institute for Advanced Studies (RIAS), Baltimore, USA., 1958 y 1963.

Viajó por primera vez a México en septiembre de 1959, para participar en el *Symposium de Ecuaciones Diferenciales Ordinarias y sus Aplicaciones*, que tuvo lugar en la Ciudad Universitaria de la UNAM, del 7 al 13 de septiembre. Visitó México por segunda ocasión, de 1963 a 1964 para trabajar en el Departamento de Matemáticas del CINVESTAV -IPN, a invitación de José Adem. Volvió a México para trabajar en la Escuela Superior de Física y Matemáticas del IPN y el Instituto Mexicano del Petróleo en el periodo 1966-1971. Trabajó en Sudamérica de 1971 a 1985 en donde fué director de las tesis doctorales de Pablo M. Salzberg (1975) en Chile y de Ramón Gómez (1985) en Venezuela. Regresó a México por cuarta vez, ahora al Departamento de Matemáticas de la Universidad Autónoma Metropolitana-Iztapalapa en 1987, donde dirigió la tesis doctoral de Luis Aguirre Castillo (2003).

En colaboración con J. André, Peter desarrolló la primera teoría matemática de sistemas de control discontinuos [2]. El concepto de robustez también apareció por primera vez en este contexto en [3].

Usando un argumento de inducción, dió la primera demostración del principio de persistencia de la estabilidad asintótica bajo perturbaciones [4]. Este resultado ha sido una herramienta útil para cimentar una teoría más general de bifurcaciones del tipo Hopf.

Con J. Auslander, desarrolló la teoría de la estabilidad de sistemas dinámicos combinando la teoría de las prolongaciones de Poincaré con la teoría clásica de A.M. Lyapunov [5]. En colaboración con P.Tulley, dió la primera demostración puramente topológica del teorema de Poincaré-Bendixson sobre sistemas dinámicos en el plano [6]. También dió una condición necesaria y suficiente para la estabilidad bajo perturbaciones persistentes, complementando trabajos anteriores de I.G. Malkin y otros [7].

Parcialmente, en colaboración con su estudiante de doctorado, P.M. Salzberg y G. Dankert, Seibert estableció una teoría abstracta de la estabilidad de Lyapunov, culminando con una condición necesaria y suficiente para la existencia de una función de Lyapunov (generalizada) en presencia de un atractor [8].

En 1968, en respuesta a un problema propuesto por J.S. Florio, Peter inventó un método para reducir el problema de la estabilidad de un sistema compuesto de subsistemas al problema correspondiente para los subsistemas, aplicando un principio del umbral, [9],[10].

En trabajos recientes, aparecen en el *resumen* expresiones como "sistemas generales en el sentido de Seibert". También se usa la notación inventada por él.

Peter, sus colaboradores y estudiantes, desarrollaron dos teorías de bifurcaciones en sistemas dinámicos que generalizan la bifurcación clásica de (Poincaré-Andronov-) Hopf. Para un resumen de estos resultados ver [9]. Colaboró con Juan Héctor Arredondo Ruíz, en una nueva línea de investigación en la que estamos estudiando la aplicación de su teoría de reducción a sistemas de ecuaciones diferenciales parciales acopladas.

Peter demostró que se puede hacer mucho desarrollando teorías abstractas mediante el uso de la topología elemental de espacios métricos, que se aplican a problemas concretos donde otros métodos fallan. Véanse, por ejemplo, las referencias [12] y [13].

Los trabajos de Peter aparecieron en revistas de 11 paises en cinco idiomas que él dominó: Alemán, Inglés, Francés, Ruso y Español. Han tenido un fuerte impacto en Rusia, particularmente en la escuela de control de Moscú e Italia. Su influencia se extiende hasta China en donde han retomado el estudio de los conjuntos límite prolongacionales, introducidos por Seibert.

Además de sus trabajos en matemáticas se ocupó de dos aspectos relevantes en la lingüistica. Por un lado la construcción del primer lenguaje completamente sintético, WUXI, [14]. Y por otra parte, desarrolló un método fonemoestadístico para la comparación de lenguajes. Además Peter Seibert era aficionado al alpinismo, la fotografía de paisajes y la fotocomposición.

Su mayor contribución a la matemática, de acuerdo con su propia opinión, es el principio general de reducción vía el principio del umbral para la estabilidad en sistemas semidinámicos [9]. Hasta hoy sus trabajos tienen más de doscientas cuarenta citas. Peter, fué un hombre sencillo, humanista, reacio a la adulación. Vivió solitario, como un lobo estepario, descanse en paz.

Referencias

[1] On a problem of Mazurkiewicz concerning the boundary of a covering surface, Proc. Nat. Acad. Sci., U.S.A., **45**, 1959, 50-54.

[2] Über stückweise lineare Differentialgleichungen, die bei Regelungsproblemen auftreten. I. & II. Arch. Math. (Basel) **7** , 1956,148–156 & 157–164. (con J. André).

[3] After end-point motions of general discontinuous control systems and their stability properties, Automat. and Remote Control, Proc. Int. Congr. IFAC, Moscú, 1960; Butterworth, 1962, 919-922 (con J. André).

[4] Prolongations and generalized Liapunov functions, Nonlin. Diff. Eqs. Nonlin. Mech., Proc. Int. Sympos., Colorado Springs, 1961; Acad. Press, 1963, 454-462 (con J. Auslander).

[5] Prolongations and stability in dynamical systems, Ann. Inst. Fourier (Grenoble) **14** (2), 1964, 237-267 (con J. Auslander).

[6] On dynamical Systems in the plane, Archiv. d. Math. **18**, 1967, 290-292 (con P. Tulley).

[7] Estabilidad bajo perturbaciones y su generalización en sistemas dinámicos, Acta Mexicana Ci. Tec. **2**, 1968, 154-165.

[8] On the existence of Liapunov functions in general systems, Nonlinear Phenomena in Math. Sciences (Proc. Int. Conf., Arlington, Texas, 1980, Acad. Press, 1982, 917-926.

[9] On the Reduction to a Subspace of Stability Properties of Systems in Metric Spaces, Annali di Matematica pura ed applicata (IV), **CLXIX**, 1965, 291-320 (con J.S. Florio).

[10] Reduction theorems for uniform stability of systems in general spaces, Bol. Soc. Mat. Mexicana **3** (3), 69-88, 1997 (con J.H. Arredondo, J. Delgado y L.Aguirre).

[11] Dynamical persistence principles and bifurcations, Univ. Jagellon. Acta Math. **36** (3), 1998, 65-77 .

[12] Global Stabilization of nonlinear cascade systems, Systems and Contro Letters **14**, 1990, 347-352 (con R. Suárez).

[13] Global stabilization of a certain class of nonlinear Systems, Systems and Control Letters **14**, 1990, 17-23 (con R. Suárez).

[14] WUXI, Proyecto de una lengua sintética universal, Cemanáhuac **5**, 1992, 12-17.

Dirección del autor
Luis Aguirre Castillo
Universidad Autónoma Metropolitana,
Unidad Iztapalapa,
División de Ciencias Básicas e Ingeniería,
Departamento de Matemáticas.
Av. San Rafael Atlixco 186, Col. Vicentina
Del. Iztapalapa, C.P. 09340 México, D.F.
e-mail: `lac@xanum.uam.mx`

UAM-Iztapalapa, 25 de marzo 2010

.

Mixba'al Rev. Met. de Mat.
Nº 1, Vol. I, Junio 2010. 11 – 22

El Teorema de Hermite-Biehler en sistemas continuos, discretos y con retardo

Edgar Cristian Díaz González Baltazar Aguirre Hernández

Resumen

Se sabe que el estudio de la estabilidad de un sistema lineal continuo o discreto esta determinado por el análisis del correspondiente polinomio característico. Uno de los resultados más importantes para verificar dicha estabilidad es el Teorema de Hermite-Biehler. Existe un correspondiente Teorema de Hermite-Biehler para analizar la estabilidad de sistemas continuos y sistemas discretos. En el caso de sistemas con retardo la ecuación característica es un cuasipolinomio existe también una versión del Teorema de Hermite-Biehler para este caso. En este artículo presentamos las tres versiones de este teorema.

Palabras clave: polinomios de Hurwitz, polinomios de Schur, cuasipolinomios estables, Teorema de Hermite-Biehler.

Clasificación de la AMS: 93D09, 34D99

1. Introducción

En el análisis de la estabilidad asintótica de un sistema continuo (o discreto) es necesario que todas las raíces de su polinomio característico asociado se encuentren en $\mathbb{C}^-$ (o en $\mathbb{D}$), donde $\mathbb{C}^-$ es el conjunto de números complejos que tienen parte real negativa y $\mathbb{D}$ es el conjunto de números complejos con módulo menor que 1. En el caso de que un polinomio tenga todas sus raíces en $\mathbb{C}^-$ se dice que es un polinomio de Hurwitz y en el caso en que todas sus raíces estén en $\mathbb{D}$ se dice que es un polinomio de Schur. En el estudio de la distribución de las raíces de un polinomio sobre el plano complejo, uno de los primeros problemas fue, históricamente, el de determinar el número

de raíces reales de una ecuación; esto es, dada una ecuación con coeficientes reales, determinar por algún criterio, que dependerá de sus coeficientes, y sin resolver la ecuación, si tiene raíces reales. En caso afirmativo, cuántas raíces positivas y cuántas negativas tiene. En 1868, *Maxwell* plantea el problema matemático de la búsqueda de condiciones bajo las cuales todas las raíces de una ecuación algebraica se encuentren en $\mathbb{C}^-$, muchos matemáticos ya se habían ocupado de la determinación del número de raíces de ecuaciones algebraicas en determinados lugares (dentro y fuera del eje real, en la mitad del plano, etc) desde las primeras décadas del siglo XIX, tales como, *Cauchy*, *Sturm*, *Jacobi*, *Cayley*, *Sylvester*, y *Hermite*. De hecho, *Hermite* (1853), ya había resuelto el problema de *Maxwell*, pero sus resultados no eran conocidos fuera del mundo de las matemáticas. Los trabajos de *Routh* y *Hurwitz* dieron lugar al Criterio de $Routh - Hurwitz$. Además del Criterio de $Routh - Hurwitz$, existen otros criterios para determinar si un polinomio es Hurwitz, mencionemos, por ejemplo, las Condiciones de $Lienard - Chipart$, el Test de Estabilidad o el Teorema de $Hermite - Biehler$. Desde el punto de vista matemático estos teoremas tienen la misma dificultad pues son resultados equivalentes, aunque probablemente el Criterio de $Routh - Hurwitz$ es el más popular. Sin embargo, el Teorema de $Hermite - Biehler$ ha demostrado su potencial al ser utilizado en el estudio de la estabilidad en familias de polinomios, basta mencionar el Teorema de *Kharitonov*, ver [5].

2. El Teorema de Hermite-Biehler: caso continuo

En el análisis de la estabilidad de un sistema de ecuaciones diferenciales $\dot{x}(t) = f(x(t))$, podemos estudiar la parte lineal del sistema, $\dot{x}(t) = Ax(t)$. Bajo ciertas condiciones, la estabilidad del sistema linealizado implica la estabilidad local del sistema original. Es conocido que la estabilidad de un sistema lineal es verificada por medio del polinomio característico asociado a la matriz A.

Definición 2.1. *Sea* $P(\lambda) = \det(\lambda I - A)$*, el polinomio característico de* A*.*

Así, el problema de determinar condiciones, bajo las cuales, todas las raíces del polinomio característico se encuentren en el semiplano abierto izquierdo es de importancia fundamental en el estudio de la estabilidad del sistema lineal.

Definición 2.2. *Consideremos el polinomio real*

$$p(s) = a_0 + a_1 s + a_2 s^2 + \ldots + a_n s^n. \tag{1}$$

Podemos escribir el polinomio $p(s)$ de la siguiente forma:

$$p(s) = (a_0 + a_2 s^2 + a_4 s^4 + \ldots) + s(a_1 + a_3 s^2 + a_5 s^4 + \ldots), \tag{2}$$

evaluando el polinomio $p(s)$ en $s = i\omega$, tenemos que

$$p(i\omega) = (a_0 - a_2\omega^2 + a_4\omega^4 - \ldots) + i\ \omega(a_1 - a_3\omega^2 + a_5\omega^4 - \ldots) \tag{3}$$

y definimos los siguientes polinomios

$$\begin{aligned} p^e(\omega) &= a_0 - a_2\omega^2 + a_4\omega^4 - \ldots \\ p^o(\omega) &= a_1 - a_3\omega^2 + a_5\omega^4 - \ldots \\ p^{par}(s) &= a_0 + a_2 s^2 + a_4 s^4 + \ldots \\ p^{imp}(s) &= a_1 s + a_3 s^3 + a_5 s^5 + \ldots \end{aligned}$$

Ejemplo 2.3.

Consideremos los siguientes polinomios

$$\begin{aligned} q(s) &= 4 + s + 2s^2 + 5s^3 + 4s^4 + s^5 \quad \text{y} \\ r(s) &= 2 + 5s + 3s^2 + s^3 + 4s^4 + 2s^5 + s^6. \end{aligned}$$

Los polinomios asociados a $q(s)$ son

$$\begin{aligned} q^e(\omega) &= 4 - 2\omega^2 + 4\omega^4 \\ q^o(\omega) &= 1 - 5\omega^2 + \omega^4 \\ q^{par}(s) &= 4 + 2s^2 + 4s^4 \\ q^{imp}(s) &= s + 5s^3 + s^5 \end{aligned}$$

para el polinomio $r(s)$ tenemos que

$$\begin{aligned} r^e(\omega) &= 2 - 3\omega^2 + 4\omega^4 - \omega^6 \\ r^o(\omega) &= 5 - \omega^2 + 2\omega^4 \\ r^{par}(s) &= 2 + 3s^2 + 4s^4 + s^6 \\ r^{imp}(s) &= 5s + s^3 + 2s^5. \end{aligned}$$

Definición 2.4. *El polinomio* $p(s) = a_0 + a_1 s + a_2 s^2 + \ldots + a_n s^n$ *satisface la propiedad de la alternancia, si y sólo si*

a) los coeficientes principales de $p^{par}(s)$ *y* $p^{imp}(s)$ *tienen el mismo signo;*

b) todas las raíces de $p^e(\omega)$, $p^o(\omega)$ *son reales y las raíces positivas de* $p^e(\omega)$, $p^o(\omega)$ *se van alternando, es decir*

$$0 < \omega_{e,1} < \omega_{o,1} < \omega_{e,2} < \omega_{o,2} < \ldots \tag{4}$$

A continuación el teorema principal de esta sección.

Teorema 2.5. (Hermite-Biehler). *Un polinomio real P(s) es de Hurwitz, si y sólo si, satisface la propiedad de la alternancia.*

Ver [2] y [7] para una demostración. Concluimos esta sección presentando el siguiente ejemplo que ilustra el Teorema 2.1.

Ejemplo 2.6.

Verificar si el siguiente polinomio es de Hurwitz.

$$\begin{aligned} p(s) &= 147 + 574s + 1363s^2 + 2103s^3 + 2402s^4 + 2015s^5 + 1293s^6 \\ &\quad +614s^7 + 221s^8 + 57s^9 + 10s^{10} + s^{11} \end{aligned}$$

Calculamos

$$\begin{aligned} p(i\omega) &= 147 + 574i\omega - 1363\omega^2 - 2103i\omega^3 + 2402\omega^4 + 2015i\omega^5 \\ &\quad -1293\omega^6 - 614i\omega^7 + 221\omega^8 + 57i\omega^9 - 10\omega^{10} - i\omega^{11}, \end{aligned}$$

así, tenemos que

$$\begin{aligned} p^e(\omega) &= 147 - 1363\omega^2 + 2402\omega^4 - 1293\omega^6 + 221\omega^8 - 10\omega^{10} \\ p^o(\omega) &= 574 - 2103\omega^2 + 2015\omega^4 - 614\omega^6 + 57\omega^8 - \omega^{10}. \end{aligned}$$

Ahora

$$\begin{aligned} p^{par}(\omega) &= 147 + 1363\omega^2 + 2402\omega^4 + 1293\omega^6 + 221\omega^8 + 10\omega^{10} \\ p^{imp}(\omega) &= 574\omega + 2103\omega^3 + 2015\omega^5 + 614\omega^7 + 57\omega^9 + \omega^{11}. \end{aligned}$$

Así, el inciso a) de la propiedad de la alternancia se satisface. Ahora verificamos el inciso b).

$$\begin{aligned} p^e(\omega) = 0 &\iff \omega = \pm 0.373, \pm 0.868, \pm 1.375, \pm 2.287, \pm 3.752 \\ p^o(\omega) = 0 &\iff \omega = \pm 0.65, \pm 1.088, \pm 1.788, \pm 2.847, \pm 6.639 \end{aligned}$$

agrupando las raíces tenemos que

$$\omega_{e,1} = 0.373, \omega_{e,2} = 0.868, \omega_{e,3} = 1.375, \omega_{e,4} = 2.287, \omega_{e,5} = 3.752$$

y

$$\omega_{o,1} = 0.65, \omega_{o,2} = 1.088, \omega_{o,3} = 1.788, \omega_{0,4} = 2.847, \omega_{o,5} = 6.639.$$

Entonces tenemos que

$$\omega_{e,1} < \omega_{o,1} < \omega_{e,2} < \omega_{o,2} < \omega_{e,3} < \omega_{o,3} < \omega_{e,4} < \omega_{o,4} < \omega_{e,5} < \omega_{o,5}$$

satisfaciéndose así el inciso b) de la propiedad de la alternancia. Por el teorema 2.1, $p(s)$ es de Hurwitz.

3. El Teorema de Hermite-Biehler: caso discreto

Es posible obtener un teorema de alternancia con respecto a $\mathbb{D}$, que es la región de estabilidad para sistemas discretos, $x(t+1) = Ax(t)$. En este caso, la estabilidad del polinomio característico asociado al sistema discreto es equivalente a la alternancia de las raíces, de los polinomios asociados a este, a lo largo del círculo unitario.

Definición 3.1. *Un polinomio*

$$P(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$$

es un polinomio de Schur si todas sus raíces se encuentran en el círculo unitario abierto $\mathbb{D}$ del plano complejo.

Una condición necesaria para la estabilidad Schur es que $|a_n| > |a_0|$.

Definición 3.2. *Sean los polinomios $P_s(z)$ y $P_a(z)$ como sigue:*

$$P_s(z) = \frac{1}{2}[P(z) + z^n P(\frac{1}{z})] \quad y \quad P_a(z) = \frac{1}{2}[P(z) - z^n P(\frac{1}{z})]$$

definimos $P(z) = P_s(z) + P_a(z)$.

Para un polinomio real, la estabilidad de $P(z)$ es equivalente a la alternancia, en el círculo unitario, de las raíces de los polinomios $P_s(z)$ y $P_a(z)$.

Teorema 3.3. *Un polinomio real $P(z)$ es de Schur, si y sólo si, $P_s(z)$ y $P_a(z)$ satisfacen las afirmaciones siguientes*

a) *$P_s(z)$ y $P_a(z)$ son polinomios de grado n con coeficientes principales del mismo signo.*

b) *$P_s(z)$ y $P_a(z)$ tienen sólo ceros simples, los cuales están en el círculo unitario.*

c) *Los ceros de $P_s(z)$ y $P_a(z)$ se alternan en el círculo unitario.*

Ver [2] para una demostración. Ahora presentamos el siguiente ejemplo que ilustra el Teorema 2.2.

Ejemplo 3.4.

Verificar si el polinomio $P(z)$ es un polinomio de Schur

$$P(z) = z^5 + 0.2z^4 + 0.3z^3 + 0.4z^2 + 0.03z + 0.02$$

utilizando la definición 3.2, obtenemos que

$$P_s(z) = 0.51z^5 + 0.115z^4 + 0.35z^3 + 0.35z^2 + 0.115z + 0.51 \text{ y}$$

$$P_a(z) = 0.49z^5 + 0.085z^4 - 0.05z^3 + 0.05z^2 - 0.085z - 0.49.$$

Así, tenemos que el inciso a) se satisface. Ahora verificamos los incisos b) y c).

$$P_s(z) = 0 \iff z = -0.221 \pm 0.975i, 0.608 \pm 0.793i, -1$$
$$P_a(z) = 0 \iff z = -0.857 \pm 0.514i, 0.270 \pm 0.962i, 1$$

Cada una de las raíces de $P_s(z)$ y $P_a(z)$ son simples, se encuentran en el círculo unitario y se alternan. Por lo tanto, por el Teorema 2.2, $P(z)$ es un polinomio de Schur.

Ejemplo 3.5.

Verificar si el polinomio $P(z)$ es un polinomio de Schur

$$P(z) = z^5 + 2z^4 + 0.3z^3 + 0.4z^2 + 0.03z + 0.02$$

de la definición 3.2, obtenemos los siguientes polinomios

$$P_s(z) = 0.51z^5 + 1.015z^4 + 0.35z^3 + 0.35z^2 + 1.015z + 0.51$$
$$P_a(z) = 0.49z^5 + 0.985z^4 - 0.05z^3 + 0.05z^2 - 0.985z - 0.49.$$

Así, el inciso a) se satisface. Ahora verificamos el inciso b) y c).

$$P_s(z) = 0 \iff z = 0.55 \pm 0.83i, -1.35, -1, -0.74.$$
$$P_a(z) = 0 \iff z = -0.17 \pm 0.98i, -2.21, -0.45, 1.$$

Cada una de las raíces de $P_s(z)$ y $P_a(z)$ son simples, pero no se encuentran en el círculo unitario ni se alternan. Por lo tanto, por el Teorema 2.2, $P(z)$ no es un polinomio de Schur.

4. El Teorema de Hermite-Biehler: sistemas con retardo

Las ecuaciones diferenciales con retardo (EDR) son ecuaciones diferenciales en las cuales la función f depende también de términos con diferentes valores de su argumento, $\dot{x}(t) = f(t, x(t), x(t - \tau(t)))$, donde $\tau(t) > 0$, la primera ecuación de este tipo apareció en la literatura en la segunda mitad del siglo $XVIII$ ($Kondorse$, 1771), pero un estudio sistemático de las ecuaciones con retardo comenzó en el siglo XX (especialmente en las primeras décadas - $A.D.Myshkis$, en la Unión Soviética, $E.M.Wright$ y $R.Bellman$ en otros países), en relación con las necesidades de la ciencia aplicada. Las EDR tienen muchas aplicaciones en la teoría del control automático, el estudio de problemas relacionados con la combustión de cohetes en movimiento, en economía, en una serie de problemas biológicos, etc. En un estudio en sistemas reales, para una primera aproximación, se supone que el retraso es constante, $\tau(t) = \tau$. Para más información sobre sistemas con retardo ver [1], [3], [9] y [11].

Consideremos la siguiente ecuación diferencial de segundo orden con dos retardos

$$\ddot{y}(t) + a_1\, \dot{y}(t - \tau_1) + a_0\, y(t - \tau_0) = u(t). \tag{5}$$

La ecuación (5) puede ser representada en la siguiente forma matricial

$$\begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ -a_0 & 0 \end{pmatrix} \begin{pmatrix} x_1(t - \tau_0) \\ x_2(t - \tau_0) \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & -a_1 \end{pmatrix} \begin{pmatrix} x_1(t - \tau_1) \\ x_2(t - \tau_1) \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(t). \tag{6}$$

En general, un sistema lineal con l distintos retardos, $\tau_1, \ldots, \tau_l$, puede ser representado de la siguiente forma

$$\dot{x} = A_0\, x(t) + \sum_{i=1}^{l} A_i\, x(t - \tau_i) + B\, u(t). \tag{7}$$

Para discutir la estabilidad de la ecuación (5) ó los sistemas (6) y (7) es común examinar las soluciones $y(t)$, con $u(t) \equiv 0$ y estudiar el comportamiento de $y(t)$, cuando $t \to \infty$. Para este propósito, consideremos, por ejemplo, la ecuación (5) con $u(t) \equiv 0$ y que $y(t) = e^{s\,t}$ es una solución de

$$\ddot{y}(t) + a_1\,\dot{y}(t - \tau_1) + a_0\,y(t - \tau_0) \equiv 0. \tag{8}$$

Por lo tanto, tenemos que

$$(s^2 + a_1\,e^{-s\,\tau_1}s + a_0\,e^{-s\,\tau_0})e^{s\,t} \equiv 0,$$

de manera que s debe satisfacer la ecuación

$$s^2 + a_1 e^{-s\,\tau_1}s + a_0 e^{-s\,\tau_0} = 0. \tag{9}$$

La ecuación (9) es la ecuación característica de (5) u (8); la ubicación de sus raíces (o ceros) determina la estabilidad de la ecuación (5).

En las secciones anteriores se mostró el correspondiente Teorema de Hermite-Biehler para sistemas continuos y discretos, sin embargo, cuando el sistema involucra retardos no se pueden aplicar los teoremas dados. Sistemas con retardos dan lugar a funciones características, conocidas como cuasipolinomios, la ecuación característica asociada con (7) es

$$\begin{aligned} P^*(s) &= \det\left(sI - A_0 - \sum_{i=1}^{l} e^{-s\tau_i}A_i\right) \\ &= P_0(s) + \sum_{k=1}^{m} P_k(s)e^{-L_k s}. \end{aligned} \tag{10}$$

En la ecuación (10) los retardos pueden ser múltiplos enteros de un número positivo τ. En estos casos, se dice que los retardos son conmensurables y la ecuación característica toma la forma

$$P^*(s) = a_0(s) + a_1(s)e^{-\tau s} + a_2(s)e^{-2\tau s} + \cdots + a_k(s)e^{-k\tau s}, \tag{11}$$

donde los $a_i(s), i = 0, 1, \ldots, k$ son polinomios.

Definición 4.1. *Sea $P(s,t)$ un polinomio en dos variables con coeficientes reales o complejos de la siguiente forma*

$$\sum_k \sum_i a_{ik} s^i t^k = P(s,t).$$

Definición 4.2. *Sea r el grado más grande en s, p el grado más grande en t.*

$P(s,t)$ tiene un término principal, si y sólo si, $a_{rp} \neq 0$.

L.S. Pontrjagin en 1942 generalizó el Teorema de Hermite-Biehler para cuasipolinomios, $P(s, e^s)$.

Definición 4.3. *Sea f(s) una función entera de la forma*

$$\sum_{k=1}^{n} e^{\lambda_k s} P_k(s) = f(s)$$

donde $P_k(s)$ para k = 1,$\cdots$,n es un polinomio arbitrario con coeficientes reales ó complejos y los λ_k's son reales, los cuales satisfacen:

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n, \ |\lambda_1| < \lambda_n \tag{12}$$

A continuación el teorema principal de esta sección.

Teorema 4.4. *Sea $f(s) = P(s, e^s)$, donde $P(s,t)$ es un polinomio con un término principal, además*

$$f(j\omega) = f^r(\omega) + jf^i(\omega) \tag{13}$$

donde $f^r(\omega)$ y $jf^i(\omega)$ representan respectivamente la parte real e imaginaria de $f(j\omega)$. Entonces para que f(s) tenga todos sus ceros en el semiplano abierto izquierdo es necesario y suficiente que las siguientes dos condiciones se tengan:

a) *$f^r(\omega)$ y $f^i(\omega)$ tengan solo raíces simples y estas raíces se alternen.*

b) *Para todo ω en $\mathbb{R}$, $f^{i'}(\omega)f^r(\omega) - f^i(\omega)f^{r'}(\omega) > 0$.*

Ver este resultado en [2], [10] y en [11]. En la condición b) el símbolo $'$ indica derivación con respecto a ω, esta condición puede ser reemplazada por la condición más débil:

$$\acute{\mathbf{b}}) \ f^{i'}(\omega)f^r(\omega) - f^i(\omega)f^{r'}(\omega) > 0, \quad \text{para algún} \quad \omega_0 \in \mathbb{R} \tag{14}$$

Por último presentamos el siguiente ejemplo que ilustra el Teorema 3.5.

Ejemplo 4.5.

Verificar si el cuasipolinomio P(s) es Hurwitz estable

$$P(s) = d(s) + e^{-sT_1} n_1(s) + e^{-sT_2} n_2(s)$$

donde

$$\begin{aligned} d(s) &= 1 + 15s + 50s^2 + 100s^3 + 100s^4 + 200s^5 + 100s^6 + 20s^7 \\ &\quad +5s^8 + s^9 \\ n_1(s) &= 2 + 10s + 50s^2 + 50s^3 + 100s^4 + 15s^5 + 10s^6 + 10s^7 + 3s^8 \\ n_2(s) &= 3 + 24s + 55s^2 + 130s^3 + 131s^4 + 51s^5 + 35s^6 + 22s^7 + 2s^8. \end{aligned}$$

Con $T_1 = \frac{1}{10}, T_2 = \frac{3}{10}$, escribimos

$$d(j\omega) = d^e(\omega) + j\omega\ d^o(\omega) \qquad n_1(j\omega) = n_1^e(\omega) + j\omega\ n_1^o(\omega)$$

$$n_2(j\omega) = n_2^e(\omega) + j\omega\ n_2^o(\omega) \qquad P(j\omega) = P_r(\omega) + j\ P_i(\omega)$$

Así tenemos que

$$\begin{aligned} P_r(\omega) &= d^e(\omega) + cos(\omega T_1)\ n_1^e(\omega) - \omega\ sen(\omega T_1)\ n_1^o(\omega) \\ &\quad +cos(\omega T_2)\ n_2^e(\omega) - \omega sen(\omega T_2)\ n_2^o(\omega) \\ P_i(\omega) &= \omega\ d^o(\omega) + \omega\ cos(\omega T_1)\ n_1^o(\omega) - sen(\omega T_1)\ n_1^e(\omega) \\ &\quad +\omega\ cos(\omega T_2)\ n_2^o(\omega)\ - sen(\omega T_2)\ n_2^e(\omega). \end{aligned}$$

Las raíces de $P_r(\omega)$ y $P_i(\omega)$ se alternan, por lo tanto, por el teorema 3.5, el cuasipolinomio es Hurwitz.

Referencias

[1] Bellman, R. E., Cooke, K. L., *Differential-Difference Equations*, Number 6 in Mathematics in Science and Engineering, Academic Press, New York, 1963.

[2] Bhattacharayya, S.P., Chapellat, H., Keel, L.H., *Robust Control: The Parametric Approach*, Prentice-Hall, Upper Saddle River, New Jersey, 1995.

[3] Gu, K., Kharitonov, V. L. and Chen, J., *Stability of Time-Delay Systems*, Birkhäuser, Boston, 2003.

[4] Hermite, C., *Sur le nombre des racines dune équation algébrique comprise entre des limites donnés*, J. Reine Angew. Math., **52**, 1856, 39-51.

[5] Kharitonov, V.L., *Asymptotic stability of an equilibrium position of a family of systems of linear differential equations*, Differentsial'nye Uravnenija, **14**, 1978, p. 2086-2088. Translation in Differential Equations, **14**, 1979, 1483-1485.

[6] LaSalle, J.P., *The Stability and Control of Discrete Processes*, New York Inc., Springer-Verlag, New York, 1986.

[7] Loredo, C.A., *Criterios para determinar si un polinomio es polinomio Hurwitz*. Reporte de los seminarios de investigación I y II, UAM-Iztapalapa, México, D.F. 2004.

[8] Maxwell, J. C., *On governors*, Proc. Royal Soc. London, **16**, 1868, 270-283.

[9] Myshkis, A. D., *General theory of differential equations with delay*, Engl. Transl. AMS **55**, 1951, 1-62 .

[10] Pontrjagin, L.S., *On the zeros of some elementary transcendental functions*, Akad. Nauk SSSR Ser. Mat., **6**, 1942, 115-134. English translation, American Mathematical Society Translations, **2**, 1955, 95-110 .

[11] Silva, G.J., Datta, A., Bhattacharayya, S.P., *PID Controllers for Time-Delay Systems*, Boston, Birkhäuser, 2005.

[12] Vyshnegradsky, I. A., *Maxwell, I. A. Vyshnegradsky, and A. Stodola. Theory of Automatic Control, On the Direct Action Regulators*, M., AN SSSR, 1949, In D. K.(in Russian).

Dirección de los autores
Edgar Cristian Díaz González
Universidad Autónoma Metropolitana,
Unidad Iztapalapa,
División de Ciencias Básicas e Ingeniería,
Departamento de Matemáticas.
Av. San Rafael Atlixco 186, Col. Vicentina
Del. Iztapalapa, C.P. 09340 México, D.F.

e-mail: edgardazgonzlez@yahoo.com.mx

Baltazar Aguirre Hernández
Universidad Autónoma Metropolitana,
Unidad Iztapalapa,
División de Ciencias Básicas e Ingeniería,
Departamento de Matemáticas.
Av. San Rafael Atlixco 186, Col. Vicentina
Del. Iztapalapa, C.P. 09340 México, D.F.
e-mail: bahe@xanum.uam.mx

Mixba'al Rev. Met. de Mat.
Nº 1, Vol. I, Junio 2010. 23 – 40

Constructibilidad relativizada y el Axioma de Elección

Franklin C. Galindo Carlos A. Di Prisco

Resumen

El objetivo de este trabajo es presentar en un solo cuerpo tres maneras de relativizar (o generalizar) el concepto de conjunto constructible de Gödel que no suelen aparecer juntas en la literatura especializada y que son importantes en la Teoría de Conjuntos, por ejemplo para resolver problemas de consistencia o independencia. Presentamos algunos modelos resultantes de las diferentes formas de relativizar el concepto de constructibilidad, sus propiedades básicas y algunas formas débiles del Axioma de Elección válidas o no válidas en ellas.

Palabras clave: Constructibilidad, Axioma de Elección, Forcing.

Clasificación de la AMS: Primary: 03E45, Secondary: 03E25, 03E35

1. Introducción

El objetivo de este trabajo es presentar en un solo cuerpo tres maneras de relativizar (o generalizar) el concepto de conjunto constructible de Gödel que no suelen aparecer juntas en la literatura especializada y que son importantes en la Teoría de Conjuntos, por ejemplo para resolver problemas de consistencia o independencia. Presentamos los modelos resultantes de las diferentes formas de relativizar el concepto de constructibilidad, sus propiedades básicas y algunas formas débiles del Axioma de Elección válidas o no válidas en ellas. El orden de la exposición es el siguiente: En la sección 2 presentamos algunos conceptos preliminares y la clase de los conjuntos constructibles de Gödel (L), la cual es un modelo transitivo de ZFC[1] más la Hipótesis Generalizada del Continuo. En la sección 3, definimos para cada conjunto

[1]ZFC la Teoría Axiomática de conjuntos de Zermelo-Fraenkel y el Axioma de Elección. ZF es ZFC sin el Axioma de Elección

A, un modelo transitivo $L(A)$ de ZF. Tal modelo permite construir, con la ayuda del forcing de Cohen el Modelo de Feferman, un modelo transitivo de ZF que no satisface el Axioma de Elección y tal que A no pertenece al mismo. En la sección 4 describimos, para cada conjunto A, otro modelo transitivo $L(A)$ de ZF. Dicho modelo permite construir, con la ayuda del forcing de Cohen, al Modelo Básico de Cohen, un modelo transitivo de ZF que no satisface el Axioma de Elección y que contiene a A como elemento. También este $L(A)$ permite construir, con la ayuda del forcing de Levy (el Colapso de Levy) y la hipótesis "existe un cardinal inaccesible", al Modelo de Solovay, un modelo transitivo de ZF más el Principio de Elección Dependiente, donde vale que cualquier conjunto de reales es medible según Lebesgue (o Lebesgue medible), tiene la propiedad de Baire y cada subconjunto de reales no numerable contiene un subconjunto perfecto. En la sección 5 describimos, para cada conjunto A, al modelo $L[A]$ de ZFC. Finalizamos (sección 6) mencionando algunos problemas abiertos de la Teoría de Conjuntos relacionados con el tema.

2. La clase de los conjuntos constructibles de Gödel (L)

ZFC está constituida por los siguientes axiomas:

1. Axioma de Extensionalidad: Si X y Y son dos conjuntos que tienen los mismos elementos, entonces ellos son iguales.

2. Axioma de pares: Si X y Y son dos conjuntos, entonces existe un conjunto $Y = \{X, Y\}$, cuyos elementos son exactamente X y Y.

3. Axioma de comprensión: Si $P(X)$ es una propiedad bien definida, entonces para cualquier conjunto X existe un conjunto $Y = \{Z \in X : P(Z)\}$.

4. Axioma de la unión: Si X es un conjunto, entonces existe un conjunto $Y = \bigcup X$, la unión de todos los elementos de X.

5. Axioma del conjunto potencia: Para todo conjunto X, existe un conjunto $Y = P(X)$, el conjunto de los subconjuntos de X.

6. Axioma del infinito: Existe un conjunto infinito.

7. Axioma de reemplazo: Si F es una función definible, entonces para cualquier conjunto X existe un conjunto $Y = F(X) = \{F(x) : x \in X\}$.

8. Axioma de regularidad: Cualquier conjunto no vacío tiene un elemento $\in$-minimal.

9. Axioma de elección: Cualquier familia de conjuntos no vacíos tiene una función de elección.

ZF es la teoría axiomática de conjuntos constituída por los axiomas 1-8. ZFC se puede presentar como una teoría de primer orden, cuyo lenguaje tiene como único símbolo no lógico al relacional binario $\in$.

Dada una fórmula $\varphi(x)$ del lenguaje de ZFC decimos que la colección $\{x : \varphi(x)\}$ es una *clase*. Hay clases que son conjuntos, en particular todo conjunto es una clase pues si x es un conjunto, entonces la clase $\{z \in x : z = z\}$ es un conjunto por el axioma de comprensión, y $x = \{z \in x : z = z\}$ por el axioma de extensionalidad. Sin embargo, hay clases que no son conjuntos (*clases propias*), ya que el suponer que son conjuntos implica una contradicción. A continuación damos algunos ejemplos de ellas: Los ordinales, los cardinales, los conjuntos bien fundamentados, el universo y los constructibles de Gödel (clase definida por Gödel para probar la consistencia de ZF con el Axioma de Elección y la Hipótesis Generalizada del continuo[11] ,[12] , [13]).

Los ordinales (Ord) : Se dice que un conjunto x es *transitivo* si $\forall z(z \in x \rightarrow z \subseteq x)$. Un conjunto α es un *ordinal* si es transitivo y está estrictamente bien ordenado por $\in$, es decir, si es transitivo y el par $(\alpha, \in_\alpha)$ es un buen orden estricto, donde $\in_\alpha = \{(\gamma, \delta) \in \alpha \times \alpha : \gamma \in \delta\}$. $Ord := \{x : x \text{ es un ordinal}\}$. Ord está estrictamente bien ordenada por $\in$. Sean α y β dos ordinales. Se define $\alpha < \beta \leftrightarrow \alpha \in \beta$. La clase de los ordinales la podemos definir intuitivamente así: $0 := \emptyset$, $1 := \{0\}$,...,$n = \{0, \ldots, n-1\}$,..., $\omega := \{0, 1, 2, 3, \ldots\} = \mathbb{N}$, $\omega + 1 := \{0, 1, 2, 3, \ldots; \omega\}$, $\omega + 2 := \{0, 1, 2, 3, \ldots; \omega, \omega + 1\}$, $\omega + 3 := \{0, 1, 2, 3, \ldots; \omega, \omega + 1, \omega + 2\}$,..., $\omega + \omega := \{0, 1, 2, \ldots; \omega, \omega + 1, \omega + 2, \omega + 3, \ldots\}$,.... De esta forma se ve claramente que cada ordinal es igual al conjunto de los ordinales que lo preceden. Un ordinal α es *sucesor* si $\alpha = \beta + 1$, para algún ordinal β. Por ejemplo: 5, $\omega + 1$ y $\omega + 2$. Un ordinal es *límite* si no es cero ni sucesor. Por ejemplo, ω y $\omega + \omega$.

Los cardinales ($Card$): Dos conjuntos v y w son *equipotentes* si existe una función $f : v \longrightarrow w$ que sea biyectiva. Un conjunto κ es un cardinal si es un ordinal y no es equipotente a ningún ordinal menor (Es decir, si no es equipotente a ninguno de sus elementos). $Card := \{x : x \text{ es un cardinal}\}$. Cada cardinal tiene la forma $\aleph_\alpha$, para algún ordinal α, pues la clase $Card$ se puede definir por inducción tranfinita en los ordinales de la siguiente forma:

$$\begin{aligned}
\aleph_0 &:= \omega \\
\aleph_{\alpha+1} &:= (\aleph_\alpha)^+ := \{\beta \in Ord : \beta \text{ es equipotente a} \\
&\quad \text{algún subconjunto de } \aleph_\alpha\} \\
\aleph_\gamma &:= \bigcup_{\beta<\gamma} \aleph_\beta, \text{ si } \gamma \text{ es límite} \\
Card &:= \{\aleph_\alpha : \alpha \in Ord\}.
\end{aligned}$$

Un cardinal es *sucesor* si es de la forma $\aleph_{\alpha+1}$ para algún ordinal α. Por ejemplo $\aleph_1$, $\aleph_2$ y $\aleph_3$. Y es *límite* si es de la forma $\aleph_\gamma$, para algún ordinal límite γ. Por ejemplo $\aleph_\omega$ y $\aleph_{\omega+\omega}$.

Sean κ, η dos cardinales. $\kappa \leq \eta \leftrightarrow$ existe una función $f : \kappa \longrightarrow \eta$ que sea inyectiva. Sea x un conjunto. Se denotará por $|x|$ al único cardinal κ equipotente con x. Tal cardinal existe por el Axioma de Elección. x es *finito* si existe un $n \in \omega$ tal que $|x| = n$. x es *infinito* si no es finito. x es *numerable* si $|x| \leq \aleph_0$.

Cardinales regulares e inaccesibles: Sea α un ordinal límite. Decimos que $\beta < \alpha$ es *cofinal* con α si existe una función creciente $f : \beta \longrightarrow \alpha$ tal que para todo $\xi < \alpha$, existe un $\delta < \beta$ tal $f(\delta) \geq \xi$ (es decir, la imagen de f es no acotada en α). Dado α, la *cofinalidad* de α, $cof(\alpha)$, es el menor ordinal cofinal con α. Con respecto a la cofinalidad se cumple lo siguiente: $cof(\alpha)$ es el menor cardinal β tal que existe una partición de α en β pedazos cada uno de los cuales tiene cardinalidad estrictamente menor que α. Un cardinal infinito κ es *regular* si es igual a su cofinalidad. Decimos que es *singular* en caso contrario. Ejemplos: ω es regular y $\aleph_\omega$ es singular. Se puede demostrar que para cada ordinal α, el cardinal $\aleph_{\alpha+1}$ es regular. Un cardinal κ es un *límite fuerte*, si para todo cardinal $\alpha < \kappa$ se tiene que $2^\alpha < \kappa$. Un cardinal $\kappa > \omega$ es *fuertemente inaccesible* (o simplemente *inaccesible*) si es regular y límite fuerte. Decimos que κ es *débilmete inaccesible* si es regular y

límite. Es conocido que la existencia de cardinales inaccesibles no se puede demostrar en ZFC.

Los conjuntos bien fundamentados WF: Se define por inducción transfinita en los ordinales así:

$$\begin{aligned} R_0 &:= \emptyset \\ R_{\alpha+1} &:= P(R_\alpha) \\ R_\lambda &:= \bigcup_{\beta<\lambda} R_\beta,\ \lambda \text{ límite.} \\ WF &:= \bigcup_{\alpha\in Ord} R_\alpha. \end{aligned}$$

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

El universo (V): $V := \{x : x = x\}$. Por el Axioma de Fundamentación se tiene que $V = WF$.

Sea M una clase. M es transitiva si $\forall z(z \in M \to z \subseteq M)$. Ord y WF son clases transitivas. Pero $Card$ no es transitiva pues, por ejemplo, los $\aleph_\alpha$ para $\alpha > 0$ tienen como elementos ordinales que no son cardinales. Ahora definiremos a la clase de los conjuntos constructibles, pero antes daremos una definición previa que vamos a utilizar: Sea $\mathfrak{A} := \langle A, < f_i >_{i\in I}, < R_j >_{j\in J}, < a_k >_{k\in K}\rangle$ una estructura y $\mathcal{L}_{\mathfrak{A}}$ un lenguaje de primer orden para la misma. Decimos que un subconjunto $B \subseteq A$ es *definible* en $\mathfrak{A}$ si existe una fórmula $\varphi(x)$ del lenguaje $\mathcal{L}_{\mathfrak{A}}$ tal que $B = \{z \in A : \mathfrak{A} \models \varphi[z]\}$. Se dice que B es definible en $\mathfrak{A}$ con parámetros si existe fórmula $\varphi(x, x_1, \ldots, x_n)$ del lenguaje $\mathcal{L}_{\mathfrak{A}}$ y existen $a_1, \ldots, a_n \in A$ tal que $B = \{z \in A : \mathfrak{A} \models \varphi[z, a_1, \ldots, a_n]\}$.

L se define intuitivamente usando inducción transfinita sobre los ordinales de la siguiente manera [19] :

$$\begin{aligned} L_0 &:= \emptyset \\ L_{\alpha+1} &:= \{X \subseteq L_\alpha : X \text{ es definible en}(L_\alpha, \in, \langle b : b \in L_\alpha\rangle)\} \\ L_\lambda &:= \bigcup_{\beta\epsilon\lambda} L_\beta\ ;\ \lambda \text{ límite} \\ L &:= \bigcup_{\alpha\in Ord} L_\alpha \end{aligned}$$

Es claro que en el paso sucesor la expresión "X es definible en la estructura $\langle L_\alpha, \in, \langle b : b \in L_\alpha\rangle$" supone que se tiene un lenguaje de

primer orden con identidad, cuyos símbolos no lógicos son: Un símbolo relacional binario $\underline{\in}$ para la relación de pertenencia $\in$, y una constante $\underline{b}$ para cada $b \in L_\alpha$. Esta definición se puede formalizar en ZF, ver [19] y [16].

Teorema 2.1. *L es un modelo transitivo de ZF que contiene a los ordinales.*

Un demostración de este teorema puede encontrarse en [19] o [16].

Vale la pena resaltar que un importante resultado que es usado en esta prueba es el Teorema de Reflexión, el cual enunciaremos luego de la siguiente definición, y una prueba del mismo puede encontrarse en: [19] o [16].

Sean M, N dos clases tal que $M \subseteq N$ y $\varphi(x_1, \ldots, x_n)$ una fórmula. Se dice que $\varphi(x_1, \ldots, x_n)$ es absoluta para M y N si y sólo si,

$$\forall x_1 \ldots \forall x_n \in M[(M, \in) \models \varphi(x_1, \ldots, x_n) \leftrightarrow (N, \in) \models \varphi(x_1, \ldots, x_n)].$$

Teorema 2.2 (Teorema de Reflexión)**.** *Sea Z un clase, y para cada ordinal α, $Z(\alpha)$ un conjunto. Supongamos que:*

(1) $\alpha < \beta$, entonces $Z(\alpha) \subset Z(\beta)$.

(2) Si γ es un ordinal límite, entonces $Z(\gamma) = \bigcup_{\alpha<\gamma} Z(\alpha)$.

(3) $Z = \bigcup_{\alpha \in Ord} Z(\alpha)$.

Entonces: Para cualquier fórmulas $\varphi_1, \ldots, \varphi_n$,

$$\forall \alpha \exists \beta > \alpha (\varphi_1, \ldots, \varphi_n \text{ son absolutas para } Z(\beta), Z).$$

Teorema 2.3. *L es un modelo del Axioma de Elección.*

Una demostración de este resultado puede encontrarse en [19] o [16]. La idea es definir un buen orden para L utilizando la definición formal del mismo.

La Hipótesis Generalizada del Continuo, HGC, es la siguiente proposición:

$$\forall \alpha \in Ord(2^{\aleph_\alpha} = \aleph_{\alpha+1}).$$

La Hipótesis del Continuo (HC) es la HGC para el caso $\alpha = 0$, es decir, $2^{\aleph_0} = \aleph_1$.

Teorema 2.4. *L es un modelo de la HGC.*

Una prueba de este Teorema puede encontrarse en [19] o [16]. Tal prueba usa el *Axioma de Constructibilidad* el cual afirma que todo conjunto es constructible, es decir, $\forall x \exists \alpha (x \in L_\alpha)$. El Axioma de Constructibilidad se simboliza así: $V = L$. $L \models V = L$. Asumiendo $V = L$, el caso base de la HGC, $2^{\aleph_0} = \aleph_1$, se sigue del hecho de que cualquier subconjunto constructible de ω es construido en algún paso numerable, es decir, $P(\omega) \subseteq L(\omega_1)$. Dado que $\mid L(\omega_1) \mid = \omega_1$ se tiene que $2^{\aleph_0} \leq \aleph_1$. Y como por el Teorema de Cantor se tiene que $\aleph_1 \leq 2^{\aleph_0}$, se concluye que $2^{\aleph_0} = \aleph_1$. Un importante Teorema que se usa para la demostración de que $L \models HGC$ es el Teorema de Colapsamiento de Mostowski, el cual enunciamos a continuación luego de la siguiente definición:

Sean R una relación binaria sobre una clase A. R es *bien fundamentada* sobre A si y sólo si para todo conjunto $X \subseteq A : X \neq \emptyset \rightarrow \exists y \in X(\forall z \in X((z, y) \notin R))$. El y anterior se llama R-mínimal en X. R es *semejante a un conjunto* sobre A si y sólo si $\forall x \in A$, $\{y \in A : (y, x) \in R\}$ es un conjunto. R es *extensional* sobre A si y sólo si $\forall x, y \in A(\forall z \in A(zRx \leftrightarrow zRy) \rightarrow x = y)$.

Teorema 2.5 (Teorema del Colapsamiento de Mostowski)**.** *Sea A una clase y R una relación binaria sobre A tal que R es bien fundamentada, semejante a un conjunto y extensional. Entonces existe una clase transitiva M y una función biyectiva $G : A \longrightarrow M$ tal que G es un isomorfismo entre (A, R) y $(M, \in)$, es decir, $\forall x, y \in A[xRy \leftrightarrow G(x) \in G(y)]$. M y G son únicas.*

Una prueba de este Teorema puede encontrarse en [19] o [16].

3. Primera relativización del concepto de conjunto constructible de Gödel (L(A))

Sea A un conjunto. Se define intuitivamente a la clase $L(A)$ por inducción en los ordinales así:

Definicion 3.1.

$$\begin{aligned} L_0(A) &:= \emptyset \\ L_{\alpha+1}(A) &:= \{X \subseteq L_\alpha(A) : X \textit{ es definible en la estructura} \\ &\quad \langle L_\alpha(A), \in, \langle a \cap L_\alpha(A) : a \in A\rangle, \langle d : d \in L_\alpha(A)\rangle\rangle\} \\ L_\lambda(A) &:= \bigcup_{\beta \in \lambda} L_\beta(A)\ , \quad \lambda \textit{ límite} \end{aligned}$$

$$L(A) := \bigcup_{\alpha \in Ord} L_\alpha(A)$$

Es claro que en el paso sucesor la expresión "X es definible en la estructura $\langle L_\alpha(A), \in, \langle a \cap L_\alpha(A) : a \in A\rangle, \langle d : d \in L_\alpha(A)\rangle\rangle$" supone que se tiene un lenguaje de primer orden con identidad, cuyos símbolos no lógicos son: Una constante $\underline{d}$ para cada $d \in L_\alpha(A)$, un símbolo de relación unario P_a por cada conjunto $a \cap L_\alpha(A)$, donde $a \in A$; y un símbolo relacional binario $\underline{\in}$ para la relación de pertenencia $\in$. Una formalización de la definición de L(A) en ZF puede hacerse siguiendo las ideas expuestas por [19] para formalizar el concepto de L.

Teorema 3.2. *L(A) es un modelo transitivo de ZF que contiene los ordinales.*

Una prueba de este resultado se puede hacer de manera similar a la que realiza [19] para demostrar que L es un modelo transitivo de ZF que contiene a los ordinales.

Teorema 3.3. *Si A es transitivo, entonces* $A \subseteq L(A)$.

Una prueba de este resultado puede hacerse por inducción en el rango (en V) de A ($\rho(A)$). Esto implica que si A es un conjunto transitivo, entonces L(A) es el menor modelo transitivo de ZF que contiene a los ordinales y a A.

3.1. Con L(A) y el forcing de Cohen que agrega infinitos reales genéricos se pueden construir modelos tipo Feferman

Denotamos por $\mathbb{N}^{<\omega}$ al conjunto de todas las sucesiones finitas de núneros naturales, es decir, $\mathbb{N}^{<\omega} := \bigcup_{n\in\mathbb{N}} \mathbb{N}^n$, donde $\mathbb{N}^n := \{s : s : n \longrightarrow \mathbb{N}\}$. Denotamos por $\mathbb{N}^\infty$ al conjunto de todas las sucesiones de números naturales, es decir, $\mathbb{N}^\infty := \{f : f : \mathbb{N} \longrightarrow \mathbb{N}\}$. Consideremos el espacio topológico $(\mathbb{N}^\infty, t_1)$, donde t_1 es la topología generada por los conjuntos básicos de la forma $U_s := \{f \in \mathbb{N}^\infty : s \subseteq f\}$, donde $s \in \mathbb{N}^{<\omega}$. Es decir, t_1 es la topología producto de $\mathbb{N}^\infty$ que se obtine al dotar a $\mathbb{N}$ con la topología discreta. El espacio $(\mathbb{N}^\infty, t_1)$ se denomina *Espacio de Baire*. Denotamos por $2^\mathbb{N}$ al conjunto de todas las sucesiones de ceros y unos, es decir, $2^\mathbb{N} := \{f : f : \mathbb{N} \longrightarrow 2\}$. $2^\mathbb{N} \subseteq \mathbb{N}^\infty$ y con la topología producto heredada de $\mathbb{N}^\infty$ se denomina

Espacio de Cantor. Denotamos por $\mathbb{N}^{[\infty]}$ a la familia de todos los subconjuntos infinitos de $\mathbb{N}$. Consideremos el espacio topológico $(\mathbb{N}^{[\infty]}, t_2)$, donde t_2 es la topología generada por los conjuntos básicos de la forma $U_a := \{X \in \mathbb{N}^{[\infty]} : a \sqsubset X\}$, donde a es un subconjunto finito de $\mathbb{N}$ y $\sqsubset$ es la relación de segmento inicial. Es conocido que los espacios $\mathbb{N}^{\infty}$ y $\mathbb{N}^{[\infty]}$ son homeomorfos [6]. También que ellos son homeomorfos a los irracionales, considerados como un subespacio del conjunto de los números reales $\mathbb{R}$ [16]. Todos los espacios mencionados, $\mathbb{N}^{\infty}$, $2^{\mathbb{N}}$,$\mathbb{N}^{[\infty]}$ y $\mathbb{R}$ tienen la misma cardinalidad$(2^{\aleph_0})$ y son *espacios polacos*, es decir, espacios *métricos*, *separables* y *completos* [16]. A los elementos de los espacios $\mathbb{N}^{\infty}$, $2^{\mathbb{N}}$ y $\mathbb{N}^{[\infty]}$ nosotros también los llamamos *números reales*.

Presentación intuitiva del Método de forcing, una técnica para construir modelos: Sea $(P, \leq)$ un orden parcial. $D \subseteq P$ es *denso* en P si y sólo si $\forall p \in P \exists q \in D(q \leq p)$. $G \subseteq P$ es un *filtro* sobre P si y sólo si: (1) $\forall p, q \in G\, \exists r \in G(r \leq p \wedge r \leq q)$ y (2) $\forall p \in G\, \forall q \in P(p \leq q \rightarrow q \in G)$. Sean M un modelo transitivo de ZFC y $P \in M$ un orden parcial. G es un filtro *P-genérico* sobre M si y sólo si G es un filtro sobre P y para todo denso $D \subseteq P$: Si $D \in M \rightarrow G \cap D \neq \emptyset$. Sean M un modelo transitivo de ZFC, $P \in M$ un orden parcial y G un filtro P-genérico sobre M. Entonces $M[G]$, la *extensión genérica* de M, es el menor modelo transitivo de ZFC que es cerrado bajo las operaciones usuales de la Teoría de Conjuntos y contiene a $M \cup \{G\}$. (en este caso M se llama *modelo base*). Cuando el orden parcial P satisface que $\forall p \in P\, \exists\, q, r \in P(q \leq p \wedge r \leq p \wedge q \perp r)$, entonces $G \notin M$. Donde $q \perp r$ significa que q y r son incompatibles, es decir, $\neg \exists s \in P(s \leq q \wedge s \leq r)$. Es decir, cuando P tiene la condición antes descrita el nuevo modelo $M[G]$ es una extensión propia del modelo base M, por ejemplo, $M[G]$ puede tener números reales que no estan en M y en este caso se dice que el orden parcial P (o el *forcing* P) agrega nuevos reales a M. Un ejemplo de un orden parcial que agrega nuevos números reales al modelo base es el forcing de Cohen, el cual definiremos a continuación.

Forcing de Cohen que agrega un real al modelo base: El forcing de Cohen es el orden parcial $(\mathcal{C}, \leq)$ donde $\mathcal{C}$ es el conjunto de todas las secuencias finitas de ceros y unos. Y $p \leq q \leftrightarrow p \supseteq q$. Sea G un $\mathcal{C}$-genérico sobre un modelo transitivo M de ZFC. Entonces, $f = \cup G \in M[G]$ y $f : \omega \longrightarrow 2$ se llama *real de Cohen*. Sea $a := \{n : f(n) = 1\}$. El conjunto a también es llamado un real de Cohen, y se cumple que

$M[G] = M[f] = M[a]$, porque G puede ser recuperado a partir de f, $G = \{p \in \mathcal{C} : p \subset f\}$. $\mathcal{C}$ también se puede definir como el conjunto de todas las secuencias finitas de números naturales. Con el mismo orden anterior, es decir, $p \leq q \leftrightarrow p \supseteq q$. En este caso si H es un $\mathcal{C}$-genérico sobre un modelo transitivo M de ZFC, entonces la función $g : \omega \longrightarrow \omega$ tal que $g = \cup H$ es también llamada un real de Cohen. Y se cumple que $M[G] = M[f] = M[a] = M[H] = M[g]$. El forcing de Cohen $\mathcal{C}$ se puede generalizar para agregar una cantidad infinita ($\alpha \geq \omega$) de reales genéricos. Una manera de hacerlo es la siguiente: Sea M un modelo transitivo de ZFC. Sea $(\mathcal{C}_\alpha, \leq)$ el siguiente orden parcial:

$$\mathcal{C}_\alpha := \{p :| \, p \, |< \omega \wedge p \subseteq \alpha \times \omega \times 2\}.$$

$$p \leq q \leftrightarrow q \subseteq p$$

Sea G un $\mathcal{C}_\alpha$-genérico sobre M y $M[G]$ la extensión genérica resultante. Tal forcing agrega la función $F := \bigcup G = \alpha \times \omega \longrightarrow 2$. La cual permite definir α nuevos reales genéricos de la siguiente forma: Para cada $\beta < \alpha$ definimos la función $f_\alpha : \omega \longrightarrow 2$ así: $f_\alpha(n) := F(\beta, n)$, para cada $n \in \omega$. Los α nuevos reales genéricos también pueden ser definidos de la siguiente manera: Sea $\beta \in \alpha$. Se define,

$$a_\beta := \{m \in \omega : \exists p \in G(p(\beta, m) = 1)\}.$$

Los a_β asi definidos son subconjuntos de ω y estan en $M[G]$. Obviamente los f_β son las funciones características de los a_β. A la extensión $M[G]$ la denotamos a veces por $M[\{a_\beta : \beta < \alpha\}]$ para resaltar los reales genéricos que agregamos.

Sea L la clase de los conjuntos constructibles. Y sea $L[G]$ la extensión genérica de L que se obtiene utilizando el forcing de Cohen que agrega $\aleph_0$ nuevos reales genéricos ($\mathcal{C}_{\aleph_0}$). Sea A el conjunto de dichos $\aleph_0$ reales genéricos, entonces $L(A)$ en $L[G]$ es el *Modelo de Feferman*. El Modelo de Feferman es un modelo transitivo de ZF que contiene a los ordinales y tiene las siguientes dos propiedades: $A \notin L(A)$ y $L(A)$ no satisface el Axioma de Elección [9] , [17]. Estos resultados se enuncian en forma de teorema a continuación:

Teorema 3.1.1. *$A \notin L(A)$.*

Teorema 3.1.2. *$L(A)$ no satisface el Axioma de Elección.*

En [15] puede encontrarse una lista de versiones débiles del Axioma de Elección válidas en el Modelo de Feferman. También se puede encontrar en [15] una lista de formas débiles que no son válidas en el mismo. Dentro de las no válidas se encuentra por ejemplo el Teorema del Ideal Primo: Toda álgebra booleana tiene un ideal primo (o todo filtro se puede extender a un ultrafiltro) [9]. Es conocido que el Teorema del Ideal Primo es equivalente al Teorema de Compacidad (si cada subconjunto finito de Σ tiene un modelo, entonces Σ tiene un modelo) y al Teorema de Completitud (si Σ es un conjunto consistente de sentencias, entonces Σ tiene un modelo) de la Lógica de Primer Orden [17].

4. Segunda manera de relativizar el concepto de conjunto constructible de Gödel (L(A))

Sea A un conjunto. Se define intuitivamente otra clase $L(A)$ por inducción en los ordinales así:

Definicion 4.1.

$$\begin{aligned} L_0(A) &:= CT(\{A\}) \\ L_{\alpha+1}(A) &:= \{X \subseteq L_\alpha(A) : X \textit{ es definible en la estructura } \langle L_\alpha(A), \in, \langle d : d \in L_\alpha(A)\rangle\rangle\} \\ L_\lambda(A) &:= \bigcup_{\beta\in\lambda} L_\beta(A)\ , \quad \lambda \textit{ límite} \\ L(A) &:= \bigcup_{\alpha\in Ord} L_\alpha(A) \end{aligned}$$

Es claro que en el paso sucesor la expresión "X es definible en la estructura $\langle L_\alpha(A), \in, \langle d : d \in L_\alpha(A)\rangle\rangle$" supone que se tiene un lenguaje de primer orden con identidad, cuyos símbolos no lógicos son: Una constante $\underline{d}$ para cada $d \in L_\alpha(A)$ y un símbolo relacional binario $\underline{\in}$ para la relación de pertenencia $\in$. Una formalización de la definición de L(A) en ZF puede darse siguiendo las ideas expuestas en [19] para el caso de L.

Teorema 4.2. *$L(A)$ es un modelo transitivo de ZF que contiene los ordinales.*

Como en el caso anterior, la prueba es similar a la que se realiza en [19] para demostrar que L es un modelo transitivo de ZF que contiene a los ordinales.

$A \subseteq L(A)$. Esto ocurre por construcción e implica que $L(A)$ es el menor modelo transitivo de ZF que contiene a los ordinales y a A.

4.2. Con L(A) y el forcing de Cohen que agrega $\aleph_0$ reales genéricos se puede construir el Modelo Básico de Cohen

Sea L la clase de los conjuntos constructibles. Y sea $L[G]$ la extensión genérica de L que se obtiene utilizando el forcing de Cohen que agrega $\aleph_0$ nuevos reales genéricos ($\mathcal{C}_{\aleph_0}$). Sea A el conjunto de dichos $\aleph_0$ reales genéricos, entonces $L(A)$ en $L[G]$ es el *Modelo Básico de Cohen*. Este es un modelo transitivo de ZF que contiene a los ordinales y tiene las siguientes dos propiedades:$A \in L(A)$ y $L(A)$ no satisface el Axioma de Elección. Una prueba de que $L(A)$ no satisface el Axioma de Elección usa argumentos de simetría y puede realizarse procediendo análogamente a como se hace en [16] , páginas 221-223. Entre las versiones débiles del AE que satisface el Modelo Básico de Cohen se encuentra el Teorema del Ideal Primo[10]. Otra versión débil del AE válida en el Modelo Básico de Cohen es: Cualquier familia de conjuntos bien ordenables no vacíos, tiene una función de elección[17].

Otra forma débil del AE es el *Teorema de Ramsey*, el cual afirma que cualquier conjunto infinito X tiene la siguiente propiedad: Para cualquier partición del conjunto $[X]^2 := \{\{a, b\} : a, b \in X\}$ en dos pedazos, existe un subconjunto infinito Y de X tal que $[Y]^2$ está incluido en una de las piezas de la partición. El Teorema de Ramsey es falso en el modelo Básico de Cohen [3]. Una lista de versiones débiles del Axioma de elección válidas en el Modelo Básico de Cohen puede encontrarse en [15]. También se puede encontrar en [15] otra lista de versiones débiles del AE que no son válidas en tal modelo.

4.3. Con L(A) y el Colapso de Lévy más la hipótesis de que existen cardinales inaccesibles se puede construir el Modelo de Solovay

El siguiente teorema nos ofrece una técnica para colapsar un cardinal aplicando forcing:

Teorema 4.3.1. *Sea κ un cardinal regular y $\lambda > \kappa$ otro cardinal. Entonces existe un orden parcial $(P, <)$ que colapsa a λ onto κ, es decir, $(P, <)$ agrega una fución sobreyectiva de κ en λ, por lo que λ tiene cardinalidad κ en la extensión genérica. Más todavía:*

(a) Cualquier cardinal $\alpha \leq \kappa$ en el modelo base M sigue siendo un cardinal en M[G]; y

(b) Si $\lambda^{<\kappa} = \lambda$, entonces cualquier $\alpha > \lambda$ sigue siendo un cardinal en M[G]. Donde $\lambda^{<\kappa} := \bigcup\{\lambda^{\alpha} : \alpha < \lambda\}$.

Una demostración de este teorema puede encontrarse en [16]. El orden parcial $(P, <)$ utilizado es el siguiente:

P es el conjunto de todas las funciones p tal que:

(i) $dom(p) \subseteq \kappa$ y $\mid dom(p) \mid < \kappa$,

(ii) $ran(p) \subseteq \lambda$.

Y el orden $<$ es: $p < q$ si y sólo si $p \supseteq q$.

La técnica anterior nos dice como colapsar un cardinal λ en un cardinal regular menor κ. Lévy nos presenta una técnica de colapsamiento de los cardinales menores que un cardinal inaccesible λ preservando a λ, así λ es un cardinal sucesor en la extensión genérica. Esto es el contenido del siguiente teorema:

Teorema 4.3.2. *Sea κ un cardinal regular y $\lambda > \kappa$ un cardinal inaccesible. Entonces existe un orden parcial $(P, <)$ tal que:*

(a) Cualquier α, $\kappa \leq \alpha < \lambda$, tiene cardinal κ en M[G], y

(b) Cualquier cardinal $\leq \kappa$ y cualquier cardinal $\geq \lambda$ siguen siendo cardinales en M[G].

En particular $M[G] \models \lambda = \kappa^{+}$.

Una demostración de este teorema puede encontrarse en [16]. El orden parcial $(P, <)$(se denota $Col(\kappa, < \lambda)$) utilizado es el siguiente:

P son todas las funciones p cuyo dominio esta contenido en $\lambda \times \kappa$ tal que:

(i) $\mid dom(p) \mid < \kappa$,

(ii) $p(\alpha, \xi) < \alpha$ para cada $(\alpha, \xi) \in dom(p)$.

Conjuntos proyectivos y borelianos [2] : Una fórmula del lenguaje de la teoría de conjuntos es $\sum_n^1$ si es de la forma:

$$\exists x_1 \forall x_2 \exists x_3 \ldots \psi(x_1, \ldots, x_n, y_1, \ldots, y_m),$$

donde los cuantificadores $\exists x_1 \forall x_2 \exists x_3 \ldots$ tienen rango sobre $\mathbb{N}^{\infty}$(es decir: $\exists x_1 \in \mathbb{N}^{\infty} \forall x_2 \in \mathbb{N}^{\infty} \exists x_3 \in \mathbb{N}^{\infty} \ldots$), $y_1, \ldots, y_m$ son variables libres que toman valores en $\mathbb{N}^{\infty}$, y todos los cuantificadores de ψ tienen rango sobre ω.

Una fórmula es $\prod_n^1$ si es la negación de una fórmula $\sum_n^1$.

Una *fórmula es proyectiva* si es $\sum_n^1$ o $\prod_n^1$, para algún número natural n.

Decimos que un conjunto $A \subseteq \mathbb{N}^\infty$ es *proyectivo* si y sólo si existe una fórmula $\sum_n^1$ o $\prod_n^1$, $\varphi(x, y_1, \ldots, y_m)$, y $a_1, \ldots, a_m \in \mathbb{N}^\infty$, tal que $A = \{b \in \mathbb{N}^\infty : \varphi(x, a_1, \ldots, a_m)\}$. Un conjunto proyectivo $B \subseteq \mathbb{N}^\infty$ es *boreliano* si es definible con una fórmula $\sum_1^1$ y también es definible con una fórmula $\prod_1^1$.

Sea κ un cardinal inaccesible en L. Y sea $L[G]$ la extensión genérica de L que se obtiene usando el Colapso de Lévy: $Col(\aleph_0, < \kappa)$. En $L[G]$ todo conjunto proyectivo de reales es medible Lebesgue, tiene la propiedad de Baire, y si es no numerable, entonces contiene un subconjunto perfecto[22]. Sea $L(\mathbb{R})$ en $L[G](L(\mathbb{R})$ es llamado el Modelo de Solovay), donde $\mathbb{R}$ son los reales en $L[G]$. Entonces $L(\mathbb{R})$ satisface ZF, más el Principio de Elección Dependiente(DC), más "cualquier conjunto de reales es medible Lebesgue, tiene la propiedad de Baire y cada subconjunto de reales no numerable contiene un subconjunto perfecto"[22]. El Principio de Elección Dependiente es la versión débil del AE que afirma lo siguiente: Si E es una relación binaria sobre un conjunto no vacío A, y si para cualquier $a \in A$ existe un $b \in A$ tal que bEa, entonces existe una secuencia $a_0, a_1, a_2, \ldots, a_n, \ldots$ en A tal que: $a_{n+1}Ea_n$, para todo $n \in \mathbb{N}$. DC implica El Axioma de Elección Numerable (Cualquier familia numerable de conjuntos no vacíos tiene una función de elección). En [15] se puede encontrar una lista de formas débiles del Axioma de Elección que valen en el modelo de Solovay. También se puede encontrar otra lista de formas débiles que no valen en el mismo.

5. Tercera manera de relativizar el concepto de conjunto constructible de Gödel (L[A])

Definición intuitiva de L[A]:

Sea A un conjunto. Se define intuitivamente a la clase $L[A]$ por inducción sobre los ordinales así:

Definicion 5.1.

$$\begin{aligned} L_0[A] &:= \emptyset \\ L_{\alpha+1}[A] &:= \{X \subseteq L_\alpha[A] : X \textit{ es definible en la estructura} \\ &\quad \langle L_\alpha[A], \in, A \cap L_\alpha[A], \langle d : d \in L_\alpha[A]\rangle\rangle\} \end{aligned}$$

$$L_\lambda[A] := \bigcup_{\beta\in\lambda} L_\beta[A]\ , \quad \lambda \textit{ límite}$$
$$L[A] := \bigcup_{\alpha\in Ord} L_\alpha[A]$$

Es claro que en el paso sucesor la expresión "X es definible en la estructura $\langle L_\alpha[A], \in, A\cap L_\alpha[A], \langle d : d \in L_\alpha[A]\rangle\rangle$" supone que se tiene un lenguaje de primer orden con identidad, cuyos símbolos no lógicos son: Una constante $\underline{d}$ para cada $d \in L_\alpha[A]$, un símbolo de relación unario P para el conjunto $A \cap L_\alpha[A]$, y un símbolo relacional binario $\underline{\in}$ para la relación de pertenencia $\in$. Una formalización de $L[A]$ en ZF puede hacerse siguiendo las ideas expuestas en [19] para L.

Teorema 5.2. *L(A) es un modelo transitivo de ZF que contiene los ordinales.*

Como en los casos anteriores la prueba es similar a la que se realiza para demostrar que L es un modelo transitivo de ZF que contiene a los ordinales.

Teorema 5.3. *L[A] satisface el Axioma de Elección.*

Una prueba puede realizarse de manera análoga a como se hace en [19] o [16] para L.

Otras propiedades de L[A] se expresan a partir del siguiente teorema, una prueba del mismo puede encontrarse en [16] :

Teorema 5.4. *(i) L[A] satisface el Axioma* $\exists X(V = L[X])$.

(ii) Si M es un modelo transitivo de ZF que contiene a los ordinales y tal que $A \cap M \in M$, *entonces* $L[A] \subseteq M$.

(iii) Existe un ordinal α_0 *tal que para todo* $\alpha \geq \alpha_0$,

$$L[A] \models 2^{\aleph_\alpha} = \aleph_{\alpha+1}.$$

6. Algunos problemas abiertos de la teoría de conjuntos que tienen relación con este tema

1. *Definición de la Propiedad de Ramsey*: Para toda (partición) $F : \mathbb{N}^{[\infty]} \longrightarrow 2$ existe un conjunto infinito $H \subseteq \mathbb{N}$ tal que F es constante en $H^{[\infty]}$. La Propiedad de Ramsey se denota así: $\omega \rightarrow (\omega)^\omega$. Es conocido que esta propiedad es falsa si vale el

Axioma de Elección [5] y que Mathias probó que en el Modelo de Solovay vale la misma [20], de esta forma demostró que si ZFC + "existe un cardinal inaccesible" es consistente, también es consistente ZF + DC + $\omega \to (\omega)^{\omega}$.

Problema [18]:

¿ Se puede eliminar la hipótesis de la existencia del cardinal inaccesible para probar la consistencia de la propiedad de Ramsey con ZF?. Es decir, ¿ se puede construir un modelo para ZF + $\omega \to (\omega)^{\omega}$ utilizando solamente ZFC ?.

2. (Carlos Di Prisco) ¿ $\omega \to (\omega)^{\omega}$ es compatible con $\aleph_1 \leq 2^{\aleph_0}$?, donde $\aleph_1 \leq 2^{\aleph_0}$ es una versión débil del AE que significa: Existe una función inyectiva de $\aleph_1$ en el conjunto de los números reales $\mathbb{R}$.

Para fianlizar, podemos mencionar que es conocido por los trabajos de K. Gödel [11] y P. Cohen [4] que la Hipótesis del Continuo (HC) es independiente de ZFC. Pero desde una perspectiva platonista de la matemática, la HC es una proposición verdadera o falsa, entonces sigue abierto el problema sobre cómo aumentar el poder deductivo de ZFC agregando nuevos axiomas para determinar cuál es el cardinal del continuo [8], [14].

Referencias

[1] Bagaria, J., *Models of Set Theory*, Master's Course, Universidad de Barcelona, España. 2008.

[2] Bagaria, J., *The many faces of the Continuum*, Notas para un curso dictado en el XI Simposio Latinoamericano de Lógica Matemática, Universidad de los Andes, Venezuela, 1998.

[3] Blass, A., *Ramsey's theorem in the hierarchy of choice principles.* The Journal of Symbolic Logic. **42**, No. 3, 1977, 387-390.

[4] Cohen, P., *Set Theory and the Continuum Hypothesis*, W.A. Benjamin, INC. New York, 1966.

[5] Di Prisco, C., *Teoría de Conjuntos*, Universidad Central de Venezuela, Consejo de Desarrollo Científico y Humanístico, Caracas, 2009.

[6] Di Prisco, C., *Temas Avanzados de Teoría de Conjuntos*, Notas para un curso dictado en el Postgrado de Matemáticas de la Facultad de Ciencias de la Universidad Central de Venezuela, Caracas, 2001.

[7] Di Prisco, C., *Mathematics versus metamathematics in Ramsey theory of the real numbers*, Logic, Methology and Phylosophy of Science, Proccedings of the twelfth International Congress, (Petr Hajek, Luis Valdes-Villanueva, Dag Westerstalhl. Eds), Kings College Publications, London, 2005.

[8] Di Prisco, C., *Are we closer to a solution of the antinuum problem?*, Manuscrito - Rev. Int. Fil., **28**, No.2, Campinas, 2005, 331-350.

[9] Feferman, S., *Some applications of notion of forcing and generic sets* , Fundamenta Mathematicae, **56**, 1964/65, 325-345.

[10] Halpern, J, Lévy, A., *The boolean prime theorem does not imply the axiom of choice*, Axiomatic Set Theory, Proceedings of Symposia in Pure Mathematics, **13**, part 1. (D. Scott ed.), Univ. of California, Los Angeles, 1967, 83-134.

[11] Gödel, K., *La consistencia del axioma de elección y la hipótesis generalizada del continuo*, Obras Completas, Gödel, K., Alianza, Madrid, 1981.

[12] Gödel, K., *Prueba de la consistencia de la hipótesis generalizada del continuo* , Obras Completas, Gödel, K., Alianza, Madrid, 1981.

[13] Gödel, K., *La consistencia del axioma de elección y de la hipótesis generalizada del continuo con los axiomas de la teoría de conjuntos*, Obras Completas, Gödel, K., Alianza, Madrid, 1981.

[14] Gödel, K., *¿ Qué es el Problema del continuo de Cantor?*, Obras Completas, Gödel, K., Alianza, Madrid, 1981.

[15] Howard, P., Rubin, J., *Consecuences of the Axiom of Choice*, American Mathematical Society, EEUU, 1998.

[16] Jech, T., *Set Theory*, Springer, New York, 2000.

[17] Jech, T., *The Axiom Choice*, North Holland, Amsterdam, 1973.

[18] Kanamori, A., *The Higler Infinite*, Springer, New York, 1997.

[19] Kunen, K., *Set Theory. An Introduction to Independence Proofs*. North Holland, Amsterdam, 1980.

[20] Mathias, A., *Happy families*, Annals of Mathematical Logic, **12**, No. 1, 1977, p. 59-111.

[21] Ramsey, P., *On a problem of formal logic*, Proceedings of the London Mathematical Society, **30**, 1929/30, 264-286.

[22] Solovay, R., *A model of set theory where every set of reals is Lebesgue measurable* , Annals of Mathematics, **92**, 1970, 1-56.

Dirección de los autores
Franklin C. Galindo
Universidad Central de Venezuela,
Facultad de Humanidades y Educación,
Escuela de Filosofía.
Venezuela.
e-mail: franklin.galindo@ucv.ve

Carlos A. Di Prisco
Instituto Venezolano de Investigaciones Científicas
Venezuela.
e-mail: cdiprisc@ivic.ve

Mixba'al Rev. Met. de Mat.
№ 1, Vol. I, Junio 2010. 41 – 64

El grupo Ext

Héctor Gabriel Salazar Pedroza

Resumen

Sean A y C dos R-módulos. Se definirá una operación binaria $+$ en $\mathrm{Ext}(C, A)$, el conjunto de clases de equivalencia de extensiones de A por C, de tal modo que $(\mathrm{Ext}(C, A), +)$ sea un grupo abeliano y cuyo elemento identidad sea la clase de equivalencia de la extensión $B = A \oplus C$.

Palabras clave: Funtor Ext, extensiones of módulos, clases de equivalencia de extensiones.

Clasificación de la AMS: 13D07, 18G15, 20J05, 20K40

1. Introducción

Sea R un anillo conmutativo con elemento unitario. Considere la sucesión exacta

$$0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$$

de R-homomorfismos. Ésta da lugar a una sucesión exacta de homomorfismos de grupos abelianos

$$0 \to \mathrm{Hom}_R(C, D) \xrightarrow{\beta'} \mathrm{Hom}_R(B, D) \xrightarrow{\alpha'} \mathrm{Hom}_R(A, D)$$

para cualquier R-módulo D, donde el homomorfismo α' no es suprayectivo en general. En forma equivalente, los homomorfismos de A en D no se pueden "levantar", en general, a homomorfismos de B en D. Una técnica del álgebra homológica permite extender estas sucesiones exactas en forma natural: mediante el funtor Ext se puede lograr una sucesión exacta infinita

$$\cdots \to \mathrm{Ext}^n_R(C, D) \to \mathrm{Ext}^n_R(B, D) \to \mathrm{Ext}^n_R(A, D) \to \mathrm{Ext}^{n+1}_R(C, D) \to \cdots$$

que comienza con

$$0 \to \mathrm{Ext}^0_R(C, D) \to \mathrm{Ext}^0_R(B, D) \to \mathrm{Ext}^0_R(A, D) \to \mathrm{Ext}^1_R(C, D) \to \cdots$$

y donde $\mathrm{Ext}^0_R(\_, D) = \mathrm{Hom}_R(\_, D)$. Los grupos $\mathrm{Ext}^1_R(\_, D)$ proporcionan una estimación del conjunto de homomorfismos de A en D que no se pueden extender a B. En esta nota construiremos $\mathrm{Ext}^1_R(C, A)$ con todo cuidado y comprobaremos que efectivamente es un grupo abeliano. Escribiremos simplemente $\mathrm{Ext}(C, A)$ en lugar de $\mathrm{Ext}^1_R(C, A)$. Dados dos módulos A y B, escribiremos $A \leq B$ para denotar que A es un submódulo de B.

2. Preliminares

Definición 2.1. *Una **extensión de** A **por** C es una sucesión exacta corta $E : 0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ de homomorfismos.*

Definición 2.2. *Dada una extensión $E : 0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$, decimos que α (respectivamente β) **se escinde** si existe un homomorfismo $\sigma : B \to A$ (respectivamente $\rho : C \to B$) tal que $\sigma \circ \alpha = 1_A$ (respectivamente $\beta \circ \rho = 1_C$). Decimos que σ (respectivamente ρ) es una **escisión** de α (respectivamente β). Si α o β se escinde, decimos que E es una extensión **escindible** (o que **se escinde**).*

Proposición 2.3. *Dada la extensión $E : 0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$, las siguientes afirmaciones son equivalentes:*

(a) *α se escinde;*

(b) *β se escinde;*

(c) *existe un isomorfismo $B \cong A \oplus C$ tal que el diagrama*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \xrightarrow{\alpha} & B & \xrightarrow{\beta} & C & \longrightarrow & 0 \\ & & \| & & \cong\downarrow & & \| & & \\ 0 & \longrightarrow & A & \xrightarrow{\iota} & A \oplus C & \xrightarrow{\pi} & C & \longrightarrow & 0 \end{array}$$

es conmutativo, donde ι es la inyección y π es la proyección.

Demostración. (a) $\Rightarrow$ (c). Supongamos que existe $\sigma : B \to A$ tal que $\sigma \circ \alpha = 1_A$. Consideremos el homomorfismo $\theta : B \to A \oplus C$ dado por $\theta(b) = (\sigma(b), \beta(b))$. Sea $b \in B$ tal que $\theta(b) = (0, 0)$, es decir, $\sigma(b) = 0$ y $\beta(b) = 0$. Se sigue que existe $a \in A$ tal que $\alpha(a) = b$, por lo que $a = \sigma(\alpha(a)) = \sigma(b) = 0$, es decir, $b = \alpha(0) = 0$ y por lo tanto, θ es un

monomorfismo. Sea $(a, c) \in A \oplus C$. Por ser β un epimorfismo, existe $b \in B$ tal que $\beta(b) = c$. Entonces

$$\begin{aligned}\theta(b - \alpha(\sigma(b) - a)) &= (\sigma(b - \alpha(\sigma(b) - a)), \beta(b - \alpha(\sigma(b) - a))) \\ &= (\sigma(b - \alpha(\sigma(b)) + \alpha(a)), \beta(b) - \beta(\alpha(\sigma(b) - a))) \\ &= (\sigma(b) - \sigma(\alpha(\sigma(b))) + \sigma(\alpha(a)), \beta(b)) \\ &= (a, c),\end{aligned}$$

por lo que θ es un epimorfismo. Por lo tanto, θ es un isomorfismo. Además, $\theta \circ \alpha = \iota$ y $\pi \circ \theta = \beta$.

(b) $\Rightarrow$ (c). Supongamos que existe $\rho : C \to B$ tal que $\beta \circ \rho = 1_C$. Consideremos el homomorfismo $\zeta : A \oplus C \to B$ dado por $\zeta(a, c) = \alpha(a) + \rho(c)$. Sea $(a, c) \in A \oplus C$ tal que $\zeta(a, c) = 0$, es decir, $\rho(c) = -\alpha(a)$. Se sigue que $c = \beta(\rho(c)) = \beta(-\alpha(a)) = 0$, y por ser α un monomorfismo, $-\alpha(a) = \rho(0) = 0$ implica que $a = 0$, por lo que $(a, c) = (0, 0)$. Por lo tanto, ζ es un monomorfismo. Sea $b \in B$. Notemos que $b - \rho(\beta(b)) \in \ker \beta = \operatorname{Im} \alpha$. Esto implica que existe $a \in A$ tal que $\alpha(a) = b - \rho(\beta(b))$, por lo que $b = \alpha(a) + \rho(\beta(b)) = \zeta(a, \beta(b))$ y entonces ζ es un epimorfismo. Por lo tanto, ζ es un isomorfismo. Además, $\zeta \circ \iota = \alpha$ y $\beta \circ \zeta = \pi$.

(c) $\Rightarrow$ (a) y (b). Sean $\iota' : C \to A \oplus C$ y $\pi' : A \oplus C \to A$ la inclusión y la proyección respectivamente. Si $\theta : B \to A \oplus C$ es un isomorfismo, $\theta \circ \alpha = \iota$ y $\pi \circ \theta = \beta$, entonces $\sigma = \pi' \circ \theta$ y $\rho = \theta^{-1} \circ \iota'$ satisfacen $\sigma \circ \alpha = \pi' \circ \iota = 1_A$ y $\beta \circ \rho = \pi \circ \iota' = 1_C$. □

Definición 2.4. *Dos extensiones* $E : 0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ *y* $E' : 0 \to A \xrightarrow{\alpha'} B' \xrightarrow{\beta'} C \to 0$ *de* A *por* C *son* ***equivalentes*** *si existe un isomorfismo* $\psi : B \to B'$ *tal que* $\alpha' = \psi \circ \alpha$ *y* $\beta = \beta' \circ \psi$.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \xrightarrow{\alpha} & B & \xrightarrow{\beta} & C & \longrightarrow & 0 \\ & & \| & & \psi\downarrow & & \| & & \\ 0 & \longrightarrow & A & \xrightarrow{\alpha'} & B' & \xrightarrow{\beta'} & C & \longrightarrow & 0 \end{array}$$

Es de rutina comprobar que la relación "ser equivalentes" en el conjunto de extensiones de A por C es una relación de equivalencia.

Definición 2.5. $\operatorname{Ext}(C, A)$ *es el conjunto de clases de equivalencia de las extensiones de* A *por* C.

Denotaremos por $[E]$ a la clase de equivalencia de la extensión E.

Corolario 2.6. *Cualesquiera dos extensiones escindibles de A por C son equivalentes.*

Demostración. Sean $E : 0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ y $E' : 0 \to A \xrightarrow{\alpha'} B' \xrightarrow{\beta'} C \to 0$ dos extensiones escindibles de A por C. Por la Proposición 2.3 existen dos isomorfismos $\theta_1 : B \to A \oplus C$ y $\theta_2 : A \oplus C \to B'$ tales que $\theta_1 \circ \alpha = \iota$, $\pi \circ \theta_1 = \beta$, $\theta_2 \circ \iota = \alpha'$ y $\beta' \circ \theta_2 = \pi$, por lo que $\theta_2 \circ \theta_1 : B \to B'$ es un isomorfismo tal que $\theta_2 \circ \theta_1 \circ \alpha = \theta_2 \circ \iota = \alpha'$ y $\beta' \circ \theta_2 \circ \theta_1 = \pi \circ \theta_1 = \beta$, y por lo tanto, $[E] = [E']$. □

Dado que la extensión $E_0 : 0 \to A \xrightarrow{\iota} A \oplus C \xrightarrow{\pi} C \to 0$ es escindible, se cumple que $[E] = [E_0]$ para toda extensión escindible E.

Proposición 2.7. (a) *Dada una extensión $E : 0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ de A por C y un homomorfismo $\lambda : A \to A'$, existe una extensión $E' : 0 \to A' \xrightarrow{\alpha'} B' \xrightarrow{\beta'} C \to 0$ de A' por C y un homomorfismo $\mu : B \to B'$ tales que el diagrama*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \xrightarrow{\alpha} & B & \xrightarrow{\beta} & C & \longrightarrow & 0 \\ & & \lambda\downarrow & & \mu\downarrow & & \| & & \\ 0 & \longrightarrow & A' & \xrightarrow{\alpha'} & B' & \xrightarrow{\beta'} & C & \longrightarrow & 0 \end{array}$$

es conmutativo.

(b) *Si $E'' : 0 \to A' \xrightarrow{\alpha''} B'' \xrightarrow{\beta''} C \to 0$ y $\mu' : B \to B''$ cumplen las condiciones anteriores, entonces $[E'] = [E'']$, y el isomorfismo resultante $\psi : B' \to B''$ cumple $\psi \circ \mu = \mu'$.*

(c) *Si $[E] = [F]$, entonces $[E'] = [F']$.*

Demostración. (a) Consideremos el submódulo

$$K = \{\, (\lambda(a), -\alpha(a)) \mid a \in A \,\}$$

de $A' \oplus B$, y definamos $B' = (A' \oplus B)/K$, $\alpha'(a') = (a', 0) + K$, $\beta'((a', b) + K) = \beta(b)$ y $\mu(b) = (0, b) + K$ para todo $a' \in A'$, $b \in B$.

(i) *α' es un monomorfismo.* Es claro que α' es un homomorfismo. Sea $a' \in A'$ tal que $\alpha'(a') = 0 + K$, es decir, $(a', 0) \in K$. Entonces existe $a \in A$ tal que $(a', 0) = (\lambda(a), -\alpha(a))$. Dado que α es monomorfismo, $-\alpha(a) = 0$ implica $a = 0$. Por lo tanto, $a' = \lambda(a) = \lambda(0) = 0$.

(ii) *β' un es epimorfismo.* Es claro que β' es un homomorfismo. Además, dado que β es epimorfismo, se sigue inmediatamente que β' también lo es.

(iii) *E' es exacta en B'.* Si $a' \in A'$, entonces

$$\beta'(\alpha'(a')) = \beta'((a', 0) + K) = \beta(0) = 0,$$

es decir, $\operatorname{Im}\alpha' \subseteq \ker\beta'$. Por otro lado, si $(a', b) + K \in \ker\beta'$, entonces $b \in \ker\beta = \operatorname{Im}\alpha$, por lo que existe $a \in A$ tal que $\alpha(a) = b$. Ahora notemos que

$$\begin{aligned}(a', b) + K &= (a', \alpha(a)) + K \\ &= (a' + \lambda(a), 0) + (-\lambda(a), \alpha(a)) + K \\ &= (a' + \lambda(a), 0) + K \\ &= \alpha'(a' + \lambda(a)).\end{aligned}$$

Se sigue que $\ker\beta' \subseteq \operatorname{Im}\alpha'$. Por lo tanto, $\operatorname{Im}\alpha' = \ker\beta'$.

Lo anterior demuestra que $E' : 0 \to A' \xrightarrow{\alpha'} B' \xrightarrow{\beta'} C \to 0$ es una extensión de A' por C.

(iv) *μ hace conmutar el diagrama*

$$\begin{aligned}\mu(\alpha(a)) &= (0, \alpha(a)) + K \\ &= (0, \alpha(a)) + (\lambda(a), -\alpha(a)) + K \\ &= (\lambda(a), 0) + K \\ &= \alpha'(\lambda(a)),\end{aligned}$$

y

$$\begin{aligned}\beta'(\mu(b)) &= \beta'((0, b) + K) \\ &= \beta(b).\end{aligned}$$

(b) Ahora sea $E'' : 0 \to A' \xrightarrow{\alpha''} B'' \xrightarrow{\beta''} C \to 0$ otra extensión de A' por C y sea $\mu' : B \to B''$ tal que $\alpha'' \circ \lambda = \mu' \circ \alpha$ y $\beta'' \circ \mu' = \beta$. Sea $\psi : B' \to B''$ dada por $\psi((a', b) + K) = \alpha''(a') + \mu'(b)$. Claramente, ψ es un homomorfismo.

(i) *ψ es un monomorfismo.* Si $\psi((a',b)+K)=0$, entonces $\alpha''(a')=-\mu'(b)$, por lo que $-\beta(b)=-\beta''(\mu'(b))=\beta''(\alpha''(a'))=0$. Esto significa que $b\in\ker\beta=\operatorname{Im}\alpha$, por lo que existe $a\in A$ tal que $\alpha(a)=b$. Se sigue que $\alpha''(a')=-\mu'(b)=-\mu'(\alpha(a))=\mu'(\alpha(-a))=\alpha''(\lambda(-a))$, y dado que α'' es monomorfismo, $a'=\lambda(-a)$. Por lo tanto, $(a',b)+K=(\lambda(-a),-\alpha(-a))+K=0+K$.

(ii) *ψ es un epimorfismo.* Si $b''\in B''$, entonces $\beta''(b'')\in C$; al ser β epimorfismo, existe $b\in B$ tal que $\beta(b)=\beta''(b'')$, y dado que $\beta''\circ\mu'=\beta$, $\beta(b)=\beta''(\mu'(b))=\beta''(b'')$, es decir, $b''-\mu'(b)\in\ker\beta''=\operatorname{Im}\alpha''$. Se sigue que existe $a'\in A'$ tal que $\alpha''(a')=b''-\mu'(b)$. Por lo tanto, $\psi((a',b)+K)=\alpha''(a')+\mu'(b)=b''$.

(iii) $\psi(\alpha'(a'))=\psi((a',0)+K)=\alpha''(a')+\mu'(0)=\alpha''(a')$, es decir, $\psi\circ\alpha'=\alpha''$; $\beta''(\psi((a',b)+K))=\beta''(\alpha''(a'))+\beta''(\mu'(b))=0+\beta(b)=\beta'((a',b)+K)$, es decir, $\beta''\circ\psi=\beta'$.

Por lo tanto, $[E']=[E'']$. Además, $\psi(\mu(b))=\psi((0,b)+K)=\alpha''(0)+\mu'(b)=\mu'(b)$, es decir, $\psi\circ\mu=\mu'$.

(c) Por último, sean $E:0\to A\xrightarrow{\alpha}B\xrightarrow{\beta}C\to 0$ y $F:0\to A\xrightarrow{\gamma}D\xrightarrow{\delta}C\to 0$ equivalentes con $\varphi:B\to D$ isomorfismo, y sean $E':0\to A'\xrightarrow{\alpha'}B'\xrightarrow{\beta'}C\to 0$, $F':0\to A'\xrightarrow{\gamma'}D'\xrightarrow{\delta'}C\to 0$ y $\mu:D\to D'$ como en el inciso (a). Entonces $\xi=\mu\circ\varphi:B\to D'$ cumple $\gamma'\circ\lambda=\xi\circ\alpha$ y $\delta'\circ\xi=\beta$, ya que

$$\gamma'\circ\lambda=\mu\circ\gamma=\mu\circ\varphi\circ\alpha=\xi\circ\alpha,$$

y

$$\beta=\delta\circ\varphi=\delta'\circ\mu\circ\varphi=\delta'\circ\xi,$$

por lo que por el inciso (b), $[E']=[F']$. □

Denotamos por λE a la extensión E' de A' por C construida en la proposición anterior a partir de la extensión E de A por C y del homomorfismo λ.

Proposición 2.8. (a) *Dada una extensión $E:0\to A\xrightarrow{\alpha}B\xrightarrow{\beta}C\to 0$ de A por C y un homomorfismo $\nu:C'\to C$, existe una extensión $E':0\to A\xrightarrow{\alpha'}B'\xrightarrow{\beta'}C'\to 0$ de A por C' y un homomorfismo*

$\mu : B' \to B$ *tales que el diagrama*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \xrightarrow{\alpha'} & B' & \xrightarrow{\beta'} & C' & \longrightarrow & 0 \\ & & \| & & \mu\downarrow & & \nu\downarrow & & \\ 0 & \longrightarrow & A & \xrightarrow{\alpha} & B & \xrightarrow{\beta} & C & \longrightarrow & 0 \end{array}$$

es conmutativo.

(b) *Si* $E'' : 0 \to A \xrightarrow{\alpha''} B'' \xrightarrow{\beta''} C' \to 0$ *y* $\mu' : B'' \to B$ *cumplen las condiciones anteriores, entonces* $[E'] = [E'']$*, y el isomorfismo resultante* $\psi : B'' \to B'$ *cumple* $\mu \circ \psi = \mu'$*.*

(c) *Si* $[E] = [F]$*, entonces* $[E'] = [F']$*.*

Demostración. (a) Consideremos el submódulo

$$B' = \{\, (b, c') \mid \beta(b) = \nu(c') \,\}$$

de $B \oplus C'$, y definamos $\alpha'(a) = (\alpha(a), 0)$, $\beta'(b, c') = c'$ y $\mu(b, c') = b$ para todo $a \in A$, $(b, c') \in B'$.

(i) α' *es un monomorfismo.* Es claro que α' es un homomorfismo. Además, dado que α es monomorfismo, se sigue inmediatamente que α' también lo es.

(ii) β' *es un epimorfismo.* Es claro que β' es un homomorfismo. Sea $c' \in C'$; entonces $\nu(c') \in C$, y al ser β epimorfismo, existe $b \in B$ tal que $\beta(b) = \nu(c')$. Por lo tanto, $(b, c') \in B'$ y $\beta'(b, c') = c'$.

(iii) E' *es exacta en* B'. Si $a \in A$, entonces $\beta'(\alpha'(a)) = \beta'(\alpha(a), 0) = 0$, es decir, $\operatorname{Im} \alpha' \subseteq \ker \beta'$. Por otro lado, si $(b, c') \in \ker \beta'$, entonces $c' = 0$. Además, $\beta(b) = \nu(c') = \nu(0) = 0$, por lo que $b \in \ker \beta = \operatorname{Im} \alpha$. Entonces existe $a \in A$ tal que $\alpha(a) = b$, es decir, $(b, c') = (\alpha(a), 0) \in \operatorname{Im} \alpha'$. Se sigue que $\ker \beta' \subseteq \operatorname{Im} \alpha'$. Por lo tanto, $\operatorname{Im} \alpha' = \ker \beta'$.

Lo anterior demuestra que $E' : 0 \to A \xrightarrow{\alpha'} B' \xrightarrow{\beta'} C' \to 0$ es una extensión de A por C'.

(iv) μ *hace conmutar el diagrama.* $\mu(\alpha'(a)) = \mu(\alpha(a), 0) = \alpha(a)$, y $\beta(\mu(b, c')) = \beta(b) = \nu(c') = \nu(\beta'(b, c'))$.

(b) Ahora sea $E'' : 0 \to A \xrightarrow{\alpha''} B'' \xrightarrow{\beta''} C' \to 0$ otra extensión de A por C' y sea $\mu' : B'' \to B$ tal que $\mu' \circ \alpha'' = \alpha$ y $\beta \circ \mu' = \nu \circ \beta''$. Sea $\psi : B'' \to B'$ dada por $\psi(b'') = (\mu'(b''), \beta''(b''))$. Claramente, ψ es un homomorfismo.

(i) *ψ es un monomorfismo.* Si $\psi(b'') = 0$, entonces $\mu'(b'') = 0$ y $\beta''(b'') = 0$. Esto significa que $b'' \in \ker \beta'' = \operatorname{Im} \alpha''$, por lo que existe $a \in A$ tal que $\alpha''(a) = b''$. Entonces $\alpha(a) = \mu'(\alpha''(a)) = \mu'(b'') = 0$, y dado que α es monomorfismo, se sigue que $a = 0$. Por lo tanto, $b'' = \alpha''(0) = 0$.

(ii) *ψ es un epimorfismo.* Sea $(b, c') \in B'$. Por ser β'' epimorfismo, existe $b'' \in B''$ tal que $\beta''(b'') = c'$. Además, $\beta(b) = \nu(c') = \nu(\beta''(b'')) = \beta(\mu'(b''))$, por lo que $b - \mu'(b'') \in \ker \beta = \operatorname{Im} \alpha$. Esto significa que existe $a \in A$ tal que $b - \mu'(b'') = \alpha(a) = \mu'(\alpha''(a))$. Por lo tanto, $\psi(\alpha''(a) + b'') = (\mu'(\alpha''(a) + b''), \beta''(\alpha''(a) + b'')) = (b, c')$.

(iii) $\psi(\alpha''(a)) = (\mu'(\alpha''(a)), \beta''(\alpha''(a))) = (\alpha(a), 0) = \alpha'(a)$, es decir, $\psi \circ \alpha'' = \alpha'$; $\beta'(\psi(b'')) = \beta'(\mu'(b''), \beta''(b'')) = \beta''(b'')$, es decir, $\beta' \circ \psi = \beta''$.

Por lo tanto, $[E'] = [E'']$. Además, $\mu(\psi(b'')) = \mu(\mu'(b''), \beta''(b'')) = \mu'(b'')$, es decir, $\mu \circ \psi = \mu'$.

(c) Por último, sean $E : 0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ y $F : 0 \to A \xrightarrow{\gamma} D \xrightarrow{\delta} C \to 0$ equivalentes con $\varphi : D \to B$ isomorfismo, y sean $E' : 0 \to A \xrightarrow{\alpha'} B' \xrightarrow{\beta'} C' \to 0$, $F' : 0 \to A \xrightarrow{\gamma'} D' \xrightarrow{\delta'} C' \to 0$ y $\mu : D' \to D$ como en el inciso (a). Entonces $\xi = \varphi \circ \mu : D' \to B$ cumple $\xi \circ \gamma' = \alpha$ y $\beta \circ \xi = \nu \circ \delta'$, ya que

$$\alpha = \varphi \circ \gamma = \varphi \circ \mu \circ \gamma' = \xi \circ \gamma',$$

y

$$\nu \circ \delta' = \delta \circ \mu = \beta \circ \varphi \circ \mu = \beta \circ \xi,$$

por lo que por el inciso (b), $[E'] = [F']$. □

Denotamos por $E\nu$ a la extensión E' de A por C' construida en la proposición anterior a partir de la extensión E de A por C y del homomorfismo ν.

Teorema 2.9. *Sea $E : 0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ una extensión de A por C, y sean $\lambda : A \to A'$ y $\nu : C' \to C$ homomorfismos. Entonces las extensiones $\lambda(E\nu)$ y $(\lambda E)\nu$ de A' por C' son equivalentes.*

Demostración. Consideremos el siguiente diagrama:

$$\begin{array}{rccccccccc} E\nu: & 0 & \longrightarrow & A & \xrightarrow{\alpha_\nu} & B_\nu & \xrightarrow{\beta_\nu} & C' & \longrightarrow & 0 \\ & & & \downarrow{\scriptstyle \lambda} & & \downarrow{\scriptstyle \mu_\nu} & & \| & & \\ \lambda(E\nu): & 0 & \longrightarrow & A' & \xrightarrow{\alpha'_\nu} & B'_\nu & \xrightarrow{\beta'_\nu} & C' & \longrightarrow & 0 \end{array}$$

Por la Proposición 2.8, tenemos que

- $B_\nu = \{\, (b, c') \mid \beta(b) = \nu(c') \,\} \leq B \oplus C'$,
- $\alpha_\nu(a) = (\alpha(a), 0)$,
- $\beta_\nu(b, c') = c'$

para todo $a \in A$, $(b, c') \in B_\nu$, y por la Proposición 2.7,

- $B'_\nu = (A' \oplus B_\nu)/K_\nu$ donde $K_\nu = \{\, (\lambda(a), -\alpha_\nu(a)) \mid a \in A \,\} \leq A' \oplus B_\nu$,
- $\alpha'_\nu(a') = (a', (0, 0)) + K_\nu$,
- $\beta'_\nu((a', b_\nu) + K_\nu) = \beta_\nu(b_\nu)$,

para todo $a' \in A'$, $b_\nu \in B_\nu$. Por otro lado, tenemos también el siguiente diagrama:

$$\begin{array}{rccccccccc} (\lambda E)\nu: & 0 & \longrightarrow & A' & \xrightarrow{\alpha'_\lambda} & B'_\lambda & \xrightarrow{\beta'_\lambda} & C' & \longrightarrow & 0 \\ & & & \| & & \downarrow{\scriptstyle \mu_\lambda} & & \downarrow{\scriptstyle \nu} & & \\ \lambda E: & 0 & \longrightarrow & A' & \xrightarrow{\alpha_\lambda} & B_\lambda & \xrightarrow{\beta_\lambda} & C & \longrightarrow & 0 \end{array}$$

Por la Proposición 2.7, tenemos que

- $B_\lambda = (A' \oplus B)/K_\lambda$ donde $K_\lambda = \{\, (\lambda(a), -\alpha(a)) \mid a \in A \,\} \leq A' \oplus B$,
- $\alpha_\lambda(a') = (a', 0) + K_\lambda$,

- $\beta_\lambda((a', b) + K_\lambda) = \beta(b)$

para todo $a' \in A'$, $b \in B$, y por la Proposición 2.8,

- $B'_\lambda = \{\, (b_\lambda, c') \mid \beta_\lambda(b_\lambda) = \nu(c') \,\} \leq B_\lambda \oplus C'$,
- $\alpha'_\lambda(a') = (\alpha_\lambda(a'), 0)$,
- $\beta'_\lambda(b_\lambda, c') = c'$

para todo $a' \in A'$, $(b_\lambda, c') \in B'_\lambda$. Sea $\psi : B'_\lambda \to B'_\nu$ dada por

$$\psi((a', b) + K_\lambda, c') = (a', (b, c')) + K_\nu.$$

(i) *ψ está bien definida.* Sean (a'_1, b_1), $(a'_2, b_2) \in A' \oplus B$ tales que $(a'_1 - a'_2, b_1 - b_2) \in K_\lambda$ y sea $c' \in C'$. Esto significa que existe $a \in A$ tal que $\lambda(a) = a'_1 - a'_2$ y $-\alpha(a) = b_1 - b_2$. Entonces $(a'_1, (b_1, c')) - (a'_2, (b_2, c')) = (a'_1 - a'_2, (b_1 - b_2, 0)) = (\lambda(a), (-\alpha(a), 0)) \in K_\nu$, por lo que $\psi((a'_1, b_1) + K_\lambda, c') = \psi((a'_2, b_2) + K_\lambda, c')$ si $(a'_1, b_1) + K_\lambda = (a'_2, b_2) + K_\lambda$ y c' es fijo.

(ii) *ψ es un homomorfismo.* Es de rutina comprobar que ψ es un homomorfismo.

(iii) *ψ es un monomorfismo.* Supongamos que $\psi((a', b) + K_\lambda, c') = 0 + K_\nu$. Esto implica que

$$(a', (b, c')) \in K_\nu = \{\, (\lambda(a), -(\alpha(a), 0)) \mid a \in A \,\},$$

lo que a su vez significa que existe $a \in A$ tal que $a' = \lambda(a)$ y $(b, c') = (-\alpha(a), 0)$, por lo que $b = -\alpha(a)$ y $c' = 0$. Por lo tanto, $(a', b) = (\lambda(a), -\alpha(a)) \in K_\lambda$ y $((a', b) + K_\lambda, c') = (0 + K_\lambda, 0)$.

(iv) *ψ es un epimorfismo.* Es claro.

(v) $\psi(\alpha'_\lambda(a')) = \psi(\alpha_\lambda(a'), 0) = \psi((a', 0) + K_\lambda, 0) = (a', (0, 0)) + K_\nu = \alpha'_\nu(a')$, es decir, $\psi \circ \alpha'_\lambda = \alpha'_\nu$.

$\beta'_\nu(\psi((a', b) + K_\lambda, c')) = \beta'_\nu((a', (b, c')) + K_\nu) = \beta_\nu(b, c') = c' = \beta'_\lambda((a', b) + K_\lambda, c')$, es decir, $\beta'_\nu \circ \psi = \beta'_\lambda$.

Por lo tanto, $[\lambda(E\nu)] = [(\lambda E)\nu]$. □

Debido al Teorema 2.9, al construir una extensión de A' por C' utilizando una extensión E de A por C y homomorfismos $\lambda : A \to A'$ y $\nu : C' \to C$, podemos prescindir de los paréntesis y escribir simplemente $\lambda E \nu$.

3. Definición de la operación de grupo en $\mathrm{Ext}(C, A)$

Consideremos dos extensiones $E : 0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ y $E' : 0 \to A \xrightarrow{\alpha'} B' \xrightarrow{\beta'} C \to 0$ de A por C. Denotemos por $E \oplus E'$ a la extensión

$$0 \to A \oplus A \xrightarrow{(\alpha,\alpha')} B \oplus B' \xrightarrow{(\beta,\beta')} C \oplus C \to 0$$

de $A \oplus A$ por $C \oplus C$. Sean $\nabla : A \oplus A \to A$ y $\Delta : C \to C \oplus C$ los homomorfismos dados por $\nabla(a_1, a_2) = a_1 + a_2$ y $\Delta(c) = (c, c)$. Por medio de las Proposiciones 2.7 y 2.8, obtenemos una extensión $\nabla(E \oplus E')\Delta$ de A por C.

Proposición 3.1. *Si E, E', F y F' son extensiones de A por C, $[E] = [F]$ y $[E'] = [F']$, entonces $[\nabla(E \oplus E')\Delta] = [\nabla(F \oplus F')\Delta]$.*

Demostración. Sean

$$E : 0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0,$$

$$E' : 0 \to A \xrightarrow{\alpha'} B' \xrightarrow{\beta'} C \to 0,$$

$$F : 0 \to A \xrightarrow{\gamma} D \xrightarrow{\delta} C \to 0,$$

$$F' : 0 \to A \xrightarrow{\gamma'} D' \xrightarrow{\delta'} C \to 0$$

extensiones de A por C tales que $[E] = [F]$ y $[E'] = [F']$, y consideremos las extensiones

$$E \oplus E' : 0 \to A \oplus A \xrightarrow{(\alpha,\alpha')} B \oplus B' \xrightarrow{(\beta,\beta')} C \oplus C \to 0,$$

$$F \oplus F' : 0 \to A \oplus A \xrightarrow{(\gamma,\gamma')} D \oplus D' \xrightarrow{(\delta,\delta')} C \oplus C \to 0.$$

Dado que $[E] = [F]$ y $[E'] = [F']$, existen isomorfismos $\psi : B \to D$ y $\psi' : B' \to D'$ tales que $\psi \circ \alpha = \gamma$, $\delta \circ \psi = \beta$, $\psi' \circ \alpha' = \gamma'$ y $\delta' \circ \psi' = \beta'$. Entonces $(\psi, \psi') : B \oplus B' \to D \oplus D'$ es un isomorfismo con $(\psi, \psi') \circ (\alpha, \alpha') = (\gamma, \gamma')$ y $(\delta, \delta') \circ (\psi, \psi') = (\beta, \beta')$, por lo que $[E \oplus E'] = [F \oplus F']$. Por la Proposición 2.7(c), $[\nabla(E \oplus E')] = [\nabla(F \oplus F')]$, y por la Proposición 2.8, $[\nabla(E \oplus E')\Delta] = [\nabla(F \oplus F')\Delta]$. □

La Proposición 3.1 nos permite definir una operación binaria + en el conjunto $\mathrm{Ext}(C, A)$ de clases de equivalencia de extensiones de A por C.

Definición 3.2. *Si* $[E]$, $[E'] \in \mathrm{Ext}(C, A)$, *definimos* $[E] + [E'] = [\nabla(E \oplus E')\Delta]$.

Lema 3.3. *La operación binaria* $+$ *definida en* $\mathrm{Ext}(C, A)$ *es conmutativa.*

Demostración. Si $E : 0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ y $E' : 0 \to A \xrightarrow{\alpha'} B' \xrightarrow{\beta'} C \to 0$ son dos extensiones de A por C, entonces es claro que $\psi : B \oplus B' \to B' \oplus B$ dado por $\psi(b, b') = (b', b)$ es un isomorfismo, por lo que $[E \oplus E'] = [E' \oplus E]$ y por lo tanto $[\nabla(E \oplus E')\Delta] = [\nabla(E' \oplus E)\Delta]$, es decir, $[E] + [E'] = [\nabla(E \oplus E')\Delta] = [\nabla(E' \oplus E)\Delta] = [E'] + [E]$. □

Proposición 3.4. *Para toda* $[E] \in \mathrm{Ext}(C, A)$, $[E] + [E_0] = [E]$, *es decir, la clase de las extensiones escindibles es un elemento neutro respecto a* $+$.

Demostración. Procedamos a construir $\nabla(E \oplus E_0)\Delta$. Sea $D = B \oplus A \oplus C$. Entonces obtenemos $E \oplus E_0 : 0 \to A \oplus A \xrightarrow{(\alpha,\iota)} D \xrightarrow{(\beta,\pi)} C \oplus C \to 0$. Por la Proposición 2.8, obtenemos el siguiente diagrama conmutativo:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & A \oplus A & \xrightarrow{\alpha_\Delta} & D_\Delta & \xrightarrow{\beta_\Delta} & C & \longrightarrow & 0 \\
 & & \| & & \mu_\Delta \downarrow & & \Delta \downarrow & & \\
0 & \longrightarrow & A \oplus A & \xrightarrow{\alpha,\iota} & D & \xrightarrow{\beta,\pi} & C \oplus C & \longrightarrow & 0
\end{array}$$

donde

- $D_\Delta \leq D \oplus C$ tal que

$$\begin{aligned} D_\Delta &= \{\, (b, a, c_1, c_2) \mid (\beta, \pi)(b, a, c_1) = \Delta(c_2) \,\} \\ &= \{\, (b, a, c_1, c_2) \mid a \in A,\ (\beta(b), c_1) = (c_2, c_2) \,\} \\ &= \{\, (b, a, \beta(b), \beta(b)) \mid a \in A,\ b \in B \,\}; \end{aligned}$$

- $\alpha_\Delta(a_1, a_2) = (\alpha(a_1), a_2, 0, 0)$,
- $\beta_\Delta((b, a, \beta(b), \beta(b))) = \beta(b)$.

y por la Proposición 2.7, obtenemos el siguiente diagrama conmutativo:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & A \oplus A & \xrightarrow{\alpha_\Delta} & D_\Delta & \xrightarrow{\beta_\Delta} & C & \longrightarrow & 0 \\
 & & \nabla \downarrow & & \mu_\nabla \downarrow & & \| & & \\
0 & \longrightarrow & A & \xrightarrow{\alpha_\nabla} & D_\nabla & \xrightarrow{\beta_\nabla} & C & \longrightarrow & 0
\end{array}$$

donde

- $D_\nabla = (A \oplus D_\Delta)/K_\nabla$, donde K_∇ es tal que

$$\begin{aligned} K_\nabla &= \{\, (\nabla(a_1, a_2), -\alpha_\Delta(a_1, a_2)) \mid (a_1, a_2) \in A \oplus A \,\}, \\ &= \{\, (a_1 + a_2, -\alpha(a_1), -a_2, 0, 0) \mid (a_1, a_2) \in A \oplus A \,\}; \end{aligned}$$

- $\alpha_\nabla(a) = (a, 0, 0, 0, 0) + K_\nabla$,
- $\beta_\nabla((a_1, b, a_2, \beta(b), \beta(b)) + K_\nabla) = \beta(b)$.

Sea $\psi : D_\nabla \to B$ dada por

$$\psi((a_1, b, a_2, \beta(b), \beta(b)) + K_\nabla) = \alpha(a_1 + a_2) + b.$$

(i) *ψ está bien definida.* Sean

$$(a_1^1, b_1, a_2^1, \beta(b_1), \beta(b_1)),\ (a_1^2, b_2, a_2^2, \beta(b_2), \beta(b_2)) \in A \oplus D_\Delta$$

tales que su diferencia

$$(a_1^1 - a_1^2, b_1 - b_2, a_2^1 - a_2^2, \beta(b_1 - b_2), \beta(b_1 - b_2)) \in K_\nabla.$$

Esto significa que existen $a_1,\ a_2 \in A$ tales que $a_2 = a_2^2 - a_2^1$, $a_1 = a_1^1 - a_1^2 + a_2^1 - a_2^2$ y $-\alpha(a_1) = b_1 - b_2$. De esta última igualdad se sigue que $-\alpha(a_1^1 - a_1^2 + a_2^1 - a_2^2) = \alpha(a_1^2 + a_2^2) - \alpha(a_1^1 + a_2^1) = b_1 - b_2$, por lo que $\alpha(a_1^2 + a_2^2) + b_2 = \alpha(a_1^1 + a_2^1) + b_1$, es decir, $\psi((a_1^1, b_1, a_2^1, \beta(b_1), \beta(b_1)) + K_\nabla) = \psi((a_1^2, b_2, a_2^2, \beta(b_2), \beta(b_2)) + K_\nabla)$.

(ii) *ψ es un homomorfismo.* Es de rutina comprobar que ψ es un homomorfismo.

(iii) *ψ es un monomorfismo.* Supongamos que

$$\psi((a_1, b, a_2, \beta(b), \beta(b)) + K_\nabla) = 0.$$

Esto implica que $b = -\alpha(a_1 + a_2) \in \operatorname{Im}\alpha = \ker\beta$, por lo que $\beta(b) = 0$. Ahora, haciendo $a_3 = a_1 + a_2$ y $a_4 = -a_2$, tenemos que $a_1 = a_3 + a_4$, por lo que $(a_1, b, a_2, \beta(b), \beta(b)) + K_\nabla = (a_3 + a_4, -\alpha(a_3), -a_4, 0, 0) + K_\nabla = 0 + K_\nabla$.

(iv) *ψ es un epimorfismo.* Si $b \in B$, basta tomar $(0, b, 0, \beta(b), \beta(b)) + K_\nabla \in D_\nabla$.

(v) $\psi(\alpha_\nabla(a)) = \psi((a,0,0,0,0) + K_\nabla) = \alpha(a)$, es decir, $\psi \circ \alpha_\nabla = \alpha$.

$$\begin{aligned}\beta(\psi((a_1,b,a_2,\beta(b),\beta(b)) + K_\nabla) &= \beta(\alpha(a_1+a_2)+b)\\ &= \beta(b)\\ &= \beta_\nabla((a_1,b,a_2,\beta(b),\beta(b)) + K_\nabla),\end{aligned}$$

es decir, $\beta \circ \psi = \beta_\nabla$.

Por lo tanto, $[E] = [\nabla(E \oplus E_0)\Delta] = [E] + [E_0]$. □

Proposición 3.5. *Dada una extensión* $E:\ 0 \to A \overset{\alpha}{\to} B \overset{\beta}{\to} C \to 0$ *de* A *por* C, *sea* $\bar{E}:\ 0 \to A \overset{-\alpha}{\to} B \overset{\beta}{\to} C \to 0$. *Entonces* $[E] + [\bar{E}] = [E_0]$.

Demostración. Consideremos

$$E \oplus \overline{E}:\ 0 \to A \oplus A \overset{(\alpha,-\alpha)}{\to} B \oplus B \overset{(\beta,\beta)}{\to} C \oplus C \to 0.$$

Por la Proposición 2.8, obtenemos el siguiente diagrama conmutativo:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A\oplus A & \xrightarrow{\alpha_\Delta} & B_\Delta & \xrightarrow{\beta_\Delta} & C & \longrightarrow & 0\\ & & \| & & \downarrow{\scriptstyle \mu_\Delta} & & \downarrow{\scriptstyle \Delta} & & \\ 0 & \longrightarrow & A\oplus A & \xrightarrow{\alpha,-\alpha} & B\oplus B & \xrightarrow{\beta,\beta} & C\oplus C & \longrightarrow & 0 \end{array}$$

donde

- $B_\Delta \leq B \oplus B \oplus C$ es tal que

$$\begin{aligned}B_\Delta &= \{\,(b_1,b_2,c) \mid (\beta,\beta)(b_1,b_2) = \Delta(c)\,\}\\ &= \{\,(b_1,b_2,c) \mid (\beta(b_1),\beta(b_2)) = (c,c)\,\}\\ &= \{\,(b_1,b_2,\beta(b_1)) \mid \beta(b_1) = \beta(b_2)\,\}\\ &= \{\,(b_1,b_2,\beta(b_1)) \mid b_1 - b_2 \in \ker\beta\,\}\end{aligned}$$

  (notemos que esto significa que existe $a_0 \in A$ tal que $\alpha(a_0) = b_1 - b_2$),

- $\alpha_\Delta(a_1,a_2) = (\alpha(a_1), -\alpha(a_2), 0)$,

- $\beta_\Delta(b_1,b_2,\beta(b_1)) = \beta(b_1)$,

y por la Proposición 2.7, obtenemos el siguiente diagrama conmutativo:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A \oplus A & \xrightarrow{\alpha_\Delta} & B_\Delta & \xrightarrow{\beta_\Delta} & C & \longrightarrow & 0 \\ & & \nabla\downarrow & & \mu_\nabla\downarrow & & \| & & \\ 0 & \longrightarrow & A & \xrightarrow{\alpha_\nabla} & B_\nabla & \xrightarrow{\beta_\nabla} & C & \longrightarrow & 0 \end{array}$$

donde

- $B_\nabla = (A \oplus B_\Delta)/K_\nabla$, donde K_∇ es tal que

$$\begin{aligned} K_\nabla &= \{\, (\nabla(a_1, a_2), -\alpha_\Delta(a_1, a_2)) \mid (a_1, a_2) \in A \oplus A \,\}, \\ &= \{\, (a_1 + a_2, -\alpha(a_1), \alpha(a_2), 0) \mid (a_1, a_2) \in A \oplus A \,\}; \end{aligned}$$

- $\alpha_\nabla(a) = (a, 0, 0, 0) + K_\nabla$,
- $\beta_\nabla((a, b_1, b_2, \beta(b_1)) + K_\nabla) = \beta(b_1)$.

Sea $\psi : B_\nabla \to A \oplus C$ dada por

$$\psi((a, b_1, b_2, \beta(b_1)) + K_\nabla) = (a + a_0, \beta(b_1)),$$

donde $\alpha(a_0) = b_1 - b_2$.

(i) *ψ está bien definida.* Sean

$$(a_1, b_1, b_2, \beta(b_1)),\ (a_2, b_3, b_4, \beta(b_3)) \in A \oplus B_\Delta$$

tales que $b_1 - b_2 = \alpha(a_0)$ y $b_3 - b_4 = \alpha(a_0')$, y tales que su diferencia

$$(a_1 - a_2, b_1 - b_3, b_2 - b_4, \beta(b_1 - b_3)) \in K_\nabla.$$

Esto significa que existen a_3, $a_4 \in A$ tales que $a_1 - a_2 = a_3 + a_4$, $-\alpha(a_3) = b_1 - b_3$, $\alpha(a_4) = b_2 - b_4$ y $\beta(b_1 - b_3) = 0$; de esta última igualdad se infiere que $\beta(b_1) = \beta(b_3)$. Dado que $b_3 = b_1 + \alpha(a_3)$ y $b_4 = b_2 - \alpha(a_4)$, se sigue que $\alpha(a_0') = b_3 - b_4 = b_1 - b_2 + \alpha(a_3 + a_4) = \alpha(a_0) + \alpha(a_1 - a_2)$, es decir, $\alpha(a_0' - a_0) = \alpha(a_1 - a_2)$. Puesto que α es un monomorfismo, se sigue que $a_0' - a_0 = a_1 - a_2$, y por tanto $a_1 + a_0 = a_2 + a_0'$. Por consiguiente, $\psi((a_1, b_1, b_2, \beta(b_1)) + K_\nabla) = (a_1 + a_0, \beta(b_1)) = (a_2 + a_0', \beta(b_3)) = \psi((a_2, b_3, b_4, \beta(b_3)) + K_\nabla)$.

(ii) *ψ es un homomorfismo.* Es de rutina comprobar que ψ es un homomorfismo.

(iii) *ψ es un monomorfismo.* Supongamos que

$$\psi((a, b_1, b_2, \beta(b_1)) + K_\nabla) = (0, 0)$$

y que $b_1 - b_2 = \alpha(a_0)$. Esto significa que $a + a_0 = 0$ y $\beta(b_1) = 0$. Dado que $\beta(b_2) = \beta(b_1) = 0$, se sigue que $b_1,\ b_2 \in \ker\beta = \operatorname{Im}\alpha$, por lo que existen $a_1,\ a_2 \in A$ tales que $\alpha(a_1) = b_1$ y $\alpha(a_2) = b_2$. Entonces $\alpha(a_1 - a_2) = b_1 - b_2 = \alpha(a_0)$, y al ser α un monomorfismo, $a_1 - a_2 = a_0$. Esto implica que $a = a_2 - a_1$, y por lo tanto, $(a, b_1, b_2, \beta(b_1)) = (-a_1 + a_2, -\alpha(-a_1), \alpha(a_2), 0) \in K_\nabla$.

(iv) *ψ es un epimorfismo.* Sea $(a, c) \in A \oplus C$. Ya que β es un epimorfismo, existe $b \in B$ tal que $\beta(b) = c$. Entonces

$$\psi((a, b, b, \beta(b)) + K_\nabla) = (a, c).$$

(v) *El siguiente diagrama es conmutativo*:

$$\psi(\alpha_\nabla(a)) = \psi((a, 0, 0, 0) + K_\nabla) = (a, 0) = \iota(a),$$

es decir, $\psi \circ \alpha_\nabla = \iota$.

$\pi(\psi((a, b_1, b_2, \beta(b_1)) + K_\nabla) = \pi(a + a_0, \beta(b_1)) = \beta(b_1)$
$= \beta_\nabla((a, b_1, b_2, \beta(b_1)) + K_\nabla)$, es decir, $\pi \circ \psi = \beta_\nabla$.

$$\begin{array}{ccccccccc}
0 & \longrightarrow & A & \xrightarrow{\alpha_\nabla} & B_\nabla & \xrightarrow{\beta_\nabla} & C & \longrightarrow & 0 \\
 & & \| & & \psi\downarrow & & \| & & \\
0 & \longrightarrow & A & \xrightarrow{\iota} & A \oplus C & \xrightarrow{\pi} & C & \longrightarrow & 0
\end{array}$$

Por lo tanto, $[E] + [\bar{E}] = [E_0]$. □

Proposición 3.6. *La operación binaria $+$ definida en $\operatorname{Ext}(C, A)$ es asociativa.*

Demostración. Sean $[E_1]$, $[E_2]$, $[E_3] \in \operatorname{Ext}(C, A)$ tales que

$$E_i : 0 \to A \xrightarrow{\alpha_i} B_i \xrightarrow{\beta_i} C \to 0,$$

para $i = 1, 2, 3$. Procedamos a construir $([E_1] + [E_2]) + E_3$. Consideremos

$$E_1 \oplus E_2 : \ 0 \to A \oplus A \overset{(\alpha_1,\alpha_2)}{\to} B_1 \oplus B_2 \overset{(\beta_1,\beta_2)}{\to} C \oplus C \to 0.$$

Por la Proposición 2.8, obtenemos el siguiente diagrama conmutativo:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & A \oplus A & \xrightarrow{\alpha_\Delta} & B_\Delta & \xrightarrow{\beta_\Delta} & C & \longrightarrow & 0 \\
 & & \| & & \mu_\Delta \downarrow & & \Delta \downarrow & & \\
0 & \longrightarrow & A \oplus A & \xrightarrow{\alpha_1,\alpha_2} & B_1 \oplus B_2 & \xrightarrow{\beta_1,\beta_2} & C \oplus C & \longrightarrow & 0
\end{array}$$

donde

- $B_\Delta \leq B_1 \oplus B_2 \oplus C$ es tal que

$$\begin{aligned}
B_\Delta &= \{\, (b_1, b_2, c) \mid (\beta_1, \beta_2)(b_1, b_2) = \Delta(c) \,\} \\
&= \{\, (b_1, b_2, c) \mid (\beta_1(b_1), \beta_2(b_2)) = (c, c) \,\} \\
&= \{\, (b_1, b_2, \beta_2(b_2)) \mid \beta_1(b_1) = \beta_2(b_2) \,\},
\end{aligned}$$

- $\alpha_\Delta(a_1, a_2) = (\alpha_1(a_1), \alpha_2(a_2), 0)$,
- $\beta_\Delta(b_1, b_2, \beta_2(b_2)) = \beta_2(b_2)$,

y por la Proposición 2.7, obtenemos el siguiente diagrama conmutativo:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & A \oplus A & \xrightarrow{\alpha_\Delta} & B_\Delta & \xrightarrow{\beta_\Delta} & C & \longrightarrow & 0 \\
 & & \nabla \downarrow & & \mu_\nabla \downarrow & & \| & & \\
0 & \longrightarrow & A & \xrightarrow{\alpha_\nabla} & B_\nabla & \xrightarrow{\beta_\nabla} & C & \longrightarrow & 0
\end{array}$$

donde

- $B_\nabla = (A \oplus B_\Delta)/K_\nabla$, donde K_∇ es tal que

$$\begin{aligned}
K_\nabla &= \{\, (\nabla(a_1, a_2), -\alpha_\Delta(a_1, a_2)) \mid (a_1, a_2) \in A \oplus A \,\}, \\
&= \{\, (a_1 + a_2, -\alpha_1(a_1), -\alpha_2(a_2), 0) \mid (a_1, a_2) \in A \oplus A \,\};
\end{aligned}$$

- $\alpha_\nabla(a) = (a, 0, 0, 0) + K_\nabla$,
- $\beta_\nabla((a, b_1, b_2, \beta_2(b_2)) + K_\nabla) = \beta_2(b_2)$.

Ahora consideremos

$$(\nabla(E_1 \oplus E_2)\Delta) \oplus E_3 : 0 \to A \oplus A \overset{(\alpha_\nabla, \alpha_3)}{\to} B_\nabla \oplus B_3 \overset{(\beta_\nabla, \beta_3)}{\to} C \oplus C \to 0.$$

Nuevamente por la Proposición 2.8, obtenemos el siguiente diagrama conmutativo:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A \oplus A & \xrightarrow{\alpha} & B & \xrightarrow{\beta} & C & \longrightarrow & 0 \\ & & \| & & \mu\downarrow & & \Delta\downarrow & & \\ 0 & \longrightarrow & A \oplus A & \xrightarrow{\alpha_\nabla, \alpha_3} & B_\nabla \oplus B_3 & \xrightarrow{\beta_\nabla, \beta_3} & C \oplus C & \longrightarrow & 0 \end{array}$$

donde

- $B \leq B_\nabla \oplus B_3 \oplus C$ es tal que

$$\begin{aligned} B = \{&((a, b_1, b_2, \beta_2(b_2)) + K_\nabla, b_3, c) \ \mid \\ &(\beta_\nabla, \beta_3)((a, b_1, b_2, \beta_2(b_2)) + K_\nabla, b_3) = \Delta(c)\} \\ = \{&((a, b_1, b_2, \beta_2(b_2)) + K_\nabla, b_3, c) \mid (\beta_2(b_2), \beta_3(b_3)) = (c, c)\} \\ = \{&((a, b_1, b_2, \beta_2(b_2)) + K_\nabla, b_3, \beta_2(b_2)) \mid \beta_2(b_2) = \beta_3(b_3)\}, \end{aligned}$$

- $\alpha(a_1, a_2) = (\alpha_\nabla(a_1), \alpha_3(a_2), 0) = ((a_1, 0, 0, 0) + K_\nabla, \alpha_3(a_2), 0)$,
- $\beta((a, b_1, b_2, \beta_2(b_2)) + K_\nabla, b_3, \beta_2(b_2)) = \beta_2(b_2)$,

y nuevamente por la Proposición 2.7, obtenemos el siguiente diagrama conmutativo:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A \oplus A & \xrightarrow{\alpha} & B & \xrightarrow{\beta} & C & \longrightarrow & 0 \\ & & \nabla\downarrow & & \mu'\downarrow & & \| & & \\ 0 & \longrightarrow & A & \xrightarrow{\alpha'} & B' & \xrightarrow{\beta'} & C & \longrightarrow & 0 \end{array}$$

donde

- $B' = (A \oplus B)/K$, donde K es tal que

$$\begin{aligned} K = &\{(\nabla(a_1, a_2), -\alpha(a_1, a_2)) \mid (a_1, a_2) \in A \oplus A\}, \\ = &\{(a_1 + a_2, (-a_1, 0, 0, 0) + K_\nabla, -\alpha_3(a_2), 0)| \\ &(a_1, a_2) \in A \oplus A\}; \end{aligned}$$

- $\alpha'(a) = (a, K_\nabla, 0, 0) + K$,
- $\beta'((a_1, (a_2, b_1, b_2, \beta_2(b_2)) + K_\nabla, b_3, \beta_2(b_2)) + K) = \beta_2(b_2)$.

De igual manera procedamos a construir $[E_1] + ([E_2] + [E_3])$. Consideremos

$$E_2 \oplus E_3 : \ 0 \to A \oplus A \overset{(\alpha_2,\alpha_3)}{\to} B_2 \oplus B_3 \overset{(\beta_2,\beta_3)}{\to} C \oplus C \to 0.$$

Por la Proposición 2.8, obtenemos el siguiente diagrama conmutativo:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & A \oplus A & \xrightarrow{\gamma_\Delta} & D_\Delta & \xrightarrow{\delta_\Delta} & C & \longrightarrow & 0 \\
 & & \| & & \nu_\Delta \downarrow & & \Delta \downarrow & & \\
0 & \longrightarrow & A \oplus A & \xrightarrow{\alpha_2,\alpha_3} & B_2 \oplus B_3 & \xrightarrow{\beta_2,\beta_3} & C \oplus C & \longrightarrow & 0
\end{array}$$

donde

- $D_\Delta \leq B_2 \oplus B_3 \oplus C$ es tal que

$$\begin{aligned}
D_\Delta &= \{\, (b_2, b_3, c) \mid (\beta_2, \beta_3)(b_2, b_3) = \Delta(c) \,\} \\
&= \{\, (b_2, b_3, c) \mid (\beta_2(b_2), \beta_3(b_3)) = (c, c) \,\} \\
&= \{\, (b_2, b_3, \beta_2(b_2)) \mid \beta_2(b_2) = \beta_3(b_3) \,\}
\end{aligned}$$

- $\gamma_\Delta(a_1, a_2) = (\alpha_2(a_1), \alpha_3(a_2), 0)$,
- $\delta_\Delta(b_2, b_3, \beta_2(b_2)) = \beta_2(b_2)$,

y por la Proposición 2.7, obtenemos el siguiente diagrama conmutativo:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & A \oplus A & \xrightarrow{\gamma_\Delta} & D_\Delta & \xrightarrow{\delta_\Delta} & C & \longrightarrow & 0 \\
 & & \nabla \downarrow & & \nu_\nabla \downarrow & & \| & & \\
0 & \longrightarrow & A & \xrightarrow{\gamma_\nabla} & D_\nabla & \xrightarrow{\delta_\nabla} & C & \longrightarrow & 0
\end{array}$$

donde

- $D_\nabla = (A \oplus D_\Delta)/L_\nabla$, donde L_∇ es tal que

$$\begin{aligned}
L_\nabla &= \{\, (\nabla(a_1, a_2), -\gamma_\Delta(a_1, a_2)) \mid (a_1, a_2) \in A \oplus A \,\}, \\
&= \{\, (a_1 + a_2, -\alpha_2(a_1), -\alpha_3(a_2), 0) \mid (a_1, a_2) \in A \oplus A \,\};
\end{aligned}$$

- $\gamma_\nabla(a) = (a, 0, 0, 0) + L_\nabla$,
- $\delta_\nabla((a, b_2, b_3, \beta_2(b_2)) + L_\nabla) = \beta_2(b_2)$.

Ahora consideremos

$$E_1 \oplus (\nabla(E_2 \oplus E_3)\Delta) : 0 \to A \oplus A \overset{(\alpha_1,\gamma_\nabla)}{\to} B_1 \oplus D_\nabla \overset{(\beta_1,\delta_\nabla)}{\to} C \oplus C \to 0.$$

Nuevamente por la Proposición 2.8, obtenemos el siguiente diagrama conmutativo:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A \oplus A & \xrightarrow{\gamma} & D & \xrightarrow{\delta} & C & \longrightarrow & 0 \\ & & \| & & \nu\downarrow & & \Delta\downarrow & & \\ 0 & \longrightarrow & A \oplus A & \xrightarrow{\alpha_1,\gamma_\nabla} & B_1 \oplus D_\nabla & \xrightarrow{\beta_1,\delta_\nabla} & C \oplus C & \longrightarrow & 0 \end{array}$$

donde

- $D \leq B_1 \oplus D_\nabla \oplus C$ es tal que

$$\begin{aligned} D = \{&(b_1, (a, b_2, b_3, \beta_2(b_2)) + L_\nabla, c) \ \mid \\ &(\beta_1, \delta_\nabla)\ (b_1, (a, b_2, b_3, \beta_2(b_2)) + L_\nabla) = \Delta(c)\} \\ = \{&\, (b_1, (a, b_2, b_3, \beta_2(b_2)) + L_\nabla, c) \mid (\beta_1(b_1), \beta_2(b_2)) = (c, c)\, \} \\ = \{&\, (b_1, (a, b_2, b_3, \beta_2(b_2)) + L_\nabla, \beta_2(b_2)) \mid \beta_1(b_1) = \beta_2(b_2)\, \} \end{aligned}$$

- $\gamma(a_1, a_2) = (\alpha_1(a_1), \gamma_\nabla(a_2), 0) = (\alpha_1(a_1), (a_2, 0, 0, 0) + L_\nabla, 0)$,
- $\delta(b_1, (a, b_2, b_3, \beta_2(b_2)) + L_\nabla, \beta_2(b_2)) = \beta_2(b_2)$,

y nuevamente por la Proposición 2.7, obtenemos el siguiente diagrama conmutativo:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A \oplus A & \xrightarrow{\gamma} & D & \xrightarrow{\delta} & C & \longrightarrow & 0 \\ & & \nabla\downarrow & & \nu'\downarrow & & \| & & \\ 0 & \longrightarrow & A & \xrightarrow{\gamma'} & D' & \xrightarrow{\delta'} & C & \longrightarrow & 0 \end{array}$$

donde

- $D' = (A \oplus D)/L$, donde L es tal que

$$\begin{aligned} L &= \{\, (\nabla(a_1, a_2), -\gamma(a_1, a_2)) \mid (a_1, a_2) \in A \oplus A\, \}, \\ &= \{(a_1 + a_2, -\alpha_1(a_1), (-a_2, 0, 0, 0) + L_\nabla, 0) \mid \\ &\qquad (a_1, a_2) \in A \oplus A\}\,; \end{aligned}$$

- $\gamma'(a) = (a, 0, L_\nabla, 0) + L$,
- $\delta'((a_1, b_1, (a_2, b_2, b_3, \beta_2(b_2)) + L_\nabla, \beta_2(b_2)) + L) = \beta_2(b_2)$.

Sea $\psi : B' \to D'$ dada por

$$\psi((a_1, (a_2, b_1, b_2, \beta_2(b_2)) + K_\nabla, b_3, \beta_2(b_2)) + K)$$
$$= (a_1, b_1, (a_2, b_2, b_3, \beta_2(b_2)) + L_\nabla, \beta_2(b_2)) + L.$$

(i) *ψ está bien definida.* Sean $(a_1, (a_2, b_1, b_2, \beta_2(b_2))+K_\nabla, b_3, \beta_2(b_2))+K$ y $(a_1', (a_2', b_1', b_2', \beta_2(b_2')) + K_\nabla, b_3', \beta_2(b_2')) + K$ dos elementos de B' tales que

$$(a_1-a_1', (a_2-a_2', b_1-b_1', b_2-b_2', \beta_2(b_2-b_2'))+K_\nabla, b_3-b_3', \beta_2(b_2-b_2'))\in K.$$

Esto significa que $\beta_2(b_2 - b_2') = 0$, y que existen a_3, $a_4 \in A$ tales que $a_1 - a_1' = a_3 + a_4$, $-\alpha_3(a_4) = b_3 - b_3'$, y

$$(a_2 - a_2' + a_3, b_1 - b_1', b_2 - b_2', \beta_2(b_2 - b_2')) \in K_\nabla,$$

y esto último significa a su vez que existen a_5, $a_6 \in A$ tales que $a_2 - a_2' + a_3 = a_5 + a_6$, $-\alpha_1(a_5) = b_1 - b_1'$, y $-\alpha_2(a_6) = b_2 - b_2'$. Sean $a_3' = a_5$, $a_4' = a_3 + a_4 - a_5$, $a_5' = a_6$ y $a_6' = a_4$. Entonces

$$\begin{aligned} a_2 - a_2' + a_4' &= a_5 + a_6 - a_3 + a_3 + a_4 - a_5 \\ &= a_6 + a_4 \\ &= a_5' + a_6', \end{aligned}$$

$$\begin{aligned} b_2 - b_2' &= -\alpha_2(a_6) \\ &= -\alpha_2(a_5'), \end{aligned}$$

$$\begin{aligned} b_3 - b_3' &= -\alpha_3(a_4) \\ &= -\alpha_3(a_6'), \end{aligned}$$

por lo que $(a_2 - a_2' + a_4', b_2 - b_2', b_3 - b_3', \beta_2(b_2 - b_2')) \in L_\nabla$, es decir,

$$(a_2 - a_2', b_2 - b_2', b_3 - b_3', \beta_2(b_2 - b_2')) + L_\nabla = (-a_4', 0, 0, 0) + L_\nabla;$$

además,

$$\begin{aligned} a_1 - a_1' &= a_3 + a_4 \\ &= a_5 + a_3 + a_4 - a_5 \\ &= a_3' + a_4', \end{aligned}$$

$$\begin{aligned} b_1 - b_1' &= -\alpha_1(a_5) \\ &= -\alpha_1(a_3'), \end{aligned}$$

por lo que

$$(a_1 - a_1', b_1 - b_1', (a_2 - a_2', b_2 - b_2', b_3 - b_3', \beta_2(b_2 - b_2')) + L_\nabla, \beta_2(b_2 - b_2'))$$

$\in L$, es decir,

$$(a_1, b_1, (a_2, b_2, b_3, \beta_2(b_2)) + L_\nabla, \beta_2(b_2)) + L =$$

$$(a_1', b_1', (a_2', b_2', b_3', \beta_2(b_2')) + L_\nabla, \beta_2(b_2')) + L,$$

y por lo tanto,

$$\psi((a_1, (a_2, b_1, b_2, \beta_2(b_2)) + K_\nabla, b_3, \beta_2(b_2)) + K) =$$
$$\psi((a_1', (a_2', b_1', b_2', \beta_2(b_2')) + K_\nabla, b_3', \beta_2(b_2')) + K).$$

(ii) ψ *es un homomorfismo.* Es de rutina comprobar que ψ es un homomorfismo.

(iii) ψ *es un monomorfismo.* Supongamos que

$$\psi((a_1, (a_2, b_1, b_2, \beta_2(b_2)) + K_\nabla, b_3, \beta_2(b_2)) + K) \in L.$$

Esto significa que $\beta_2(b_2) = 0$, y que existen a_3, $a_4 \in A$ tales que $a_1 = a_3 + a_4$, $b_1 = -\alpha_1(a_3)$ y $(a_2 + a_4, b_2, b_3, \beta_2(b_2)) \in L_\nabla$, donde esto implica a su vez que existen a_5, $a_6 \in A$ tales que $a_2 + a_4 = a_5 + a_6$, $b_2 = -\alpha_2(a_5)$ y $b_3 = -\alpha_3(a_6)$. Sean $a_3' = a_1 - a_6$, $a_4' = a_6$, $a_5' = a_3$ y $a_6' = a_5$. Dado que $a_2 + a_4 = a_5 + a_6$, $a_2 + (a_3 + a_4) = a_3 + a_5 + a_6$, es decir, $a_2 + a_1 = a_5' + a_6' + a_6$; se sigue que $a_2 + a_3' = a_2 + (a_1 - a_6) = a_5' + a_6'$. Además, $b_1 = -\alpha_1(a_5')$, $b_2 = -\alpha_2(a_6')$, por lo que $(a_2 + a_3', b_1, b_2, \beta_2(b_2)) \in K_\nabla$, es decir, $(a_2, b_1, b_2, \beta_2(b_2)) + K_\nabla = (-a_3', 0, 0, 0) + K_\nabla$. También, $a_1 = (a_1 - a_6) + a_6 = a_3' + a_4'$ y $b_3 = -\alpha_3(a_4')$. Por lo tanto, $(a_1, (a_2, b_1, b_2, \beta_2(b_2)) + K_\nabla, b_3, \beta_2(b_2)) \in K$.

(iv) *ψ es un epimorfismo.* Es claro.

(v) *El siguiente diagrama es conmutativo*:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \xrightarrow{\alpha'} & B' & \xrightarrow{\beta'} & C & \longrightarrow & 0 \\ & & \| & & \psi\downarrow & & \| & & \\ 0 & \longrightarrow & A & \xrightarrow{\gamma'} & D' & \xrightarrow{\delta'} & C & \longrightarrow & 0 \end{array}$$

$\psi(\alpha'(a)) = \psi((a, K_\nabla, 0, 0) + K) = (a, 0, L_\nabla, 0) + L = \gamma'(a)$, es decir, $\psi \circ \alpha' = \gamma'$.

$$\begin{aligned} \delta'(\psi((a_1, (a_2, b_1, b_2, \beta_2(b_2)) + K_\nabla, b_3, \beta_2(b_2)) + K) = \\ \delta'((a_1, b_1, (a_2, b_2, b_3, \beta_2(b_2)) + L_\nabla, \beta_2(b_2)) + L) = \\ \beta_2(b_2) = \\ \beta'((a_1, (a_2, b_1, b_2, \beta_2(b_2)) + K_\nabla, b_3, \beta_2(b_2)) + K), \end{aligned}$$

es decir, $\delta' \circ \psi = \beta'$. Por lo tanto,

$$([E_1] + [E_2]) + [E_3] = [E_1] + ([E_2] + [E_3]).$$

□

De este modo hemos demostrado que $(\mathrm{Ext}(C, A), +)$ **es un grupo abeliano**.

Referencias

[1] Dummit, D.S., Foote, R.M., *Abstract Algebra. Third Edition.* John Wiley & Sons Inc., Hobken, New Jersey, 2004.

[2] Grillet, P.A., *Abstract Algebra. Second Edition.* Graduate Texts in Mathematics, **242**, Springer, New York, 2007.

[3] Hilton, P.J., Stammbach, U. *A Course in Homological Algebra*, Graduate Texts in Mathematics, **4**, Springer, New York, 1971.

[4] Mac Lane, S. *Homology*, Die Grundlehren der mathematischen Wissenschaften, **114**, Springer, Berlin Heidelberg, 1963.

[5] Weibel, C.A., *An Introduction to Homological Algebra.* Cambridge University Press, England, 1994.

Dirección del autor
Héctor Gabriel Salazar Pedroza
Universidad Autónoma Metropolitana,
Unidad Iztapalapa,
División de Ciencias Básicas e Ingeniería,
Departamento de Matemáticas.
Av. San Rafael Atlixco 186, Col. Vicentina
Del. Iztapalapa, C.P. 09340 México, D.F.
e-mail: `hgabrielsp@yahoo.com`

Mixba'al Rev. Met. de Mat.
Nº 1, Vol. I, Junio 2010. 65 – 88

Espacios L_p no conmutativos

Gildardo Barrientos Roberto Quezada

Resumen

Siguiendo la referencia [2], presentamos una construcción elemental de los espacios L_p no conmutativos para matrices complejas $n \times n$.

Palabras clave: Valores singulares, normas de Schatten, normas de Ky Fan, normas Unitariamente Invariantes, desigualdades tipo Minkowski.

Clasificación de la AMS: 15A60, 81P45.

1. Introducción.

Los espacios L_p no conmutativos, también llamados ideales de Schatten, son una herramienta fundamental en Análisis Matemático. Sus aplicaciones recientes en Probabilidad Cuántica y Teoría Cuántica de la Información son muy interesantes, al respecto, consúltense por ejemplo las referencias [3], [4], [5] y [9]. En este trabajo presentamos una construcción elemental de estos espacios en dimensión finita, es decir para matrices complejas $n \times n$, siguiendo la referencia [2]. Incluiremos sólo aquellas demostraciones que son ilustrativas e interesantes, el lector puede consultar la demostración de cada resultado incluido en este trabajo en [1]. Para el caso de dimensión infinita se puede consultar la referencia [6].

Los conceptos de mayorización y matriz estocástica son centrales en este trabajo. Introducimos estos conceptos en la Sección 2 y discutimos algunas de sus consecuencias importantes. En la Sección 3 discutimos los conceptos de función simétrica y calibrada y norma unitariamente invariante, con ayuda de los cuales definimos las p-normas de Schatten. En la Sección 4 presentamos las desigualdades tipo Minkowski de E. Carlen y E. Lieb [3], [4], así como su aplicación para demostrar la subaditividad fuerte de la entropía de von Neumann.

2. Mayorización y matrices biestocásticas.

2.4. Mayorización.

Sea $x = (x_1, ..., x_n) \in \mathbb{R}^n$, denotaremos por $x^{\downarrow}$ y $x^{\uparrow}$ a los vectores obtenidos a partir de x reordenando las coordenadas en forma decreciente y en forma creciente respectivamente, es decir:

(i) $x^{\downarrow} = (x_1^{\downarrow}, ..., x_n^{\downarrow})$ donde $x_1^{\downarrow} \geq x_2^{\downarrow} \geq ... \geq x_n^{\downarrow}$

(ii) $x^{\uparrow} = (x_1^{\uparrow}, ..., x_n^{\uparrow})$ donde $x_1^{\uparrow} \leq x_2^{\uparrow} \leq ... \leq x_n^{\uparrow}$.

Obsérvese que

$$x_j^{\uparrow} = x_{n-j+1}^{\downarrow} \qquad j = 1, ..., n.$$

Definición 2.1. *Dados x,$y \in \mathbb{R}^n$ decimos que x es* ***mayorizado*** *por y, escribiremos $x \prec y$, si se cumplen las condiciones:*

$$\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow}, \text{ para cada } 1 \leq k < n \text{ y} \tag{1}$$

$$\mathrm{Tr}\,(x) := \sum_{j=1}^{n} x_j^{\downarrow} = \sum_{j=1}^{n} y_j^{\downarrow} = \mathrm{Tr}\,(y). \tag{2}$$

A la última igualdad le llamaremos condición de la traza.

Ejemplo 2.2. *Si $x_j \geq 0$ y $\sum_{j=1}^{n} x_j = 1$, entonces*

$$\left(\frac{1}{n}, ..., \frac{1}{n}\right) \prec (x_1, ...x_n) \prec \mathbb{1} = (1, 0, ..., 0).$$

La noción de mayorización ocurre naturalmente en varios contextos de la vida real. Por ejemplo, en Física: supóngase que $\mathrm{Tr}\,(x) = \sum_{j=1}^{n} x_j = 1$ y que un sistema físico se describe mediante una cadena de Markov con espacio de estados $S = \{1, \cdots, n\}$, si $x_i \geq 0$ es la probabilidad de que el sistema se encuentre en el estado i; entonces dadas dos distribuciones de probabilidades $x = (x_1, \cdots, x_n)$, $y = (y_1, \cdots y_n)$, tenemos que

$$x \prec y \Longrightarrow H(x) \geq H(y)$$

donde $H(x) = -\sum_{i=1}^{n} x_i \log(x_i)$ es la entropía de Shannon, véase la Proposición 3.9. Esto significa que la distribución y es más desordenada (o caótica) que x.

Otro ejemplo ocurre en Economía, si $x_1, ..., x_n$, $y_1, ..., y_n$ denotan los ingresos de $n-$individuos, entonces $x \prec y$ significa que hay una distribución más equitativa en el estado x que en el estado y. El ejemplo anterior ilustra este hecho.

Denotaremos por $\mathbf{e}$ al vector $(1, 1, ..., 1)$ y para cualquier subconjunto I de $\{1, 2, ..., n\}$, $\mathbf{e}_I$ denotará la función indicadora de I y $|I|$ su cardinalidad. Nótese que podemos escribir $\mathrm{Tr}\,(x) := \sum_{j=i}^{n} \langle x, \mathbf{e}\rangle$, donde $\langle \cdot, \cdot \rangle$ es el producto interior en $\mathbb{R}^n$.

Las siguientes proposiciones se demuestran fácilmente a partir de la Definición 2.1.

Proposición 2.3. *Sea $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$, para cada $1 \leq k \leq n$, se tiene que*

$$\sum_{j=1}^{k} x_j^{\downarrow} = \max_{|I|=k} \langle x, \mathbf{e}_I \rangle.$$

Proposición 2.4. *$x \prec y$ si y sólo si para cada subconjunto I de $\{1, 2, ..., n\}$ existe J tal que $|I| = |J|$,*

$$\langle x, \mathbf{e}_I \rangle = \langle y, \mathbf{e}_J \rangle \qquad y \qquad \mathrm{Tr}\,(x) = \mathrm{Tr}\,(y).$$

Definición 2.5. *Diremos que un vector x es **submayorizado débilmente** por un vector y si $\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow}$, $k = 1, 2, ..., n$. Escribiremos $x \prec_\omega y$.*

*De manera análoga diremos que un vector x es **supermayorizado débilmente** por un vector y si $\sum_{j=1}^{k} x_j^{\uparrow} \geq \sum_{j=1}^{k} y_j^{\uparrow}$, $k = 1, 2, ..., n$. Y escribiremos $x \prec^\omega y$.*

No es difícil demostrar las siguientes propiedades.

Proposición 2.6. *(i) $x \prec y$ si y sólo si $x \prec_\omega y$, $x \prec^\omega y$.*

(ii) *Si* α *es un número positivo, entonces* $x \prec_\omega y \Rightarrow \alpha x \prec_\omega \alpha y$, $x \prec^\omega y \Rightarrow \alpha x \prec^\omega \alpha y$.

(iii) $x \prec_\omega y \Leftrightarrow -x \prec^\omega -y$.

(iv) $\forall \alpha \in \mathbb{R}$, $x \prec y \Rightarrow \alpha x \prec \alpha y$.

Cada una de las relaciones $\prec, \prec_\omega, \prec^\omega$ es reflexiva y transitiva, pero ninguna define un orden parcial. Por ejemplo, si $x \prec y$ y $y \prec x$, sólo podemos decir que $Py = x$, donde P es una matriz de permutación, es decir, si $P = (p_{ij})_{1 \leq i,j \leq n}$ entonces $p_{ij} = \delta_{\sigma(i)j}$, donde $\sigma \in S_n$. Escribiremos $P \in S_n$ si P es una matriz de permutación. Se demuestra fácilmente que tenemos una relación de equivalencia si definimos $x \sim y$ cuando $x = Py$ para alguna $P \in S_n$. Si denotamos por $\mathbb{R}^n_{sym} = \mathbb{R}^n / \sim$, entonces claramente $\prec$ define un orden parcial en $\mathbb{R}^n_{sym}$. Esta relación también es un orden parcial sobre el conjunto $\{x \in \mathbb{R}^n : x_1 \geq \cdots \geq x_n\}$. La misma afirmación se cumple para las relaciones $\prec_\omega$ y $\prec^\omega$.

Para $a, b \in \mathbb{R}$ escribiremos

$$a \vee b = \text{máx}(a, b), \quad a \wedge b = \text{mín}(a, b).$$

Y para $x, y \in \mathbb{R}^n$ definimos

$$x \vee y = (x_1 \vee y_1, x_2 \vee y_2, ..., x_n \vee y_n), \quad x \wedge y = (x_1 \wedge y_1, x_2 \wedge y_2, ..., x_n \wedge y_n).$$

Además, sean $x^+ = x \vee 0, \quad |x| = x \wedge (-x)$.

Nótese que x^+ es el vector obtenido al reemplazar las coordenadas negativas de x por ceros y $|x|$ es el resultado de tomar el valor absoluto en cada coordenada del vector.

Con las definiciones anteriores tenemos el siguiente resultado, que caracteriza la mayorización sin apelar al reordenamiento de las coordenadas.

Teorema 2.7. *Sean* $x, y \in \mathbb{R}^n$, *entonces se cumplen*

(i) $x \prec_\omega y$ *si y sólo si* $\forall t \in \mathbb{R}$, $\displaystyle\sum_{j=1}^{n} (x_j - t)^+ \leq \sum_{j=1}^{n} (y_j - t)^+$.

(ii) $x \prec^\omega y$ *si y sólo si para cualquier* $t \in \mathbb{R}$,

$$\sum_{j=1}^{n} (t - x_j)^+ \leq \sum_{j=1}^{n} (t - y_j)^+.$$

(iii) $x \prec y$ *si y sólo si* $\forall t \in \mathbb{R}, \quad \sum_{j=1}^{n} |x_j - t| \leq \sum_{j=1}^{n} |y_j - t|.$

2.5. Matrices biestocásticas.

En la presente sección definiremos el concepto de matriz biestocástica y demostraremos algunos resultados que relacionan a estas matrices con la mayorización.

Definición 2.8. *Dada una matriz* $A \in \mathbf{M}_n$*, decimos que es una matriz* ***biestocástica*** *si cumple*

$$a_{ij} \geq 0, \quad \sum_{i=1}^{n} a_{ij} = 1, \quad y \ \sum_{j=1}^{n} a_{ij} = 1, \ \forall\, 1 \leq i, j \leq n.$$

La matriz A se llama **estocástica** si sólo satisface la primera y última condiciones. Las matrices estocásticas aparecen en Probabilidad como funciones de transición de cadenas de Markov.

Definición 2.9. *Diremos que*

(i) $A : \mathbb{C}^n \to \mathbb{C}^n$ *preserva la positividad si la imagen de cada vector con coordenadas no negativas es un vector con coordenadas no negativas.*

(ii) $A \in \mathbf{M}_n$ *preserva la traza si* $\mathrm{Tr}\,(Ax) = \mathrm{Tr}\,(x)$ *para cualquier* x*.*

(iii) $A \in \mathbf{M}_n$*, preserva la unidad si* $A\boldsymbol{e} = \boldsymbol{e}$*.*

Proposición 2.10. *Una matriz* $A \in \mathbf{M}_n$*, es biestocástica si y sólo si cumple con las condiciones de la Definición 2.9.*

Demostración. Supongamos que $A \in \mathbf{M}_n$ es matriz biestocástica y sea $x \in \mathbb{C}^n$ tal que $x_j \geq 0$, para toda $j = 1, ..., n$. Entonces es claro que $y_i = \sum_{j=1}^{n} a_{ij} x_j \geq 0$, pues $a_{ij} \geq 0$, para toda $1 \leq i, j \leq n$, por tanto, A preserva positividad. Ahora,

$$\mathrm{Tr}\,(Ax) = \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} a_{ij} x_j \Big) = \sum_{j=1}^{n} x_j \Big(\sum_{i=1}^{n} a_{ij} \Big) = \sum_{j=1}^{n} x_j = \mathrm{Tr}\,(x),$$

pues $\sum_{i=1}^{n} a_{ij} = 1$. Es decir A preserva traza. Finalmente,

$$A\mathbf{e} = A(1,\ldots,1) = \Big(\sum_{j=1}^{n} a_{1j},\ldots,\sum_{j=1}^{n} a_{nj}\Big) = (1,1,...,1) = \mathbf{e}.$$

Recíprocamente, si A preserva la positividad, tomando e_j el j-ésimo elemento de la base canónica $1 \leq j \leq n$, obtenemos que $Ae_j = (a_{1j}, a_{2j}, \cdots, a_{nj})$ tiene coordenadas no negativas para cada j. Entonces $a_{ij} \geq 0$ para cada $1 \leq i, j \leq n$.

Si A preserva la traza, para cada vector $x \in \mathbb{C}^n$ se tiene la identidad $\sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}x_j\right) = \sum_{j=1}^{n} x_j$, así tomando $x = e_j$ tenemos $\left(\sum_{i=1}^{n} a_{ij} - 1\right) = 0$, para toda $j = 1, 2, \ldots, n$. En consecuencia $\sum_{i=1}^{n} a_{ij} = 1$ para toda $j = 1, \ldots, n$.

Finalmente, si A preserva la unidad tenemos que $\left(\sum_{j=1}^{n} a_{ij}\right)_i = (1,1,...,1)$ de donde se obtiene que $\sum_{j=1}^{n} a_{ij} = 1$, para toda $i = 1, \ldots, n$. Esto demuestra que la matriz es biestocástica. □

Teorema 2.11 (Birkhoff)**.** *Las matrices biestocásticas de tamaño $n \times n$ son un conjunto cerrado bajo la multiplicación de matrices y la toma de adjuntos, además este conjunto es convexo y sus puntos extremos son las matrices de permutación.*

Teorema 2.12. *Una matriz $A \in \mathbf{M}_n$ es biestocástica si y sólo si $Ax \prec x$ $\forall x$.*

Proposición 2.13. *Sea A una matriz biestocástica, entonces, todos sus valores propios tienen módulo ≤ 1, 1 es valor propio de A y $\|A\| = 1$, donde $\|A\| = \sup_{\|x\|=1} \|Ax\|$.*

Demostración. Sea A matriz biestocástica, por la Proposición 2.10 preserva la unidad y, por lo tanto, 1 es valor propio de A con vector propio asociado $\mathbf{e} = (1,1,...,1)$. Por el Teorema 2.12 tenemos que $Ax \prec x$, $\forall x$. Sea x vector propio de A con valor propio asociado λ,

entonces $\lambda x \prec x$ y por (iii) de la Proposición 2.7, lo anterior se cumple si y sólo si $\forall t \in \mathbb{R}$

$$\sum_{j=1}^{k} |(\lambda x)_j - t| \leq \sum_{j=1}^{k} |x_j - t|, \;\; 1 \leq k \leq n.$$

Tomando $t = 0$ se obtiene que $|\lambda| \sum_{j=1}^{k} |x_j| \leq \sum_{j=1}^{k} |x_j|$, $1 \leq k \leq n$. Consecuentemente $|\lambda| \leq 1$.

Finalmente, $\|A\| = \text{máx}\{|\lambda_j| : \lambda \text{ es valor propio de A}\} = 1$, pues 1 es valor propio de A. □

Proposición 2.14. *Si $A \in \mathbf{M}_n$ es matriz biestocástica, entonces*

$$|Ax| \leq A(|x|),$$

donde $|x| = (|x_1|, ..., |x_n|)$ y escribiremos $x \leq y$ siempre que $x_i \leq y_i$ para toda $i = 1, ..., n$.

Proposición 2.15. *(i) Dados $x, y \in \mathbb{R}^2$ y $0 \leq t \leq 1$ se tiene*

$$(x_1, x_2) = (ty_1 + (1-t)y_2, (1-t)y_1, +ty_2) \quad \textit{si y sólo si } x \prec y.$$

(ii) Sean $x, y \in \mathbb{R}^n$ tales que x se obtiene promediando cualesquiera dos coordenadas de y como en (i), dejando fijas a las demás coordenadas. Entonces $x \prec y$.

Esto motiva la siguiente definición.

Definición 2.16. *Diremos que una aplicación lineal T sobre $\mathbb{R}^n$, es una t-transformación si existen $0 \leq t \leq 1$ e índices i, j tales que:*

$$Ty = (y_1, ..., \underbrace{ty_i + (1-t)y_j}_{i}, ..., \underbrace{(1-t)y_i + ty_j}_{j}, ..., y_n). \qquad (3)$$

Proposición 2.17. *Dada una $t-$transformación, entonces $Ty \prec y$ para toda $y \in \mathbb{R}^n$.*

Teorema 2.18. *Para toda $x, y \in \mathbb{R}^n$, las siguientes afirmaciones son equivalentes*

(i) $x \prec y$;

(ii) x *se obtiene a partir de* y *mediante un número finito de t-transformaciones;*

(iii) x *pertenece a la envolvente convexa de todos los vectores obtenidos al permutar las coordenadas de* y*;*

(iv) $x = Ay$ *para alguna matriz biestocástica* A*.*

Proposición 2.19. *(i) Si* $U = (u_{ij})$ *es una matriz unitaria, entonces la matriz* $(|u_{ij}|^2)$ *es biestocástica. Le llamaremos matriz uni-estocástica y se llamará ortoestocástica si* U *es ortogonal real.*

(ii) Si $x = Ay$ *para alguna matriz biestocástica* A*, entonces existe una matriz ortoestocástica* B *tal que* $x = By$

Teorema 2.20 (Lema de Schur)**.** *Sea* $A_{n\times n}$ *una matriz hermitiana* $(A = A^*)$*; sea* $\operatorname{diag}(A)$ *el vector formado con los elementos en la diagonal de* A *y sea* $\lambda(A)$ *el vector cuyas coordenadas son los valores propios de* A *especificados en cualquier orden. Entonces* $\operatorname{diag}(A) \prec \lambda(A)$*.*

Demostración. Como A es matriz hermitiana, por el teorema espectral existe una matriz U unitaria tal que $A = U\operatorname{diag}(\lambda(A))U^*$. Es decir,

$$A = \begin{pmatrix} \sum_{j=1}^n u_{1j}\lambda_j\overline{u_{1j}} & \dots & \sum_{j=1}^n u_{1j}\lambda_j\overline{u_{nj}} \\ \vdots & & \vdots \\ \sum_{j=1}^n u_{nj}\lambda_j\overline{u_{1j}} & \dots & \sum_{j=1}^n u_{nj}\lambda_j\overline{u_{nj}} \end{pmatrix}. \tag{4}$$

De esta manera $\operatorname{diag}(A) = \left(\sum_{j=1}^n u_{1j}\lambda_j\overline{u_{1j}}, \dots, \sum_{j=1}^n u_{nj}\lambda_j\overline{u_{nj}}\right)$, entonces podemos escribir

$$a_{ii} = \sum_j |u_{ij}|^2\lambda_j.$$

Esto implica que $\operatorname{diag}(A) = S\lambda(A)$, donde S es la matriz $S = (|u_{ij}|^2)$, que es biestocástica por la parte (i) de la Proposición 2.19. Ahora, una aplicación de la parte (i) del Teorema 2.18 nos permite concluir que $\operatorname{diag}(A) \prec \lambda(A)$. □

Teorema 2.21 (Principio del máximo de Ky Fan)**.** *Si* $\lambda(A)$ *es el vector de valores propios de una matriz hermitiana* A*, entonces para*

toda $k = 1, 2, ..., n$

$$\sum_{j=1}^{k} \lambda_j^{\downarrow} = \text{máx} \sum_{j=1}^{k} \langle x_j, Ax_j \rangle. \tag{5}$$

Donde el máximo es tomado sobre todas las $k-$*uplas de vectores ortonormales* $\{x_1, \cdots, x_k\}$ *en* $\mathbb{C}^n$.

Demostración. Sea $\{x_1, ..., x_k\}$ un conjunto ortonormal en $\mathbb{C}^n$, que podemos completar a una base ortonormal $\{x_1, ..., x_k, ..., x_n\}$ en $\mathbb{C}^n$. Si (a_{ij}) es la matriz de A respecto de esta base, entonces $a_{ij} = \langle x_i, Ax_j \rangle$, $1 \leq i, j \leq n$. Por el Lema de Schur, tenemos que $\sum_{j=1}^{m} \langle x_j, Ax_j \rangle^{\downarrow} \leq \sum_{j=1}^{m} \lambda_j^{\downarrow}$, $1 \leq m \leq n$. En particular, si $\{x_1, ..., x_k\}$ es un subconjunto ortonormal de vectores propios de A, entonces

$$\sum_{j=1}^{k} \langle x_j, Ax_j \rangle = \sum_{j=1}^{k} \langle x_j, \lambda_j x_j \rangle = \sum_{j=1}^{k} \lambda_j^{\downarrow}.$$

Es decir, la suma se satura alcanzando su máximo. Esto demuestra el Teorema. □

Proposición 2.22. *El principio del máximo de Ky Fan y el Lema de Schur son equivalentes.*

Demostración. En la demostración del resultado anterior se vió que el Lema de Schur implica el Principo del máximo de Ky Fan. Ahora, el Principio del máximo de Ky Fan implica que para cada $1 \leq k \leq n$ se tiene

$$\sum_{j=1}^{k} \lambda_j^{\downarrow} = \text{máx} \sum_{j=1}^{k} \langle x_j, Ax_j \rangle \geq \sum_{j=1}^{k} \langle x_j, Ax_j \rangle^{\downarrow} = \sum_{j=1}^{k} a_{ii}^{\downarrow}. \tag{6}$$

Además si $k = n$ el máximo del lado izquierdo se alcanza en la base ortonormal $\{x_1, \cdots, x_n\}$ de los vectores propios de A y la desigualdad anterior se convierte en igualdad. Es decir la condición de la traza también se cumple. Esto demuestra la proposición. □

Proposición 2.23. *Sean* A, B *matrices hermitianas, entonces para toda* $k = 1, ..., n$

$$\sum_{j=1}^{k} \lambda_j^{\downarrow}(A + B) \leq \sum_{j=1}^{k} \lambda_j^{\downarrow}(A) + \sum_{j=1}^{k} \lambda_j^{\downarrow}(B). \tag{7}$$

Demostración. Usando el Teorema 2.21

$$\sum_{j=1}^{k} \lambda_j^{\downarrow}(A+B) = \text{máx} \sum_{j=1}^{k} \langle x_j, (A+B)x_j \rangle \leq \text{máx} \sum_{j=1}^{k} \langle x_j, Ax_j \rangle \quad (8)$$

$$+ \text{máx} \sum_{j=1}^{k} \langle x_j, Bx_j \rangle = \sum_{j=1}^{k} \lambda_j^{\downarrow}(A) + \sum_{j=1}^{k} \lambda_j^{\downarrow}(B), \quad (9)$$

donde $\{x_1, \cdots, x_k\}$ es un subconjunto ortornormal en $\mathbb{C}^n$. □

Dada una matriz A de tamaño $n \times n$, decimos que B es la raíz cuadrada de la matriz A si $B \cdot B = A$ y el valor absoluto $|A|$ se define mediante la relación $|A| = \sqrt{A^*A}$. A los valores propios de la matriz $|A|$ tomando en cuenta sus multiplicidades se les llaman valores singulares de A.

Proposición 2.24.

(a) Para cualquier matriz A, sea $\widetilde{A}$ la matriz:

$$\widetilde{A} = \begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix}.$$

Entonces $\widetilde{A}$ es hermitiana y sus valores propios son los valores singulares de A junto con sus negativos.

(b) Sean $s_1(A), \ldots, s_n(A)$ los valores singulares de A ordenados en forma decreciente, entonces para cualesquiera matrices A, B y todo $k = 1, 2, \ldots, n$ se cumple

$$\sum_{j=1}^{k} s_j(A+B) \leq \sum_{j=1}^{k} s_j(A) + \sum_{j=1}^{k} s_j(B).$$

Demostración. Es fácil ver que $\widetilde{A}$ es hermitiana. Sea λ un valor propio de $\widetilde{A}$, con vector propio (X_1, X_2). Entonces tenemos que $AX_2 = \lambda X_1$ y $A^*X_1 = \lambda X_2$, como $\widetilde{A}$ es hermitiana $\lambda \in \mathbb{R}$, consecuentemente $|A|^2X_2 = A^*AX_2 = \lambda A^*X_1 = \lambda^2 X_2$. Por lo tanto $|\lambda| = \sqrt{\lambda^2}$ es un valor singular de A. Reciprocamente, si $W^*AQ = S$ es la descomposición de valores singulares de A, es decir, $S = \text{diag}(s_1(A), ..., s_n(A))$, W es la matriz unitaria cuyas columnas W_j son los vectores propios de

A^*A y Q es la matriz unitaria cuyas columnas Q_j son los vectores propios de AA^* correspondientes a los valores propios $s_j^2(A)$, $1 \leq j \leq n$. Entonces para cada j tenemos que

$$AQ_j = s_j(A)W_j, \quad \text{y} \quad A^*W_j = s_j(A)Q_j,$$

es decir, $s_j(A)$ y $-s_j(A)$ son valores propios de $\tilde{A}$ con vectores propios asociados (W_j, Q_j) y $(-W_j, Q_j)$, respectivamente. Esto demuestra (a).

Sean $\widetilde{A}$ y $\widetilde{B}$ como en (a). Por la Proposición 2.23 tenemos que

$$\sum_{j=1}^{k} \lambda_j^{\downarrow}(\widetilde{A} + \widetilde{B}) \leq \sum_{j=1}^{k} \lambda_j^{\downarrow}(\widetilde{A}) + \sum_{j=1}^{k} \lambda_j^{\downarrow}(\widetilde{B}).$$

Pero por el inciso anterior, para $1 \leq k \leq n$ podemos escribir

$$\sum_{j=1}^{k} s_j(A + B) \leq \sum_{j=1}^{k} s_j(A) + \sum_{j=1}^{k} s_j(B),$$

pues los negativos de los valores propios de $|A|$ aparecen después de la posición n. □

Si en la proposición anterior, tomamos $k = 1$ obtenemos que $s_1(A+B) \leq s_1(A) + s_1(B)$, que es la desigualdad del triángulo para la norma de operadores, i.e., $||A + B|| \leq ||A|| + ||B||$.

3. Normas simétricas sobre $\mathbb{C}^n$

3.6. Normas en $\mathbb{C}^n$

Para cada $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$ y cada $1 \leq p \leq \infty$

$$||x||_p = \left(\sum_{j=1}^{n} |x_j|^p \right)^{1/p}$$
$$||x||_\infty = \max_{1 \leq i \leq n} |x_i|.$$

Es bien conocido que estas relaciones definen normas en $\mathbb{C}^n$.

Las siguientes propiedades se demuestran fácilmente.

Proposición 3.1. *Para toda $x \in \mathbb{C}^n$ y toda $1 \leq p \leq \infty$ se cumplen*

(i) Propiedad del módulo o calibración. $\|x\|_p = \| \, |x| \, \|_p$ *, donde el vector* $|x|$ *se define mediante* $|x| = (|x_1|, \ldots, |x_n|)$.

(ii) Propiedad de monotonía. $\|x\|_p \leq \|y\|_p$ *siempre que* $|x| \leq |y|$ *donde ésta última desigualdad significa que* $|x_j| \leq |y_j|$ *para toda* $1 \leq j \leq n$.

(iii) Propiedad de simetría. $\|x\|_p = \|Px\|_p \ \forall P \in S_n$.

Observación 3.2. *Supongamos que el espacio es de dimensión infinita y está dotado de la p-norma, es decir se trata del espacio* ℓ_p. *Las primeras dos condiciones en la proposición 3.1 se cumplen claramente pues las sucesiones en* ℓ_p *son absolutamente convergentes. Y por el teorema de reordenamiento de series la última condición en la Proposición 3.1 también se cumple. Entonces la proposición también es válida en los espacios* ℓ_p.

Proposición 3.3. *Una norma en* $\mathbb{C}^n$ *es calibrada si y sólo si es monótona.*

3.7. Normas simétricas y calibradas

Por la propiedad de calibración vista en la sección anterior, las normas simétricas y calibradas en $\mathbb{C}^n$ están determinadas por aquellas en $\mathbb{R}^n$.

Definición 3.4. *Una función* $\Phi : \mathbb{R}^n \to \mathbb{R}_+$ *es simétrica y calibrada si*

(i) Φ *es una norma.*

(ii) $\Phi(Px) = \Phi(x)$ *para toda* $x \in \mathbb{R}^n$ *y toda* $P \in S_n$.

(iii) $\Phi(\varepsilon_1 x_1, \ldots, \varepsilon_n x_n) = \Phi(x_1, \ldots, x_n)$, *donde* $\varepsilon_j = \pm 1$, $j = 1, 2, ..., n$.

(iv) $\Phi(\mathbf{1}) = \Phi(1, 0, \ldots, 0) = 1$, *decimos que* Φ *esta normalizada.*

Por la condición *(iii)* una función simétrica y calibrada está determinada por sus valores en $\mathbb{R}^n_+$

Definición 3.5. *Sea* $x \in \mathbb{R}^n$ *y supóngase que sus coordenadas son ordenadas de tal forma que* $|x_1| \geq |x_2| \geq \cdots \geq |x_n|$. *Para* $k = 1, 2, ..., n$ *la norma de Ky Fan se define como*

$$\Phi_{(k)}(x) = \sum_{j=1}^{k} |x_j|. \tag{10}$$

Por las propiedades del módulo, la función $\Phi_{(k)}$ es claramente una función simétrica y calibrada. Además $\Phi_{(1)}(x) = |x_1| = \|x\|_\infty$ y $\Phi_{(n)}(x) = \|x\|_1$. Usaremos también $\|\cdot\|_{(k)}$ para referirnos a la k-ésima norma de Ky Fan, con paréntesis en el subíndice para no confundirla con la p-norma.

Proposición 3.6.

(i) Toda función simétrica y calibrada es continua.

(ii) Sea Φ una función simétrica y calibrada. Entonces para toda $x \in \mathbb{R}^n$, $\|x\|_\infty \leq \Phi(x) \leq \|x\|_1$.

(iii) Sea Φ función simétrica y calibrada y sean $0 \leq t_j \leq 1$, entonces

$$\Phi(t_1 x_1, ..., t_n x_n) \leq \Phi(x).$$

(iv) Toda función Φ simétrica y calibrada es monótona en $\mathbb{R}^n_+$.

(v) Sean $x, y \in \mathbb{R}^n$ tales que $|x| \leq |y|$, entonces $\Phi(x) \leq \Phi(y)$.

Consideraremos funciones $\Phi : \mathbb{R}^n \to \mathbb{R}^m$, cuyo dominio es todo $\mathbb{R}^n$ o bien un subconjunto de este espacio que es convexo e invariante bajo la permutación de sus coordenadas. Dados dos vectores en $\mathbb{R}^n$, $x = (x_1, ..., x_n)$, $y = (y_1, ..., y_n)$, escribiremos $x \leq y$ si $x_j \leq y_j$, $1 \leq j \leq n$. Retomando el concepto mayorización presentado en la sección anterior introducimos ahora la siguiente.

Definición 3.7. *Diremos que una función $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ es*

(i) ***isotona*** *si*

$$x \prec y \Rightarrow \Phi(x) \prec_\omega \Phi(y) \tag{11}$$

(ii) ***fuertemente isotona*** *si*

$$x \prec_\omega y \Rightarrow \Phi(x) \prec_\omega \Phi(y); \tag{12}$$

(iii) ***monótona creciente*** *si $x \leq y$ implica $\Phi(x) \leq \Phi(y)$;*

(iv) ***convexa*** *si*

$$\Phi(tx + (1-t)y) \leq t\Phi(x) + (1-t)\Phi(y), \;\; 0 \leq t \leq 1, \;\; x, y \in \mathbb{R}^n.$$

Si $m = 1$ la submayorización $\prec_\omega$ en el lado derecho de (11) se convierte en desigualdad. En este caso las funciones isotonas también se llaman ***Schur-convexas*** *o bien* ***S-convexas****.*

Teorema 3.8. *Sea $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ una aplicación convexa. Supóngase que para toda $P \in S_n$ existe $P' \in S_m$ tal que*

$$\Phi(Px) = P'\Phi(x) \qquad \forall x \in \mathbb{R}^n. \tag{13}$$

Entonces Φ es isotona. Si además Φ es monótona creciente entonces es fuertemente isotona.

Proposición 3.9. *La función $-H$, donde H es la entropía de Shannon, es una función isotona.*

Demostración. Para $x \geq 0$ sea $\varphi(x) = x\log(x)$, con la convención de que $0\log(0) = 0$. Como φ es una función convexa se tiene que $-H(x) = \sum_{j=1}^{n} \varphi(x_j)$ es convexa. Además es claro que $-H(x)$ es invariante bajo permutaciones de las coordenadas de x.

Supóngase que $x \prec y$, entonces por el inciso (iii) del Teorema 2.18, para $1 \leq j \leq n$, existen $P_j \in S_n$ y $0 \leq t_j \leq 1$ con $\sum_{j=1}^{n} t_j = 1$ tales que $x = \sum_{j=1}^{n} t_j P_j y$. Por la convexidad de $-H$ y su invarianza bajo permutaciones se obtiene que

$$-H(x) \leq -\sum_j t_j H(P_j y) = -H(y).$$

□

Proposición 3.10. *Sean $x, y \in \mathbb{R}^n_+$. Si $x \prec_\omega y$ entonces $\Phi(x) \leq \Phi(y)$ para toda función Φ simétrica y calibrada.*

Demostración. Por ser norma Φ es convexa, por la simetría se cumple la identidad (13) para toda $P \in S_n$, con $P' = Id$. Entonces por el Teorema 3.8, Φ es fuertemente isotona, es decir $\Phi(x) \leq \Phi(y)$. □

Proposición 3.11. *Sea $\Phi_{(k)}(\cdot)$ la k-ésima norma de Ky Fan, para $k = 1, 2, ..., n$. Si $\Phi_{(k)}(x) \leq \Phi_{(k)}(y)$ para toda $k = 1, 2, ..., n$ y para toda $x, y \in \mathbb{R}^n$, entonces $\Phi(x) \leq \Phi(y)$ para toda función Φ simétrica y calibrada y para toda $x, y \in \mathbb{R}^n$.*

Demostración. Supóngase que $\Phi_{(k)}(x) \leq \Phi_{(k)}(y)$ para toda $k = 1, 2, \ldots, n$ por la Definición 3.5 tenemos que $|x| \prec_\omega |y|$, entonces de la Proposición 3.10 se tiene que para toda Φ función simétrica y calibrada $\Phi(|x|) \leq \Phi(|y|)$. En consecuencia $\Phi(x) \leq \Phi(y)$ por la propiedad de calibración. □

Proposición 3.12. *(Desigualdad de Young) Sean $p, q, r \in \mathbb{R}_+$ tales que $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$. Entonces para cada función simétrica y calibrada Φ y cada x, $y \in \mathbb{R}^n$ se cumple*

$$[\Phi(|x \cdot y|^r)]^{1/r} \leq [\Phi(|x|^p)]^{1/p} [\Phi(|y|^q)]^{1/q}. \tag{14}$$

La desigualdad que se obtiene tomando $r = 1$ en (14) es la desigualdad de Hölder para funciones simétricas y calibradas. Si tomamos $r = 1$ y $p = 2$ en (14) obtenemos la desigualdad de Cauchy-Schwarz para funciones simétricas y calibradas.

Si $\Phi = \Phi_{(n)}$ y $r = 1$ la desigualdad (14) se reduce a la bien conocida desigualdad de Hölder para las p-normas en $\mathbb{C}^n$.

Teorema 3.13 (Desigualdad de Minkowski)**.** *Sea Φ una función simétrica y calibrada, sea $p \geq 1$, entonces para toda $x, y \in \mathbb{R}^n$ se cumple*

$$[\Phi(|x + y|^p)]^{1/p} \leq [\Phi(|x|^p)]^{1/p} + [\Phi(|y|^p)]^{1/p}. \tag{15}$$

Nótese que si $\Phi = \Phi_{(n)}$ tenemos

$$\left(\sum_{j=1}^{n} |x_j + y_j|^p\right)^{1/p} \leq \left(\sum_{j=1}^{n} |x_j|^p\right)^{1/p} + \left(\sum_{j=1}^{n} |y_j|^p\right)^{1/p} \tag{16}$$

que es la desigualdad de Minkowski para la p-norma.

3.8. Normas en $\mathbb{C}^n$ unitariamente invariantes

En esta sección $\mathbb{C}^n$ denotará el espacio de Hilbert $\mathbb{C}^n$ dotado del producto interno $\langle \cdot, \cdot \rangle$ y con la norma asociada $\| \cdot \|$. Si $A \in M_n$, $\|A\|$ denotará su norma como operador, definida como

$$\|A\| = \sup_{\|x\|=1} \|Ax\|_{\mathbb{C}^n}. \tag{17}$$

Si $A \in M_n$, su módulo $|A|$ se define por $|A| = (A^*A)^{1/2}$, donde $(A^*A)^{1/2}$ es la raíz cuadrada positiva del operador positivo A^*A.

Denotaremos por $s(A) = (s_1(A), s_2(A), \ldots, s_n(A))$ al vector de valores singulares de A ordenados de manera que $s_1(A) \geq s_2(A) \geq \cdots \geq s_n(A)$.

Proposición 3.14. *Sea A operador lineal en $\mathbb{C}^n$, entonces*

$$\|A\| = \| \, |A| \, \| = s_1(A). \tag{18}$$

Proposición 3.15. *Si U, V son operadores unitarios sobre $\mathbb{C}^n$ entonces*

$$|U\,A\,V| = V^*|A|V \tag{19}$$

y además

$$\|A\| = \|U\,A\,V\|. \tag{20}$$

Demostración. Sean $U\,y\,V$ operadores unitarios, entonces

$$\begin{aligned} |U\,A\,V|^2 &= (U\,A\,V)^*(U\,A\,V) = (V^*\,A^*\,U^*)(U\,A\,V) \\ &= V^*\,A^*\,Id\,A\,V = V^*\,|A|^2\,V = V^*\,|A|\,VV^*\,|A|\,V \\ &= (V^*\,|A|\,V)^2. \end{aligned}$$

Así, extrayendo la raíz cuadrada tenemos $|U\,A\,V| = V^*|A|\,V$. Ahora usaremos esta relación para demostrar la segunda identidad de la proposición. Sea $x \in \mathbb{C}^n$

$$\begin{aligned} \|U\,A\,Vx\|^2 = \|V^*|A|\,Vx\|^2 &= \langle V^*|A|\,Vx, V^*|A|\,Vx\rangle \\ &= \langle V\,V^*|A|\,Vx, |A|\,Vx\rangle = \langle |A|\,Vx, |A|Vx\rangle. \end{aligned}$$

Como V es unitario se tiene que $\|Vx\| = \|x\|$; además $\| \, |A| \, \| = \|A\|$, por lo tanto $\|A\| = \|U\,A\,V\|$. □

Una norma en los operadores sobre $\mathbb{C}^n$ que satisface la propiedad (20) para todo U, V operadores unitarios, se llama unitariamente invariante. Denotaremos a tales normas mediante $|||\cdot|||$ y las normalizaremos de manera que $|||\mathrm{diag}(1, 0, \cdots, 0)||| = 1$.

El siguiente teorema relaciona las normas unitariamente invariantes con las funciones simétricas y calibradas, mediante los valores singulares.

Teorema 3.16. *(i) Sea Φ una función simétrica y calibrada sobre $\mathbb{R}^n$. Definimos la función $|||\cdot|||_\Phi$ sobre $\mathbf{M}_{n\times n}$ como sigue*

$$|||A|||_\Phi = \Phi(s(A)). \tag{21}$$

Entonces $|||\cdot|||_\Phi$ *es una norma.*

(ii) Sea $|||\cdot|||$ *una norma unitariamente invariante. Definimos la función* $\Phi_{|||\cdot|||}$ *sobre* $\mathbb{R}^n$ *mediante*

$$\Phi_{|||\cdot|||}(x) = |||\mathrm{diag}\,(x)|||. \tag{22}$$

Donde $\mathrm{diag}\,(x)$ *es la matriz diagonal con coordenadas* $x_1, x_2, \ldots, x_n$. *Entonces* $\Phi_{|||\cdot|||}$ *es una función simétrica y calibrada.*

Las funciones simétricas y calibradas conducen a varios ejemplos de normas unitariamente invariantes en el espacio de las matrices complejas $n \times n$. Dos clases de tales normas son:

(1) Las p-normas de Schatten, para $1 \leq p \leq \infty$,

$$\|A\|_p = \Phi_p(s(A)) = \left[\sum_{j=1}^{n} (s_j(A))^p\right]^{1/p} \tag{23}$$

$$\|A\|_\infty = \Phi_\infty(s(A)) = s_1(A) = \|A\|. \tag{24}$$

(2) Las k-normas de Ky Fan

$$\|A\|_{(k)} = \sum_{j=1}^{k} s_j(A) = \Phi_{(k)}(s(A)) \qquad \text{para } 1 \leq k \leq n. \tag{25}$$

Hemos visto que $\|A\|_{(1)} = \|A\|_\infty = \|A\|$ y $\|A\|_{(n)} = \|A\|_1 = \| \, |A| \, \|_1 = \mathrm{Tr}\,(|A|)$. En la base que diagonaliza a $|A|$ tenemos que

$$(\mathrm{Tr}|A|^p)^{\frac{1}{p}} = \Big(\sum_{j=1}^{n} s_j(A)^p\Big)^{\frac{1}{p}} = \|A\|_p. \tag{26}$$

La traza es un funcional sobre el espacio de matrices complejas $n \times n$ con propiedades similares a las de la integral. Debido a esta analogía y en vista de la identidad (26), al espacio $\mathbf{M}_{n\times n}(\mathbb{C})$ provisto con la norma $\|\cdot\|_p$ le llamaremos **espacio** L_p **no conmutativo** y lo denotaremos por $L_p(\mathbb{C}^n)$.

Igual que en el caso conmutativo el espacio $L_2(\mathbb{C}^n)$ tiene un estructura de espacio de Hilbert, con el producto interno de Hilbert-Schmidt que se define mediante

$$\langle A, B\rangle_{HS} = \mathrm{Tr}\big(A^*B\big).$$

La norma inducida por este producto interno es la norma Hilbert-Schmidt,

$$\|A\|_{HS} = \mathrm{Tr}\big(A^*A\big)^{\frac{1}{2}} = \mathrm{Tr}(|A|^2)^{\frac{1}{2}} = \Big(\sum_j (s_j(A))^2\Big)^{\frac{1}{2}} = \|A\|_2,$$

que coincide con la 2−norma de Schatten. A esta norma también se le llama norma de Frobenius. Si $A = (a_{ij})_{1\leq i,j\leq n}$ entonces después de un cálculo simple se ve que $\|A\|_2 = \left(\sum_{ij} |a_{ij}|^2\right)^{1/2}$. Es decir, la norma de Hilbert-Schmidt es la norma euclidiana de la matriz A pensada como vector en $\mathbb{C}^{n^2}$.

Proposición 3.17. *Si $1 \leq p < \infty$ y $\frac{1}{p} + \frac{1}{q} = 1$, entonces se cumple la desigualdad de Hölder para las p-normas de Schatten.*

Demostración. Se sigue de (14) con $r = 1$, $\Phi = \Phi_p$ y la identidad (23). □

El siguiente Teorema muestra la relevancia de las normas de Ky Fan.

Teorema 3.18 (Dominación de Ky Fan)**.** *Sean $A, B \in \mathbf{M}_{n\times n}$; si $\|A\|_{(k)} \leq \|B\|_{(k)}$ para $k = 1, 2, \ldots, n$. Entonces*

$$|||A||| \leq |||B|||,$$

para todas las normas unitariamente invariantes.

Demostración. Sea $|||\cdot|||$ una norma unitariamente invariante y $\Phi := \Phi_{|||\cdot|||}$ la función simétrica y calibrada asociada. Por la descomposición de valores singulares, para cualquier matriz A se tiene que $\mathrm{diag}(s(A)) = W^*AQ$ con W y Q matrices unitarias, consecuentemente $|||A||| = |||W^*AQ||| = |||\mathrm{diag}(s(A))|||$. Por otra parte, del Teorema 3.16 obtenemos que $|||A|||_\Phi = \Phi_{|||\cdot|||}(s(A)) = |||\mathrm{diag}(s(A))|||$, entonces

$$|||A||| = |||A|||_\Phi = \Phi_{|||\cdot|||}(s(A)).$$

Sean A, B matrices de tamaño $n \times n$. Si $\|A\|_{(k)} \leq \|B\|_{(k)}$ para $k = 1, 2, \ldots, n$; por la ecuación (25) tenemos que

$$\sum_{j=1}^{k} s_j(A) \leq \sum_{j=1}^{k} s_j(B) \qquad \text{para toda } 1 \leq k \leq n.$$

Entonces $s(A) \prec_\omega s(B)$. Por la Proposición 3.10 se tiene que $|||A||| = \Phi_{|||\cdot|||}(s(A)) \leq \Phi_{|||\cdot|||}(s(B)) = |||B|||$.

□

Recordemos que de la Proposición 3.6 se cumple que $\Phi_{(1)}(x) \leq \Phi(x) \leq \Phi_{(n)}(x)$ para toda $x \in \mathbb{C}^n$ y toda función simétrica y calibrada Φ. Entonces por el Teorema 3.16 tenemos que $\|A\|_\infty = \|A\| \leq |||A||| \leq \|A\|_{(n)} = \|A\|_1$, para toda $A \in \mathbf{M}_{n\times n}$ y toda norma unitariamente invariante $|||\cdot|||$.

4. Desigualdades tipo Minkowski.

La noción de traza parcial es central en esta parte, para una discusión más amplia de este operador el lector puede consultar, por ejemplo, las referencias [7], [8].

Definición 4.1. *Sean $\mathcal{H}_1$ y $\mathcal{H}_2$ espacios de Hilbert de dimensión finita. Sea $\{e_i\}_i$ una base ortonormal de $\mathcal{H}_2$. Sea ρ un operador sobre $\mathcal{H}_1 \otimes \mathcal{H}_2$. La traza parcial de ρ es el operador $\mathrm{Tr}_2 \rho \in \mathcal{B}(\mathcal{H}_1)$ definido para todos u, $v \in \mathcal{H}_1$ mediante*

$$\langle u, \mathrm{Tr}_2 \rho v\rangle_{\mathcal{H}_1} = \sum_i \langle u \otimes e_i, \rho(v \otimes e_i)\rangle_{\mathcal{H}_1\otimes\mathcal{H}_2}. \tag{27}$$

Se puede demostrar que el operador $\mathrm{Tr}_2 \rho$ no depende de la base ortonormal utilizada. Además, la traza parcial es un operador $\mathrm{Tr}_2 : \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2) \to \mathcal{B}(\mathcal{H}_1)$, que preserva la traza y la positividad. De hecho, es un operador completamente positivo, veáse por ejemplo, [7] y [8].

Denotaremos mediante $\rho_1 = \mathrm{Tr}_2 \rho$, a la traza parcial del operador $\rho \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ y de manera análoga $\rho_2 = \mathrm{Tr}_1 \rho$.

La siguientes desigualdades tipo Minkowski se deben a E. A. Carlen y E. Lieb, [3], [4]. Para las condiciones de igualdad consúltese [9]. Demostraremos sólo la primera desigualdad, la demostración de la segunda es más difícil.

Teorema 4.2 (Desigualdad Tipo Minkowski I)**.** *Sea $A \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ positivo. Entonces, para toda $p \geq 1$*

$$\left(\mathrm{Tr}_2(\mathrm{Tr}_1 A)^p\right)^{1/p} \leq \mathrm{Tr}_1(\mathrm{Tr}_2 A^p)^{1/p}. \tag{28}$$

Si $0 < p \leq 1$, la desigualdad se invierte.

Demostración. Por la desigualdad de Hölder, Proposición 3.17, para $\frac{1}{p}+\frac{1}{q}=1$ tenemos que

$$\mathrm{Tr}\,(B\mathrm{Tr}_1 A) \leq (\mathrm{Tr}\,B^q)^{\frac{1}{q}}\mathrm{Tr}\,((\mathrm{Tr}_1 A)^p)^{\frac{1}{p}} \tag{29}$$

para $B \in \mathcal{B}(\mathcal{H}_2)$ positivo. Tomando $B = \sum_j b_j|\beta_j\rangle\langle\beta_j|$, donde $b_j = \alpha\lambda_j$ con $\alpha = \left(\sum_j \lambda_j^q\right)^{\frac{1}{q}}$ y $\{\lambda_j\}$, $\{\beta_j\}$ los valores propios y la base ortonormal que diagonaliza a $\mathrm{Tr}_1 A$, respectivamente, es fácil ver que se cumple la igualdad en (29) y que $\mathrm{Tr}\,B^q = 1$. Entonces tenemos

$$\begin{aligned}\mathrm{Tr}\left((\mathrm{Tr}_1 A)^p\right)^{\frac{1}{p}} &= \mathrm{Tr}\left(B\mathrm{Tr}_1(A)\right) = \mathrm{Tr}\left((I\otimes B)A\right)\\ &= \sum_{i,j}\langle e_i\otimes\beta_j, (I\otimes B)A(e_i\otimes\beta_j)\rangle\\ &= \sum_{i,j}\langle e_i\otimes B\beta_j, A(e_i\otimes\beta_j)\rangle,\end{aligned}$$

donde $\{e_i\}$ es base ortonormal de $\mathcal{H}_1$.

Podemos estimar el lado derecho como sigue

$$\begin{aligned}\sum_{i,j}\langle e_i\otimes B\beta_j, A(e_i\otimes\beta_j)\rangle &\leq \sum_i\left(\sum_j\lambda_j^q\right)^{\frac{1}{q}}\left(\sum_j(\langle e_i\otimes\beta_j, A(e_i\otimes\beta_j)\rangle)^p\right)^{\frac{1}{p}}\\ &= \sum_i\left(\sum_j(\langle e_i\otimes\beta_j, A\,(e_i\otimes\beta_j)\rangle)^p\right)^{\frac{1}{p}}.\end{aligned}$$

Una aplicación del Teorema espectral en dimensión finita permite demostrar que

$$\langle e_i\otimes\beta_j, A(e_i\otimes\beta_j)\rangle \leq \left(\langle e_i\otimes\beta_j, A^p(e_i\otimes\beta_j)\rangle\right)^{\frac{1}{p}},$$

de donde se obtiene

$$\begin{aligned}(\mathrm{Tr}\,(\mathrm{Tr}_1 A^p))^{\frac{1}{p}} &\leq \sum_i\left(\sum_j\left(\langle e_i\otimes\beta_j, A^p(e_i\otimes\beta_j)\rangle\right)^{\frac{1}{p}}\right)\\ &= \sum_i\left(\langle e_i, \mathrm{Tr}_2 A^p e_i\rangle\right)^{\frac{1}{p}}.\end{aligned}$$

Si se elige la base $\{e_j\}_j$ como los valores propios de $\mathrm{Tr}_2 A^p$. Entonces el lado derecho de la desigualdad anterior es $\mathrm{Tr}\,(\mathrm{Tr}_2 A^p)^{\frac{1}{p}}$. Con esto se concluye la demostración. □

Teorema 4.3 (Desigualdad Tipo Minkowski II)**.** *Para $1 \leq q \leq p \leq 2$ y todo A operador positivo sobre $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$,*

$$\mathrm{Tr}_3 \left(\mathrm{Tr}_2 \left((\mathrm{Tr}_1 A^q)^{\frac{p}{q}} \right) \right)^{\frac{q}{p}} \leq \mathrm{Tr}_3 \left(\mathrm{Tr}_1 [(\mathrm{Tr}_2 A^p)^{\frac{q}{p}}] \right). \tag{30}$$

Para $0 \leq p \leq 1$, y cualquier $q \geq p$ la desigualdad se invierte.

4.9. Subaditividad Fuerte de la entropía de von Neumann

Consideremos una función $f(p) = \rho^p$, ρ un estado, i.e., un operador positivo de traza unitaria. Por cálculo funcional podemos escribir $\rho^p = \sum_j \lambda_j^p |e_j\rangle\langle e_j|$, entonces

$$\begin{aligned} \frac{d}{dp}(Tr\rho^p)\Big|_{p=1} &= \frac{d}{dp}\Big[Tr\Big(\sum_j \lambda_j^p |e_j\rangle\langle e_j|\Big)\Big]_{p=1} \\ &= \frac{d}{dp}\Big[\sum_j \lambda_j^p Tr\big(|e_j\rangle\langle e_j|\big)\Big]_{p=1} \\ &= \frac{d}{dp}\Big[\sum_j \lambda_j^p\Big]_{p=1} = \sum_j \lambda_j \log \lambda_j = -S(\rho). \end{aligned}$$

La serie de McLaurin de $f(x) = a^{x+1}$ alrededor de 0 tiene la forma $f(x) = a + xa\ln a + a(\ln a)^2 \frac{x^2}{2!} + \ldots$. Entonces para una matriz definida positiva A y $\varepsilon > 0$ suficientemente pequeño, tenemos

$$A^{1\pm\varepsilon} = A \pm \varepsilon\, A \log\, A \pm \varepsilon^2 \frac{A(\log\, A)^2}{2!} \pm \ldots \tag{31}$$

Teorema 4.4. *(Subaditividad fuerte de la entropía de von Neumann) Sea ρ_{123} un estado sobre el espacio de Hilbert $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$. Entonces*

$$S(\rho_{1,3}) + S(\rho_{2,3}) \geq S(\rho_{1,2,3}) + S(\rho_3),$$

donde S es la entropía de von Neumann.

Demostración. Sea $\rho = \rho_{1,2,3} \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3)$ un estado. Usaremos las siguientes notaciones: $\rho_{2,3} = \mathrm{Tr}_1 \rho$, $\rho_3 = \mathrm{Tr}_1 \mathrm{Tr}_2 \rho$, etc. Aplicamos traza parcial sobre el segundo factor en (31),

$$Tr_2(\rho^{1+\varepsilon}) = Tr_2\rho + \varepsilon\, Tr_2\rho \log\, \rho + \mathcal{O}(\varepsilon^2) = \rho_{1,3} + \varepsilon\, Tr_2\rho \log\, \rho + \mathcal{O}(\varepsilon^2).$$

Por otro lado, sabemos que $\frac{1}{1+\varepsilon} = \sum_{n\geq 0}(-\varepsilon)^n = 1-\varepsilon+\mathcal{O}(\varepsilon^2)$. Así

$$\begin{aligned}
&\Big[Tr_2(\rho^{1+\varepsilon})\Big]^{\frac{1}{1+\varepsilon}} = \big(\rho_{1,3}+\varepsilon\, Tr_2\rho\log\,\rho+\mathcal{O}(\varepsilon^2)\big)^{\frac{1}{1+\epsilon}}\\
=&\rho_{1,3}+\varepsilon\, Tr_2\rho\log\,\rho-\varepsilon\big(\rho_{1,3}+\varepsilon\, Tr_2\rho\log\,\rho\big)\log(\rho_{1,3}+\varepsilon\, Tr_2\rho\log\,\rho\big)\\
&\quad+\mathcal{O}(\varepsilon^2)\\
=&\rho_{1,3}+\varepsilon\, Tr_2\rho\log\,\rho-\epsilon\rho_{1,3}\log\rho_{1,3}+\mathcal{O}(\varepsilon^2).
\end{aligned}$$

Y tomando Tr_1 y Tr_3 obtenemos que

$$Tr_3\, Tr_1\Big[Tr_2(\rho^{1+\varepsilon})\Big]^{\frac{1}{1+\varepsilon}} = 1-\varepsilon S(\rho)+\varepsilon S(\rho_{1,3})+\mathcal{O}(\varepsilon^2). \tag{32}$$

De manera análoga se obtiene que

$$Tr_3\Big[Tr_2\big(\rho_{2,3}^{1+\varepsilon}\big)\Big]^{\frac{1}{1+\varepsilon}} = 1-\varepsilon S(\rho_{2,3})+\varepsilon S(\rho_3)+\mathcal{O}(\varepsilon^2) \tag{33}$$

Ahora, aplicando el Teorema 4.2, con $A=\rho$, $p=1+\varepsilon$, y usando que la traza parcial es una transformación completamente positiva, de (32) y (33) obtenemos,

$$1-\varepsilon S(\rho)+\varepsilon S(\rho_{1,3})+\mathcal{O}(\varepsilon^2) \leq 1-\varepsilon S(\rho_{2,3})+\varepsilon S(\rho_3)+\mathcal{O}(\varepsilon^2)$$

Equivalentemente

$$1-S(\rho)+S(\rho_{1,3})+\frac{\mathcal{O}(\varepsilon^2)}{\varepsilon} \leq 1-S(\rho_{2,3})+S(\rho_3)+\frac{\mathcal{O}(\varepsilon^2)}{\varepsilon}.$$

Haciendo $\varepsilon\to 0$, obtenemos la subaditividad fuerte para la entropía de von Neumann. □

5. Agradecimiento

Agradecemos al árbitro por varias observaciones y sugerencias que nos permitieron mejorar este artículo.

Referencias

[1] Barrientos Sánchez, G., *Espacios L_p no conmutativos y aplicaciones*, Tesis de Maestría, Posgrado en Matemáticas, UAM-I, México, 2010.

[2] Bhatia, R., *Matrix Analysis*, Springer-Verlag, New York U.S.A., 1997.

[3] Carlen, E.A. and Lieb, E., *A Minkowski type trace inequality and strong subadditivity of quantum entropy*, Advances in the Mathematical Science, AMS Transl., **189**, Series 2, 1999, 59-68.

[4] Carlen, E.A. and Lieb, E., *A Minkowski type trace inequality and strong subadditivity of quantum entropy II: convexity and concavity*, Lett. math. Phys., **83**, 2008, 107-126.

[5] Gibilisco, P., Imparato, D. and Isola, T., *Schrödinger Equation, L^p-Duality and the Geometry of Wigner-Yanase-Dyson Information*, in Quantum Probability and Related Topics, Proceedings of the 28th Conference, 2008, p. 157-164.

[6] McCarthy, C. A., $\mathcal{C}_p$, Israel J. Math., **5**, 1967, p. 249-271.

[7] Pantaleón, L. y Quezada, R., *Una introducción a la Teoría cuántica de la información*, Texto para el *Primer Coloquio del Departamento de Matemáticas*, Departamento de Matemáticas UAM-I, México, 2008.

[8] Petz, D.,*Quantum Information Theory and Quantum Statistics*, Springer, Berlin Heidelberg, 2008.

[9] Ruskai, M. B. and Jenčová, A., *A Unified Treatment of Convexity of Relative Entropy and Related Trace Functions, with Conditions for Equality*, arxiv:0903.2895v1 [quant-ph] 17 Mar 2009.

Dirección de los autores
Gildardo Barrientos
Universidad Autónoma Metropolitana,
Unidad Iztapalapa,
División de Ciencias Básicas e Ingeniería,
Departamento de Matemáticas.
Av. San Rafael Atlixco 186, Col. Vicentina
Del. Iztapalapa, C.P. 09340 México, D.F.
e-mail: gil_barrientos@hotmail.com

Roberto Quezada
Universidad Autónoma Metropolitana,
Unidad Iztapalapa,
División de Ciencias Básicas e Ingeniería,
Departamento de Matemáticas.
Av. San Rafael Atlixco 186, Col. Vicentina
Del. Iztapalapa, C.P. 09340 México, D.F.
e-mail: roqb@xanum.uam.mx

Mixba'al Rev. Met. de Mat.
Nº 1, Vol. I, Junio 2010. 89 – 102

Las conjeturas de Weil para curvas elípticas

Marco Antonio Sánchez Mirafuentes

Resumen

En este trabajo se presentan las conjeturas de *Weil* y se da una demostración en el caso de una curva elíptica. Estas conjeturas tienen antecedentes en los trabajos de Gauss, concretamente en las *Disquisitiones* Art. 358, donde se da el número de soluciones de congruencias de las forma $ax^3 - by^3 \equiv 1 (\mathrm{mod}\, p)$, con p un número primo de la forma $p = 3n+1$. Tiempo después en los trabajos de *Artin* y *Hasse* se formula la conjetura de Riemann para campos de funciones, demostrada por Hasse en el caso elíptico y por Weil en el caso de una curva arbitraria. En la demostración que se presenta a continuación son importantes las nociones de morfismo *separable*, el morfismo de *Frobenius* y el apareamiento de *Weil*.

Palabras clave: variedad proyectiva, función zeta, curva algebraica, curva elíptica, diferenciales, morfismo de Frobenius, isogenia, apareamiento de Weil, módulo de Tate, cohomología ℓ-ádica.

Clasificación de la AMS: 14H52, 11G20

1. Funciones zeta y las conjeturas de Weil

Denotaremos la cardinalidad de un conjunto C por $N(C)$. Si k es un campo finito con q elementos, que también denotaremos por $\mathbb{F}_q$, se tiene el lema siguiente

Lema 1.1. *Sea k un campo finito con q elementos, entonces la cardinalidad de $\mathbb{P}^n(k)$ es $1 + q + \ldots + q^n$.*

Demostración. Definiendo $\varphi : k^{n+1} - \{0\} \to \mathbb{P}^n(k)$ como

$$\varphi(x_0, \ldots x_n) = [x_0, \ldots x_n],$$

se tiene que φ es suprayectiva, y dado un punto cualquiera $[x_0, \dots x_n]$ en $\mathbb{P}^n(k)$ observe que

$$\varphi^{-1}[x_0, \dots x_n] = \{(\lambda x_0, \dots \lambda x_n) : \lambda \in k - \{0\}\}.$$

ya que si $(x'_0, \dots x'_n) \in \varphi^{-1}[x_0, \dots x_n]$ entonces existe un $\lambda \neq 0$ en k tal que $x'_i = \lambda x_i$ para cada i; así el conjunto $\varphi^{-1}[x_0, \dots x_n]$ tiene $q-1$ elementos. Puesto que φ es suprayectiva y $N(k^{n+1} - \{0\}) = q^n - 1$, se tiene que $q^{n+1} - 1 = (q-1)N(\mathbb{P}^n(k))$, es decir, $N(\mathbb{P}^n(k)) = 1 + q + \dots + q^n$. □

Recuerde ahora que si k es un campo finito, con q elementos, para cada $s \in \mathbb{N}$ existe un campo k_s, y sólo uno, tal que $k \subseteq k_s$ y $N(k_s) = q^s$.

Ahora sea X una variedad proyectiva en $\mathbb{P}^n(k)$. Como los polinomios homogéneos que definen a X pertenecen a

$$k[x_0, \dots, x_n] \subseteq k_s[x_0, \dots, x_n],$$

podemos pensar a X como una variedad proyectiva en $\mathbb{P}^n(k_s)$ y por lo tanto tiene sentido considerar el número $N_s = N(X)$ de puntos de la variedad X vista en $\mathbb{P}^n(k_s)$.

Definición 1.2. *Si k es un campo finito con q elementos, la función zeta de la variedad proyectiva $X \subseteq \mathbb{P}^n(k)$ es la serie:*

$$Z_X(u) = \exp\Big(\sum_{s=1}^{\infty} \frac{N_s u^s}{s}\Big),$$

donde $\exp(u) = \sum_{s=0}^{\infty} \frac{u^s}{s!}$. *La variable u toma valores en $\mathbb{C}$, por lo tanto podemos pensar a Z_X como una serie formal de potencias o, utilizando la desigualdad siguiente (que se obtiene del lema 1.1)*

$$N_s \leq \frac{q^{s(n+1)} - 1}{q^s - 1} \leq (n+1)q^{sn},$$

vemos que $f(u) = \sum_{s=1}^{\infty} \frac{N_s u^s}{s}$ *se puede comparar con la serie* $g(u) = \sum_{s=1}^{\infty} \frac{(q^n u)^s}{s}$, *y concluir que Z_X es holomorfa en alguna vecindad del 0.*

Las conjeturas de Weil. Sea X una variedad proyectiva lisa de dimensión n definida sobre $k = \mathbb{F}_q$, y sea $Z_X(u)$ la función zeta de X.

(1) **La racionalidad de la función zeta.** Z_X es una función racional, es decir, es un cociente de polinomios con coeficientes racionales.

(2) **Ecuación funcional.** Sea E el número de autointersección de la diagonal Δ de $X \times X$. Entonces Z_X satisface la ecuación funcional siguiente:

$$Z_X\Big(\frac{1}{q^n u}\Big) = \pm q^{n\frac{E}{2}} u^E Z_X(u).$$

(3) **Analogía con la hipótesis de Riemann.** Es posible escribir

$$Z_X(U) = \frac{P_1(u)P_3(u)\dots P_{2n-1}(u)}{P_0(u)P_2(u)\dots P_{2n}(u)},$$

donde $P_0(u) = 1 - u$; $P_{2n} = 1 - q^n u$; y para cada $1 \leq i \leq 2n-1$, $P_i(u)$ es un polinomio con coeficientes enteros que se puede escribir de la forma siguiente:

$$P_i(u) = \prod(1 - \alpha_{ij} u),$$

donde los α_{ij} son enteros algebraicos con $|\alpha_{ij}| = q^{\frac{1}{2}}$.

Estas conjeturas aparecen por primera vez en el artículo de *A. Weil*, Number of solutions of equations over finite fields, Bull. Amer. Math. Soc. 55 (1949), 497 - 508; el artículo en cuestión empieza tratando el caso de variedades algebraicas de la forma $V(a_0 X_0^n + \ldots + a_r X_r^n)$ donda cada $a_i \in k$; de aquí obtiene la racionalidad de la función zeta asociada a tal variedad y después enuncia las conjeturas en el caso general. Cabe señalar que *A. Weil* probó sus conjeturas en el caso de curvas, es decir, variedades algebraicas de dimensión 1.

Ejemplo 1.3.

Calculamos la función zeta de $\mathbb{P}^n$ usando el lema 1.1. En efecto, como el resultado es independiente del campo finito usado, esto significa que $N_s = 1 + q^s + \ldots + q^{sn}$, por lo que:

$$\sum_{s=1}^{\infty} \frac{(1 + q^s + \ldots + q^{sn})u^s}{s} = \sum_{s=1}^{\infty} \frac{u^s}{s} + \sum_{s=1}^{\infty} \frac{(uq)^s}{s} + \ldots \sum_{s=1}^{\infty} \frac{(uq^n)^s}{s}.$$

Y recordando que $-\log(1-u) = \sum_{s=1}^{\infty} \frac{u^s}{s}$ obtenemos:

$$Z_{\mathbb{P}^n}(u) = \frac{1}{(1-u)(1-qu)\dots(1-q^n u)}.$$

En particular esto muestra que $Z_{\mathbb{P}^n}$ es una función racional, en concordancia con la conjetura de Weil correspondiente.

2. Curvas algebraicas

Por una *curva* o *curva algebraica* entenderemos una variedad *proyectiva* (irreducible) de dimensión 1. No sólo nos interesamos por las variedades proyectivas sino también por los morfismos entre ellas y en este sentido se tiene el teorema siguiente.

Teorema 2.1. *Si $\phi : C_1 \to C_2$ es un morfismo de curvas, entonces ϕ es constante o suprayectivo.*

Demostración. Como $\phi : C_1 \to C_2$ es regular sabemos que $\phi(C_1)$ es un subconjunto algebraico de C_2. Si $\phi(C_1)$ es finito, entonces sólo puede tener un punto, ya que en caso contrario como ϕ es continua, podríamos escribir a C_1 como unión disjunta de conjuntos algebraicos, contradiciendo que C_1 es irreducible y, en este caso, ϕ es constante.

Ahora si $\phi(C_1)$ no fuera finito entonces $\dim(\phi(C_1)) > 0$ además $\dim(\phi(C_1)) \leq \dim(C_2) = 1$ por lo tanto $\dim(\phi(C_1)) = \dim(C_2)$, y como $\phi(C_1) \subseteq C_2$ se tiene que, ([1] capítulo I ejercicio 1.10), $\phi(C_1) = C_2$ en particular, ϕ es suprayectiva. □

Ahora supongamos que tenemos curvas C_1/k y C_2/k y $\phi : C_1 \to C_2$ un morfismo no constante definido sobre k. Entonces ϕ induce una función $\phi^* : k(C_2) \to k(C_1)$ entre los campos de funciones $k(C_2)$ y $k(C_1)$ mediante $\phi^*(f) = f \circ \phi$.

Puesto que ϕ no es constante, por el teorema 2.1, ϕ es suprayectiva, por lo que ϕ^* es inyectiva. Además, ϕ^* es un morfismo de campos que deja fijo al campo k. Por lo anterior ya podemos suponer que $k(C_1)$ es una extensión de $\phi^* k(C_2)$, y el teorema siguiente nos da una caracterización de tal extensión.

Teorema 2.2. *Sean C_1/k y C_2/k dos curvas y $\phi : C_1 \to C_2$ un morfismo de curvas no constante. Entonces $k(C_1)$ es una extensión finita de $\phi^* k(C_2)$.*

Demostración. Usando ϕ^* podemos pensar que se tiene la inclusión de campos $k(C_2) \subseteq k(C_1)$. Ambos campos son extensiones finitamente generadas y de grado de trascendencia 1 sobre k. De la torre de campos siguiente

$$\begin{array}{c} k(C_1) \\ | \\ k(C_2) \\ | \\ k \end{array}$$

se sigue que $\operatorname{gr}\operatorname{tr}(k(C_1)/k) = \operatorname{gr}\operatorname{tr}(k(C_2)/k) + \operatorname{gr}\operatorname{tr}(k(C_1)/k(C_2))$ y, ya que $\operatorname{gr}\operatorname{tr}(k(C_1)/k) = \operatorname{gr}\operatorname{tr}(k(C_2)/k) = 1$, se tiene que $\operatorname{gr}\operatorname{tr}(k(C_1)/k(C_2)) = 0$. Por lo tanto la extensión es algebraica y, puesto que las extensiones son finitamente generadas sobre k, se tiene que $k(C_1)/k(C_2)$ es finita. □

Con base en los Teoremas 2.1 y 2.2 se tiene la definición siguiente.

Definición 2.3. *Sea $\phi : C_1 \to C_2$ un morfismo de curvas definidas sobre k. Si ϕ es constante se define el grado de ϕ, que denotaremos por* $\operatorname{gr}(\phi)$*, igual a 0. Si ϕ no es constante diremos que ϕ es finito y se define el grado de ϕ como:*

$$\operatorname{gr}(\phi) = [k(C_1) : \phi^* k(C_2)]$$

Decimos que ϕ es separable si $k(C_1)/\phi^* k(C_2)$ es separable, y el grado de separabilidad de ϕ, denotado por $\operatorname{gr}_s \phi$*, es el grado de separabilidad de $k(C_1)/\phi^* k(C_2)$.*

Definición 2.4. *Sea C una curva. El espacio de formas diferenciales (meromorfas) en C, denotado por Ω_C, es el $\overline{k}(C)$-espacio vectorial generado por los símbolos de la forma dx para $x \in \overline{k}(C)$, donde $\overline{k}$ es una cerradura algebraica de k, que satisfacen las relaciones siguientes:*

(1) $d(x+y) = dx + dy$ para todo $x, y \in \overline{k}(C)$.

(2) $d(xy) = xdy + ydx$ para todo $x, y \in \overline{k}(C)$.

(3) $da = 0$ para cada $a \in \overline{k}$.

Sean C_1 y C_2 dos curvas y $\varphi : C_1 \to C_2$ un morfismo no constante. Entonces la función natural $\varphi^* : \overline{k}(C_2) \to \overline{k}(C_1)$ induce una función $\varphi^* : \Omega_{C_2} \to \Omega_{C_1}$ dada por:

$$\varphi^*\Big(\sum f_i\, dx_i\Big) = \sum (\varphi^* f_i)\, d(\varphi^* x_i). \qquad (*)$$

El teorema siguiente, cuya demostración puede verse en [3] capítulo II proposición 4.2, enuncia algunas propiedades del espacio de formas diferenciales sobre una curva C.

Teorema 2.5. *Sea C una curva. Entonces:*

(1) Ω_C es un $\overline{k}(C)$-espacio vectorial de dimensión 1.

(2) Sea $x \in \overline{k}(C)$. Entonces $dx \in \overline{k}(C)$ es una base para Ω_C si y sólo si $\overline{k}(C)/\overline{k}(x)$ es una extensión finita separable. Aquí $\overline{k}(x)$ denota el campo de funciones racionales en x.

Proposición 2.6. *Sean C_1, C_2 curvas y $\phi : C_1 \to C_2$ un morfismo no constante. Entonces ϕ es separable si y sólo si la función $\varphi^* : \Omega_{C_2} \to \Omega_{C_1}$ es inyectiva.*

Demostración. Por el teorema 2.5 se puede escoger un $y \in \overline{k}(C_2)$ tal que $\{dy\}$ es una base del $\overline{k}(C_2)$-espacio vectorial Ω_{C_2} y $\overline{k}(C_2)/\overline{k}(y)$ es separable. Ahora $\phi^* : \overline{k}(C_2) \to \overline{k}(C_1)$ es un $\overline{k}$-monomorfismo, por lo que $\phi^*\overline{k}(C_2)$ es separable sobre $\phi^*\overline{k}(y) = \overline{k}(\phi^* y)$, esta última igualdad por que ϕ^* es un $\overline{k}$-monomorfismo.

Se afirma que $\phi^* : \Omega_{C_2} \to \Omega_{C_1}$ es inyectiva si y sólo si $d(\phi^* y) \neq 0$. Para ver esto supongamos que ϕ^* es inyectiva; por $(*)$ se tiene que $\phi^*(dy) = d(\phi^* y)$ y puesto que $dy \neq 0$, ya que ϕ^* inyectiva, se tiene que $d(\phi^* y) \neq 0$. Recíprocamente supongamos que $d(\phi^* y) \neq 0$ y sea $f \in \Omega_{C_2}$ tal que $\phi^*(f) = 0$. Entonces $f = \eta dy$, con $\eta \in \overline{k}(C_2)$, así $\phi^*(f) = \phi^*(\eta) d(\phi^* y)$, por $(*)$. Como ϕ^* es un monomorfismo de campos se sigue que $\eta = 0$ por lo que $f = 0$. Con esto la afirmación queda completamente probada.

Ahora, $d(\phi^* y) \neq 0$ si y sólo si $d(\phi^* y)$ es una base para Ω_{C_1}, por el teorema 2.5 (1), y $d(\phi^* y)$ es base de Ω_{C_1} si y sólo si $\overline{k}(C_1)/\overline{k}(\phi^* y)$ es separable, por el teorema 2.5 (2).

Puesto que $\overline{k}(\phi^* y) \subseteq \phi^*\overline{k}(C_2) \subseteq \overline{k}(C_1)$ y $\phi^*\overline{k}(C_2)/\overline{k}(\phi^* y)$ es separable, entonces $\overline{k}(C_1)/\overline{k}(\phi^* y)$ es separable si y sólo si $\overline{k}(C_1)/\phi^*\overline{k}(C_2)$ es separable. Por lo tanto $\overline{k}(C_1)/\phi^*\overline{k}(C_2)$ es separable si y sólo si ϕ^* es inyectiva. □

3. Curvas elípticas

Una *curva elíptica* sobre un campo k es un par (E, O), donde E es una curva de género 1 definida sobre k, y $O \in E$ es un punto distinguido de E y con coordenadas en k.

Ejemplo 3.1.

Sea K un campo, de característica distinta de 2. La curva

$$y^2 = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$$

donde los α_i son distintos entre sí, es una curva elíptica, ya que por [3] capítulo II ejemplo 5.7, tal curva tiene género 1 (ver la figura siguiente de una curva elíptica en $\mathbb{C}$, aunque solo veamos la parte real de tal curva, aquí $\alpha_1 = -1$, $\alpha_2 = 0$ y $\alpha_3 = 1$.)
Una curva elíptica E se puede identificar con una curva dada por una ecuación de Weierstrass

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

con coeficientes en K, ver [3] capítulo III proposición 3.1. Una diferencial importante asociada a tal curva elíptica

$$\omega = \frac{dx}{2y + a_1x + a_3} = \frac{dy}{3x^2 + 2a_2x + a_4 - a_1y}$$

se le llama *la diferencial invariante* asociada a la ecuación de Weierstrass E.

La estructura de grupo. Sea E una curva elíptica con un punto base O, dos puntos $P, Q \in E$ y L la recta que une a P y Q si $P \neq Q$ y si $P = Q$, entonces L es la recta tangente a E en P. Sea R el tercer punto de intersección de L con E, que existe por el teorema de Bezout. Sea $\tilde{L}$ la recta que une a O con R. Entonces $R' = P \oplus Q$ es el tercer punto dado por la intersección de $\tilde{L}$ con E, que existe gracias al teorema de Bezout (ver la figura (1)). Se demuestra que, bajo esta operación, E es un grupo abeliano, ver, por ejemplo, [3] capítulo III proposición 2.2.

Definición 3.2. *Sean E_1 y E_2 curvas elípticas. Una isogenia entre E_1 y E_2 es un morfimo de curvas $\phi : E_1 \to E_2$ tal que $\phi(O_1) = O_2$, donde O_1 y O_2 son los puntos distinguidos de E_1 y E_2, respectivamente. Diremos que E_1 y E_2 son isógenas si existe una isogenia $\phi : E_1 \to E_2$ tal que $\phi(E_1) \neq O$, y así, por el teorema 2.1, ϕ es suprayectiva.*

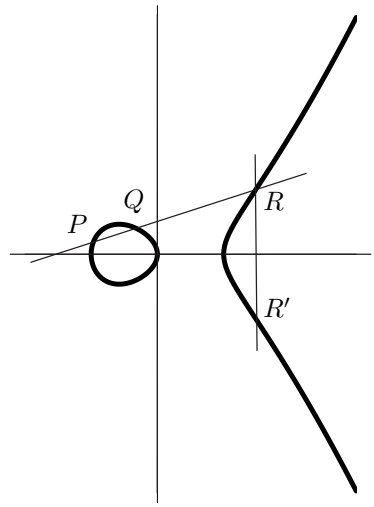

Figura 1: La operación de grupo en una curva elíptica

Ejemplo 3.3.

Sea $m \in \mathbb{Z}$ y E una curva elíptica. Se define la isogenia *multiplicación* por m, $[m] : E \to E$ de la manera siguiente:

$$[m](P) = P + \ldots + P \,(m\text{-veces}).$$

Ejemplo 3.4.

El morfismo de Frobenius. Sea K un campo de característica $p > 0$ y $q = p^r$, r un número natural. Si E/K es una curva elíptica dada por una ecuación de Weierstrass C, $E^{(q)}/K$ es la curva definida por el ideal $I(E^{(q)})$ generado por

$$\{f^{(q)} \mid f \in I(C)\}$$

donde $f(q)$ es el polinomio cuyos coeficientes son los coeficientes de f elevados a la potencia q-ésima y el *morfismo de Frobenius*, $\phi_q : E \to E^{(q)}$, es el dado por

$$(x, y) \mapsto (x^q, y^q)$$

Aquí $E^{(q)}$ está dada por una ecuación de Weierstrass y se puede probar, ver [3] capítulo III ejemplo 4.6, que $E^{(q)}$ es una curva elíptica. En particular si $K = \mathbb{F}_q$, entonces el morfismo potencia q-ésima es la identidad por lo que $E^{(q)} = E$ y ϕ_q es un endomorfismo de E, llamado

el *endomorfismo de Frobenius*. Nótese que los puntos fijos del endomorfismo de Frobenius son los puntos de E con coordenadas en K, es decir, el núcleo del morfismo $1 - \phi_q : E \to E$ consiste de tales puntos y sólo ellos.
Puede mostrarse, ver por ejemplo [3] capítulo III teorema 4.8, que si E es una curva elíptica y $\phi : E \to E$ es una isogenia, entonces ϕ define un morfismo (endomorfismo) del grupo abeliano E en sí mismo. El teorema siguiente nos da información sobre la cardinalidad de las fibras de ϕ.

Teorema 3.5. *Sean E_1 y E_2 curvas elípticas y $\phi : E_1 \to E_2$ una isogenia entre ellas, no constante. Entonces para todo $Q \in E_2$, excepto un número finito, se tiene que $N(\phi^{-1}(Q)) = \mathrm{gr}_s\, \phi$.*

Una demostración del teorema anterior puede encontrarse en [3] capítulo II proposición 2.6.

Proposición 3.6. *Sean E_1 y E_2 curvas elípticas y $\phi : E_1 \to E_2$ una isogenia entre ellas no constante. Entonces si ϕ es separable se tiene que:*

$$N(\ker(\phi)) = \mathrm{gr}\, \phi.$$

Demostración. Por el teorema 3.5 sabemos que $N(\phi^{-1}(Q)) = \mathrm{gr}_s\, \phi$ excepto para un número finito de puntos Q en E_2, además como ϕ es separable se tiene que $\mathrm{gr}\, \phi = \mathrm{gr}_s\, \phi$. Si P y P_1 son dos puntos de E_2, puesto que ϕ no es constante, por el teorema 2.1 existe un $R \in E_1$ tal que $\phi(R) = P_1 - P$. De esta manera definimos una función $\psi : \phi^{-1}(P) \to \phi^{-1}(P_1)$ como $\psi(S) = S + R$. Obsérvese que $\phi(S) = P$ y $\phi(R) = P_1 - P$. De este modo $\phi(\psi(S)) = \phi(S) + \phi(R) = P$, es decir, ψ está bien definida.

Si $S_1, S_2 \in E_1$ cumplen que $\psi(S_1) = \psi(S_2)$, se tiene que $R + S_1 = R + S_2$ por lo que $S_1 = S_2$ y así ψ es inyectiva. Sea $P_3 \in \phi^{-1}(P_1)$; notemos que $P_3 - R \in \phi^{-1}(P)$ y que $\psi(P_3 - R) = P_3$ por lo que ψ es suprayectiva. Por lo tanto para todo $Q \in E_2$ se tiene que $N(\phi^{-1}(Q)) = \mathrm{gr}_s\, \phi$, en particular $\ker(\phi) = \phi^{-1}(0)$ y con esto termina la demostración. □

Para la demostración del teorema siguiente consultar [3] capítulo III teorema 5.2.

Teorema 3.7. *Sean E y E' curvas elípticas, ω una diferencial invariante de E, y $\phi, \psi : E' \to E$ dos isogenias. Entonces*

$$(\phi + \psi)^*(\omega) = \phi^*(\omega) + \psi^*(\omega)$$

Corolario 3.8. *Sea ω una diferencial invariante de una curva elíptica E y sea m un entero. Entonces*

$$[m]^*(\omega) = m\omega$$

Demostración. El corolario es cierto para $m = 0$ y $m = 1$. Si en el teorema 3.7 ponemos $\phi = [m]$ y $\psi = [1]$ se obtiene que

$$[m+1]^*\omega = [m]^*\omega + \omega$$

así, por inducción, obtenemos el resultado deseado. □

Proposición 3.9. *Sean E una curva elíptica definida sobre $\mathbb{F}_q$, $q = p^l$, $\phi : E \to E$ el morfismo de Frobenius y $m, n \in \mathbb{Z}$. Entonces, el morfismo $m+n\phi : E \to E$ es separable si y sólo si $p \nmid m$. En particular $1 - \phi$ es separable.*

Demostración. Sea ω una diferencial invariante en E. Por la proposición 2.6, un morfismo $\psi : E \to E$ es separable si y sólo si $\psi^*(\omega) \neq 0$, es decir si y sólo si $\psi^* : \Omega_E \to \Omega_E$ es inyectiva. Considere el morfismo $\psi = m + n\phi$. Usando el teorema 3.7 y el corolario 3.8, se obtiene que:

$$(m + n\phi)^*(\omega) = m^*(\omega) + (n\phi)^*(\omega) = m\omega + n\phi^*\omega$$

Pero $\phi^*(\omega) = 0$, puesto que $\phi^* dx = d(x^q) = qdx^{q-1} = 0$. Entonces;

$$(m + n\phi)^*(\omega) = m\omega$$

Ahora $m\omega = 0$ si y sólo si $p \mid m$, ya que por el teorema 2.5 se tiene que Ω_E es un $\overline{\mathbb{F}_q}(E)$- espacio vectorial, de dimensión 1, además $\overline{\mathbb{F}_q}(E)$ tiene característica q y, ciertamente, $\omega \neq 0$ por lo que ω es una base de Ω_E, de aquí se sigue nuestra afirmación ya que $(m + n\phi)^*(\omega) \neq 0$ si y sólo si $p \nmid m$. □

4. Las conjeturas de Weil

Probaremos ahora las conjeturas de Weil, en el caso de una curva elíptica. Para empezar sea ℓ un número primo distinto de $\mathrm{car}(K)$, donde K es un campo finito. Se recuerda que se tiene una representación $\mathrm{End}(E) \to \mathrm{End}(T_\ell(E))$ (donde T_ℓ es el *módulo de Tate* de E) dada por

$$\psi \mapsto \psi_\ell$$

ver [3] capítulo III sección 7 y, en esta misma sección proposición 7.1, se tiene que $T_\ell(E) \cong \mathbb{Z}_\ell \times \mathbb{Z}_\ell$ como $\mathbb{Z}_\ell$-módulo. Así es posible escribir a ψ_ℓ como una matriz 2×2, con coeficientes en $\mathbb{Z}_\ell$ y se tiene que

$$\det(\psi_\ell), \mathrm{Tr}\,(\psi_\ell) \in \mathbb{Z}_\ell.$$

En la proposición siguiente aparece el *módulo de Tate* de K, $T_\ell(\mu)$, cuya definición puede verse en [3], p.91.

Proposición 4.1. *Si $\psi \in \mathrm{End}(E)$, entonces*

$$\det(\psi_\ell) = \mathrm{gr}(\psi) \text{ y } \mathrm{Tr}\,(\psi_\ell) = 1 + \mathrm{gr}(\varphi) - \mathrm{gr}(1-\varphi)$$

En particular, $\det(\psi_\ell)$ *y* $\mathrm{Tr}\,(\psi_\ell)$ *son enteros e independientes de* ℓ*.*

Demostración. Sea v_1, v_2 una $\mathbb{Z}_\ell$-base para $T_l(E)$, y escribamos la matriz de ψ_ℓ para esta base

$$\psi_\ell = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Usando el apareamiento de Weil, alternante, no degenerado, (ver [3] capítulo III proposición 8.3)

$$e : T_\ell(E) \times T_\ell(E) \to T_\ell(\mu)$$

se tiene que

$$\begin{aligned} e(v_1, v_2)^{\mathrm{gr}(\psi)} &= e((\mathrm{gr}(\psi))v_1, v_2) \\ &= e(\widehat{\psi}_l \psi_l v_1, v_2) \qquad (1) \\ &= e(\psi_\ell v_1, \psi_\ell v_2) \qquad (2) \\ &= e(av_1 + cv_2, bv_1 + dv_2) \\ &= e(v_1, v_2)^{ad-bc} \\ &= e(v_1, v_2)^{\det(\psi_\ell)} \end{aligned}$$

la igualdad (1) es por [3] capítulo III proposición 6.1(a) y la igualdad (2) es por [3] capítulo III proposición 8.3 y proposición 6.2(f). Puesto que e es no degenerado, se sigue que $\mathrm{gr}(\psi) = \det(\psi_\ell)$. Finalmente para una matriz A de tamaño 2×2, se tiene que:

$$\mathrm{Tr}\,(A) = 1 + \det(A) - \det(1 - A).$$

y de aquí se sigue la última afirmación.[1] □

[1] donde 1 en $\det(1 - A)$ denota la matriz identidad.

Considere ahora el morfismo de Frobenius $\phi : E \to E$. Por el ejemplo 3.4 y las proposiciones 3.6 y 3.9

$$N(E(K)) = \mathrm{gr}(1 - \phi).$$

De modo análogo, para cada $n \geq 1$ tenemos que:

$$N(E(K_n)) = \mathrm{gr}(1 - \phi^n).$$

De la proposición 4.1, el polinomio característico de ϕ_ℓ tiene coeficientes enteros, y este polinomio se puede factorizar sobre los complejos, es decir:

$$\det(T - \phi_\ell) = T^2 - \mathrm{Tr}\,(\phi_\ell)T + \det(\phi_\ell) = (T - \alpha)(T - \beta)$$

Además, puesto que para cada número racional $\frac{m}{n}$ se tiene que:

$$\det\left(\frac{m}{n} - \phi_\ell\right) = \frac{\det(m - n\phi_\ell)}{n^2} = \frac{\mathrm{gr}(m - n\phi)}{n^2} \geq 0$$

se sigue que el polinomio cuadrático $\det(T - \phi_\ell)$ tiene una raíz doble o raíces complejas conjugadas. De esto se sigue que $|\alpha| = |\beta|$, y ya que

$$\alpha\beta = \det(\phi_\ell) = gr(\phi) = q$$

se concluye que $|\alpha| = |\beta| = \sqrt{q}$ y así el polinomio característico de ϕ_ℓ^n esta dado por:

$$\det(T - \phi_\ell^n) = (T - \alpha^n)(T - \beta^n).$$

De aquí se sigue que:

$$\begin{aligned} N(E(K_n)) &= \mathrm{gr}(1 - \phi^n) \\ &= \det(1 - \phi_\ell^n) \\ &= 1 - \alpha^n - \beta^n + q^n. \end{aligned}$$

El resultado principal es:

Teorema 4.2 (Conjeturas de Weil para curvas elípticas)**.** *Sea K un campo con q elementos E/K una curva elíptica. Entonces existe un número entero a tal que*

$$Z_{E/K}(T) = \frac{1 - aT + qT^2}{(1 - T)(1 - qT)}.$$

Además

$$Z_{E/K}(\frac{1}{q}T) = Z_{E/K}(T)$$

y

$$1 - aT + qT^2 = (1 - \alpha T)(1 - \beta T)$$

con $\mid \alpha \mid = \mid \beta \mid = \sqrt{q}$.

Demostración. Calculamos la serie:

$$\begin{aligned}\log Z_{E/K}(T) &= \sum N(E(K_n))T^n/n \\ &= \sum (1 - \alpha^n - \beta^n + q^n)T^n/n \\ &= -\log(1-T) + \log(1-\alpha T) + \log(1-\beta T) \\ &\quad - \log(1-qT).\end{aligned}$$

Por lo tanto

$$Z_{E/K}(T) = \frac{(1-\alpha T)(1-\beta T)}{(1-T)(1-qT)}.$$

Así la función $Z_{E/K}$ tiene la forma deseada, ya que por las observaciones anteriores al teorema, vemos que α y β son complejos conjugados, con valor absoluto $\sqrt{q}$, además

$$a = \alpha + \beta = \mathrm{Tr}\,(\phi_\ell) = 1 + q - \mathrm{gr}(1 - \phi) \in \mathbb{Z}.$$

□

Notemos que $\mathrm{gr}(1-\phi) = N(E(K))$, es decir, para obtener a hay que contar el número de soluciones de la curva elíptica sobre el campo K.

Observaciones finales. Para el caso general de una variedad proyectiva lisa sobre un campo finito, las conjeturas de Weil fueron demostradas por los matemáticos siguientes:

- La racionalidad de la función zeta fue demostrada por B. Dwork en 1960 usando métodos de análisis p-ádico. Una demostración más acorde con lo conjeturado por Weil es la de Grothendieck de 1965 y que usa la cohomología étale de un esquema. Así los fundamentos para la geometría algebraica, propuestos por Grothendieck, son un marco de referencia más adecuado para establecer

y probar tales conjeturas, pues permitió generalizar métodos de la topología algebraica en la geometría algebraica, por ejemplo una generalización del teorema de punto fijo de Lefschetz, usando cohomología ℓ-ádica, a pesar de que en una variedad abstracta no se tenga una topología fina, como en el caso complejo. Para el caso de una curva de género arbitrario, generalizando el ejemplo de una curva elíptica, puede verse la tesis de maestría del autor [2].

- La existencia de la ecuación funcional, para cuya demostración se requirió la existencia de un teorema de dualidad de Poincaré para variedades algebraicas no singulares, en esta parte se usa no sólo la cohomología étale sino también la cohomología ℓ-ádica

- La parte más difícil de las conjeturas de Weil, el análogo de la hipótesis de Riemann, fue demostrado para una curva arbitraria por Weil y el caso general, fue demostrado por Pierre Deligne en 1974 completando el programa iniciado por Grothendieck.

Referencias

[1] Hartshorne, R., *Algebraic Geometry*. Springer-Verlag, New York, 1977.

[2] Sánchez Mirafuentes, M., *Racionalidad de la función zeta de una curva*. Tesis de Maestría en Ciencias. Universidad Autónoma Metropolitana-I., México, D. F., 2009.

[3] Silverman J. H., *The Arithmetic of Elliptic Curves*. Springer-Verlag, New York, 1986.

Dirección del autor
Marco Antonio Sánchez Mirafuentes
Universidad Autónoma Metropolitana,
Unidad Iztapalapa,
División de Ciencias Básicas e Ingeniería,
Departamento de Matemáticas.
Av. San Rafael Atlixco 186, Col. Vicentina
Del. Iztapalapa, C.P. 09340 México, D.F.

UAM-Iztapalapa

Posgrados :
Maestría y Doctorado en Matemáticas

pmat@xanum.uam.mx

http://pmat.izt.uam.mx/

LÍNEAS DE INVESTIGACIÓN

Teoría de anillos y módulos.
Teoría de conjuntos y lógica.
Geometría algebraica.
Geometría diferencial y Riemanniana.
Teoría de números.
Teoría de códigos y criptografía.
Análisis geométrico.
Física matemática.
Análisis diferencial.
Matemáticas discretas, combinatoria y gráficas.
Dinámica de fluidos computacional.
Resolución numérica de ecuaciones en derivadas parciales.
Métodos matemáticos en finanzas y economía.
Control y sistemas dinámicos.
Mecánica celeste, sistemas hamiltonianos y aplicaciones a la física.
Persistencia y bifurcación en sistemas dinámicos.
Control, estabilidad y robustez de sistemas estocásticos.
Metodología estadística.
Estadística asintótica.
Topología de conjuntos, grupos topológicos y Cp-teoría.
Métodos geométricos en mecanica. Dinámica de vórtices. Mécanica celeste

Maestría en Ciencias Matemáticas Aplicadas e Industriales (MACMAI)

info@mcmai.info

http://www.mcmai.info/

LÍNEAS DE INVESTIGACIÓN

Códigos y Criptografía
Control y Sistemas Dinámicos
Combinatoria y Optimización
Estadística
Métodos Matemáticos en Finanzas
Modelación y Simulación Computacional

UNIVERSIDAD AUTÓNOMA METROPOLITANA, UNIDAD IZTAPALAPA
DEPARTAMENTO DE MATEMÁTICAS, EDIFICIO ANO DEL T, 3ER PISO, TEL. 58044656 ext 307 FAX. 5804 4660
AV. SAN RAFAEL ATLIXCO No 186, COL. VICENTINA, IZTAPALAPA, MEXICO, D.F.

Contenido