UNIVERSIDAD NACIONAL DE CÓRDOBA

FACULTAD DE MATEMÁTICA, ASTRONOMÍA Y FÍSICA

SERIE "B"

TRABAJOS DE MATEMÁTICA

N° 64/2014

VII Encuentro Nacional de Álgebra

Notas de Cursos

4 al 8 de Agosto 2014, Córdoba, Argentina.

Daniel Jaume – Vanesa Meinardi – Leandro Vendramin
Martín Mereb – Nicolás Sirolli

Editores: Jorge G. Adrover

CIUDAD UNIVERSITARIA – (5000) CÓRDOBA
REPÚBLICA ARGENTINA

Prefacio

Los Encuentros Nacionales de Álgebra vienen realizándose en las Sierras de Córdoba, periódicamente y con gran éxito, desde 2003 cuando tuvo lugar el primero de ellos. El segundo Encuentro elENA II se realizó en 2004 y a partir de éste se hicieron en forma bianual: elENA III (2006), elENA IV (2008), elENA V (2010), elENA VI (2012).

El *Séptimo Encuentro Nacional de Álgebra*, elENA VII, se llevará a cabo desde el 4 al 8 de Agosto de 2014, en el Hotel del Lago, situado en la localidad de La Falda, Sierras de Córdoba. Como es habitual, elENA VII contará con la presencia de numerosos matemáticos del país y también del extranjero. En esta ocasión el encuentro retoma su formato de cinco días, el cual permite una mayor interacción entre los investigadores formados y los alumnos de licenciatura que buscan en el álgebra y temas afines un probable tema de investigación para un posible doctorado, como así también la posibilidad para los estudiantes de doctorado y posdoctorado de absorber conocimientos de temas básicos y también avanzados de la matemática actual.

En este volumen ponemos a disposición de los asistentes al encuentro, cinco de los nueve cursos que se dictarán en el elENA VII. Hemos priorizado en esta oportunidad el hecho de que estas notas estén disponibles una semana antes del encuentro para mayor comodidad y aprovechamiento de los concurrentes. Como es habitual, las notas están disponibles al público en general de manera gratuita en la página web de Publicaciones de la FaMAF (Universidad Nacional de Córdoba, Argentina), como así también en la página web del encuentro. En esta ocasión, por razones ecológicas, las presentes notas no se imprimirán.

En nombre del Comité Organizador agradecemos a todos los cursistas, muy especialmente a aquellos que, de manera desinteresada, se tomaron el enorme esfuerzo de redactar estas notas.

Emilio Lauret
Córdoba, 28 de Julio de 2014

Contenidos

Cursos de Nivel Básico

Cursos de Nivel Intermedio

Curso de Nivel Avanzado

MÉTODOS ALGEBRAICOS EN COMBINATORIA

DANIEL A. JAUME

Índice

1. Método (cota) del Álgebra Lineal

1.1. Villa (Im)Par. El método de la cota lineal consiste en la siguiente idea (muy simple): si uno desea acotar el tamaño de un conjunto, se asocia los elementos del conjunto con los elementos de un espacio vectorial V de dimensión *baja*, y se muestra que los vectores asociados son linealmente independientes, y por lo tanto el conjunto no puede tener más elementos que la dimensión del espacio V.

En honor a Babai y Frankl, pioneros en este campo, empezaremos con el problema de Villa Par.

Villa Par tiene 32 habitantes. Quienes tienen el vicio compulsivo de formar clubs (grandes o pequeños). Pero los clubs no se pueden formar de forma arbitraria. En gobierno municipal ha establecido las siguientes reglas que todo club debe satisfacer:

1. Cada club debe tener un número par de miembros.
2. Cada par de clubes debe tener en común un número par de miembros.
3. No pueden haber dos clubes formados por exactamente las mismas personas.

Los ciudadanos de Villa Par desean formar tantos clubes como puedan, respetando las reglas pues son bueno vecinos. Para que se den una idea del grado de obsesión que tienen, ellos registran inclusive el *club vacío*, que no posee miembros (¡ya que cero es par!). La pregunta que nos hacemos es:

Date: 25 de Julio de 2014.

¿Cuántos clubes se pueden formar en Villa Par?

Podemos hallar una cota inferior rápidamente, si hacemos un par de suposiciones adicionales, pues a mayor número de requisitos, menos elementos los cumplirán. Primera condició todo el mundo en Villa Par está casado. Segunda condición adicional: a fin de evitar la infidelidad (ilusos) los esposos deben pertenecer a los mismos clubes. (observe que con estas reglas adicionales, las reglas 1 y 2 son superfluas).

Con estas nuevas reglas se pueden formar $2^{16} = 65{,}536$ clubes ¿Por qué? Pues cada club se forma al tomar 16 decisiones binarias: la pareja 1 se asocia o no, la pareja 2 se asocia o no, ..., la pareja 16 se asocia o no. El conjunto de decisiones es igual al conjunto de clubes que se pueden formar en Villa Par (Casados y Esposados).

Entonces el número de clubes posibles en Villa Par es de más de 65.536. Lo cual es asombrosamente alto para una pequeã villa de 32 habitantes.

Después de un tiempo la situación se vuelve inmanejable y el municipio decide cambiar las reglas a fin de que no halla demasiados clubes. El intendente cree que no se necesita cambiar demasiado las reglas, de hecho, según él basta cambiar una sola palabra:

1. Cada club debe tener un número **impar** de miembros.
2. Cada par de clubes debe tener en común un número par de miembros.
3. No pueden haber dos clubes formados por exactamente las mismas personas.

Este cambio, motiva que el pueblo cambie su nombre a **Villa Impar**

Un breve análisis hace ver que última regla es redundante (¿Por qué?). Por lo que la ordenanza que regirá la formación de clubes en Villa Impar es:

(a) Cada club debe tener un número **impar** de miembros.
(b) Cada par de clubes debe tener en común un número par de miembros.

Con esta ordenanza es claro que se pueden armar al menos 32 clubes, por ejemplo:

- 32 clubes de 1 persona.
- 32 clubes de 31 personas.
- 31 clubes de 7 personas y uno de 31.

De hecho hay muchas formas de lograr 32 clubes que satisfagan la regla. Pero, curiosamente, es imposible formar 33 o clubes que satisfagan estas reglas.

Y más curiosamente, todas las soluciones conocidas (al menos para el autor) de este problema puramente combinatorio involucran álgebra lineal.

La idea es trabajar en el espacio ${\mathbb{F}_2}^{32}$ de dimensión 32. Supongamos que tengamos $C_1, \ldots, C_m$ clubs. A cada club le asociamos un vector de incidencia v_i con la siguiente regla: la entrada j-ésima del vector v_i es 1 si el ciudadano j es socio del club C_i y cero en todo otro caso (estamos asumiendo que hemos numerado los ciudadanos de Villa Impar del 1 al 32).

Es claro que el producto interno de dos vectores de incidencia da:

$$\langle v_i, v_j \rangle = |C_i \cap C_j|.$$

O si pensamos en la aritmética sobre el cuerpo $\mathbb{F}_2$

$$\langle v_i, v_j \rangle = \begin{cases} 1, & \text{si } i = j, \\ 0, & \text{si } i \neq j. \end{cases}$$

Ahora vamos a probar que estos vectores son linealmente independientes en ${\mathbb{F}_2}^{32}$. Consideremos:

$$\lambda_1 v_1 + \cdots + \lambda_m v_m = 0.$$

donde cada $\lambda_i \in \mathbb{F}_2$. Por supuesto, debemos probar que todos los escalares son nulos. Par ver el valor de λ_i vamos a multiplicar ambos lados por v_i:

$$\langle \lambda_1 v_1 + \cdots + \lambda_i v_i + \cdots + \lambda_m v_m, v_i \rangle = \langle 0, v_i \rangle.$$

Distribuyendo

$$\langle \lambda_1 v_1, v_i \rangle + \cdots \langle \lambda_i v_i, v_i \rangle + \cdots + \langle \lambda_m v_m, v_i \rangle = 0$$

de donde concluimos que $\lambda_i = 0$. Por lo tanto los vectores $v_1, \ldots, v_m$ linealmente independientes. Consecuentemente $m \leq 32$.

Ahora podemos establecer fácilmente el siguiente teorema:

Teorema 1.1 (Teorema de Villa Impar u Oddtown Theorem)**.** *En toda ciudad con n habitantes, no se pueden formar más de n clubes bajo las reglas:*

(a) *Cada club debe tener un número **impar** de miembros.*
(b) *Cada par de clubes debe tener en común un número par de miembros.*

Daremos otra demostración del Teorema de Villa Impar, usando la independencia lineal vía un argumento de rango. Recordemos brevemente algunas de las (des)igualdades de rango más importantes. Sean $\mathbf{A}, \mathbf{B}$ matrices sobre un cuerpo arbitrario $\mathbb{F}$, de tamaños adecuados para cada caso. Entonces:

1. $rk(\mathbf{A}) = rk(\mathbf{A}^T)$.
2. $rk(\mathbf{AB}) = rk(\mathbf{B}) - \dim(N(\mathbf{A}) \cap R(\mathbf{B}))$.
3. $rk(\mathbf{A}) - rk(\mathbf{B}) \leq rk(\mathbf{A} + \mathbf{B}) \leq rk(\mathbf{A}) + rk(\mathbf{B})$.
4. $rk(\mathbf{A} + rk(\mathbf{B}) - n \leq rk(\mathbf{AB}) \leq \text{mín}\{rk(\mathbf{A}, rk(\mathbf{B})\}$.

Tenemos por un lado n habitantes los cuales pueden ser identificados con $[n] := \{1, \ldots, n\}$, el segmento inicial de los primeros n enteros positivos. Por el otro tenemos un conjunto de clubes, el que puede ser identificado en con una familia de subconjuntos de $[n]$.

$$\mathcal{C} := \{C_1, \ldots, C_m\}$$

donde cada $C_i \in 2^{[n]}$.

Construimos, como es habitual en combinatoria, la matriz $\mathbf{M}$ de incidencia habitantes-clubes. Las filas de esta matriz están indizadas por los habitantes, las columnas por los clubes.

$$m_{i,C_j} = \begin{cases} 1, & \text{si } i \in C_j \\ 0, & \text{en todo otro caso.} \end{cases}$$

Por ejemplo tenemos 5 habitantes y los clubes son

$$\mathcal{C} := \{C_1 = \{1, 2, 3\}, C_2 = \{1, 2, 4\}, C_3 = \{5\}\}.$$

Tenemos que la matriz de incidencia es

$$\mathbf{M} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

De nuevo queremos ver que los vectores de incidencia de los clubes, i.e. (*id est*, latín, esto es) los vectores columnas de la matriz de incidencia M, son linealmente independientes sobre $\mathbb{F}_2$.

Una forma de probar esto es mostrando que el rango de la matriz de incidencia es al menos m, el número de columnas de la matriz.

Consideremos la matriz

$$\mathbf{A} = \mathbf{M}^T\mathbf{M},$$

la cual llamaremos matriz de intersección del sistema $\mathcal{C}$ de clubes. El nombre proviene siguiente hecho (haciendo las cuentas en $\mathbb{Q}$ o $\mathbb{R}$ o $\mathbb{C}$):

$$\mathbf{A}_{i,j} = |C_i \cap C_j|.$$

Pero, haciendo las cuentas en $\mathbb{F}_2$ tenemos que $\mathbf{A} = \mathbf{I}_m$, donde $\mathbf{I}_m$ representa la matriz identidad de orden m.

Ahora

$$m = rk(\mathbf{I}_m)) = rk(\mathbf{A}) = rk(\mathbf{M}^T\mathbf{M}) \leq rk(\mathbf{M}) \leq n.$$

1.2. Teorema de Graham-Pollak. Sea $G = (V,\ E)$ un grafo (finito), y sean $V_1, \ldots, V_r$ conjuntos de vértices de G. Denotemos por G_i al subgrafo de G inducido por los vertices en V_i. Si los grafos $G_i, \ldots, G_r$ son lado-disjuntos y entre todos contienen a todos los lados de G, diremos que los grafos $G_i, \ldots, G_r$ forman una **descomposición lado-disjunta** de G.

Un Clique bipartito es un grafo completo bipartito $K_{A,B} = (A \cup B, E)$ con $A \cap B = \emptyset$ y $E = A \times B$.

Sea $bd(K_n)$ el número más pequeño tal que el grafo completo K_n, puede ser descompuesto en $bd(K_n)$ cliques bipartitos lado-disjuntos. Esta cantidad esta bien definida, pues E forma una colección de $\frac{n(n-1)}{2}$ cliques bipartitos lado-disjuntos. Luego $bd(K - n) \leq \frac{n(n-1)}{2}$. Pero obviamente que podemos hacerlo mejor, por ejemplo K_5 puede ser descompuesto en 4 estrellas (grafos donde uno de los vertices está unido al resto, y no hay otras conexiones):

1. $K_{\{1\},\{2,3,4,5\}}$,
2. $K_{\{2\},\{3,4,5\}}$,
3. $K_{\{3\},\{4,5\}}$,
4. $K_{\{4\},\{5\}}$.

Por lo que claramente $f(5) \leq 4$. En general esta partición en cliques bipartitos funciona, y K_n puede ser descompuesto en $n-1$ cliques bipartitos lado-disjuntos: K_{A_i,B_i}, con $A_i = \{i\}$ y $B_i = \{i+1, \ldots, n\}$, para $i = 1, \ldots, n-1$. Esta es una forma de realizar la descomposición en $n-1$ cliques bipartitos lado-disjuntos, pero hay muchas otras.

Si bien ahora es claro que $f(n) \leq n-1$, el siguiente resultado clásico de Graham y Pollak (1971) nos dice que no podemos mejorar más.

Teorema 1.2. *Los lados de K_n no pueden ser descompuestos en menos de $n-1$ cliques bipartitos lado-disjuntos.*

Este es un resultado importante con multiples aplicaciones a la teoría de comunicaciones y computación, por lo que ha sido objeto de multiples demostraciones. Hasta hace muy poco, sólo se conocían demostraciones algebraicas. Recién en 2008 Vishwanathan y 2013 Chung dieron demostraciones puramente combinatorias.

Demostración. **(Trevberg 1982)** El polinomio

$$S(x_1, \ldots, x_n) := \sum_{1 \leq i < j \leq n} x_i x_j$$

nos da una representación algebraica del conjunto de lados de K_n. Y cada descomposición de K_n en cliques bipartitos lado-disjuntos nos da una forma de reescribir a $S(x)$.

Supongamos que los pares $(A_1, B_1), \ldots, (A_t, B_t)$ nos dan una descomposición de K_n es cliques bipartitos lado-disjuntos. Entonces

$$S(x_1, \ldots, x_n) = \sum_{k=1}^{t} \left(\sum_{i \in A_k} x_i \right) \left(\sum_{j \in B_k} x_j \right). \tag{1.1}$$

Ahora nos podemos olvidar del grafo, y simplemente tratar de hallar el mínimo t para el cual 1.1 vale, donde la única restricción es que $A_i \cap B_i = \emptyset$ para toda $i = 1, \ldots, t$, con $A_i, B_i \subset [n]$.

Ahora

$$\left(\sum_{i=1}^{n} x_i \right)^2 = \sum_{i=1}^{n} x_i^2 + 2 \sum_{1 \leq i < j \leq n} x_i x_j.$$

Por lo tanto

$$\sum_{i=1}^{n} x_i^2 = \left(\sum_{i=1}^{n} x_i \right)^2 - 2S(x_1, \ldots, x_n).$$

Supongamos que $t \leq n - 2$, recordando 1.1 tenemos que el sistema homogéneo de a lo más $n - 1$ ecuaciones:

$$\begin{cases} \sum_{i \in A_k} x_i &= 0, \text{ para } k = 1, \ldots, t. \\ \sum_{i=1}^{n} x_i &= 0 \end{cases}$$

tiene una solución no-trivial $\hat{x}_1, \ldots, \hat{x}_n$. Ahora, por un lado como $\sum_{i \in A_k} \hat{x}_i = 0$ y $\sum_{i=1}^{n} \hat{x}_i = 0$, tenemos que

$$\sum_{i=1}^{n} \hat{x}_i^2 = 0.$$

Pero como $\hat{x}_1, \ldots, \hat{x}_n$ es una solución no trivial:

$$\sum_{i=1}^{n} \hat{x}_i^2 > 0,$$

lo cual es ridículo. Luego $t \geq n - 1$. □

1.3. Autovalores y Teorema de Witsenhausen. Ahora abordaremos el problema general de descomponer un grafo finito G en un conjunto lado-disjunto de cliques bipartitos, de ahora en más **bicliques**. El **número de descomposición por bicliques** de un grafo G, denotado por $bd(G)$, es el mínimo número de bicliques lado-disjuntos en que se puede descomponer a G. Observe que este parámetro está bien definido para todo grafo finito. El teorema de Graham-Pollak nos dice que $bd(K_n) = n - 1$.

Sabemos que $bd(K_n) = 4$, y sean grafos B_1, B_2, B_3, B_4 una descomposición por bicliques lado-disjuntos del K_5. Por ejemplo

1. $B_1 = K_{\{1\},\{2,3,4,5\}}$,
2. $B_2 = K_{\{2\},\{3,4,5\}}$,
3. $B_3 = K_{\{3\},\{4,5\}}$,
4. $B_4 = K_{\{4\},\{5\}}$.

Denotemos por u_i y v_i a los vectores característicos de las particiones de B_i. Así:

1. $u_1 = (1, 0, 0, 0, 0)$ y $v_1 = (0, 1, 1, 1, 1)$,
2. $u_2 = (0, 2, 0, 0, 0)$ y $v_2 = (0, 0, 1, 1, 1)$,

3. $u_3 = (0,0,1,0,0)$ y $v_3 = (0,0,0,1,1)$,
4. $u_4 = (0,0,0,1,0)$ y $v_4 = (0,0,0,0,1)$.

Observe que la matriz de adyacencia de B_i como subgrafo de G se puede escribir en términos de los vectores característicos u_i y v_i:

$$\mathbf{D}(B_i) = u_i v_i^T + v_i u_i^T.$$

Por ejemplo:

$$\mathbf{D(B_3)} = u_3 v_3^T + v_3 u_3^T.$$

Pues

$$\begin{aligned} u_3 v_3^T + v_3 u_3^T &= \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} [0,0,0,1,1] + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} [0,0,1,0,0] \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \\ &= \mathbf{D(B_3)}. \end{aligned}$$

Ahora estableceremos un resultado muy general, atribuido a Witsenhausen (en algún momento de los 80s):

Teorema 1.3. *Sea G un grafo finito y A su matriz de adyacencia. Definimos $n_+(A)$ y $n_-(A)$ como el número de autovalores positivos de A y el número de autovalores negativos de A, contando multiplicidades. Entonces*

$$bd(G) \geq \text{máx}\{n_+(A), n_-(A)\}.$$

Demostración. Sean $B_1, \ldots, B_t$ una descomposición por bicliques lado-disjuntos de un grafo finito G. Sean $D(B_i)$ la matriz de adyacencia del grafo B_i visto como subgrafo de G. Entonces es claro que

$$A(G) = \sum_{i=1}^{bd(G)} D(B_i).$$

Consideremos los siguientes subespacios de $\mathbb{R}^n$

$$\begin{aligned} W &= span\,\{w \in \mathbb{R}^n \,:\, w^T u_i = 0, \forall 1 \leq i \leq bd(G)\}, \\ P &= span\,\{\text{ Autovectores asociados a autovalores positivos de } A(G)\}. \end{aligned}$$

Es claro que

$$\dim(W) \geq n - bd(G).$$

Ahora, para todo $0 \neq p \in P$, tenemos que $p^T A(G)p > 0$. Por lo que $W \cap P = \{0\}$. Recordando que al ser $A(G)$ una matriz simétrica podemos conseguir una base (ortonormal) de $\mathbb{R}^n$ con autovectores de $A(G)$, tenemos que

$$\dim(W) \leq n - \dim(P) = n - n_+(A).$$

Entonces

$$n - bd(G) \leq \dim(W) \leq n - n_+(A)$$

de donde podemos concluir que $bd(G) \geq n_+(A)$. *Mutatis mutandis* (latín: cambiando lo que haya que cambiar) se prueba que $bd(G) \geq n_-(A)$. □

1.4. Ejercicios.

1. Intercambie **par** por **impar** en las reglas **(a)** y **(b)**. Bajo estas reglas espejo: todos los clubes deben tener un número par de miembros, y dos clubes deben compartir un número impar de miembros. Probar que no más de n clubes se pueden formar en un pueblo con n habitantes.
2. Sea $\mathbf{J}_n$ la matriz cuadra de orden n de todos 1s, por ejemplo
$$\mathbf{J}_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$
Consideremos la matriz $\mathbf{J}_n - \mathbf{I}_n$, ceros en la diagonal y ceros en el resto de las posiciones, por ejemplo
$$\mathbf{J}_3 - \mathbf{I}_3 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$
Se pide:
   a) Calcular los autovalores de $\mathbf{J}_n - \mathbf{I}_n$.
   b) Calcular el determinante de $\mathbf{J}_n - \mathbf{I}_n$.
   c) Demostrar que el rango de $\mathbf{J}_n - \mathbf{I}_n$ sobre $\mathbb{F}_2$ es n si n es par y $n-1$ si n es impar.
3. Usando las reglas espejo dadas es el ejercicio 1, determinar el número máximo de clubes en un pueblo de n habitantes.
4. Sea $\mathbf{A}_{2n\times 2n}$ una matriz de ceros en la diagonal y ± 1 en el resto de las posiciones. Probar que $\mathbf{A}$ es no singular (sobre $\mathbb{R}$).
5. **Villa Impar Bipartita:** Suponga que hay m clubes rojos $R_1, \cdots, R_m$ y m clubes blancos $B_1, \cdots, B_m$ en un pueblo con n habitantes. Asuma que los clubes satisfacen las siguientes reglas:
   a) $|R_i \cap B_i|$ es impar para toda i.
   b) $|R_i \cap B_j|$ es par para toda $i \neq j$.
   Probar que $m \leq n$.
6. Hallar dos descomposiciones de K_6 y tres de K_7 en cliques bipartitos lado-disjuntos.
7. Deducir el Teorema de Graham-Pollak como corolario del Teorema de Witsenhausen.

2. Densidad Combinatoria y Ortogonalidad

2.1. Densidad Combinatoria. En la teoría computacional del aprendizaje la siguiente noción juega un rol importante:

Definición 2.1. Sea $A \subset \{0,1\}^n$, diremos que A es **(n,k)-denso** si existe un subconjunto de k coordenadas

$$S = \{i_1, \ldots, i_k\}$$

tales que la proyección de A sobre los índices en S contienen todas las posibles 2^k configuraciones (por supuesto, n y k son enteros no negativos con $0 \leq k \leq n$).

Por ejemplo, $A = \{(0,1)\}$ es (2,0)-denso, mientras que

$$B = \{(0,0,0), (0,0,1), (0,1,1), (1,1,0)\}$$

es (3,1)-denso y (3,2)-denso pero no es (3,3)-denso.

Una observación elemental, si A es (n,k)-denso y $A \subset B$, entonces B es (n,k)-denso.

El máximo k para el cual un subconjunto de $\{0,1\}^n$ es (n,k)-denso se llama **dimensión de Vapnik-Chervonenkis** del subconjunto (por brevedad: VC-dimensión). Así la VC-dimensión de B es 2.

El problema combinatorio es: dados n y k hallar un conjunto A (n,k)-denso, con la menor cantidad de vectores posibles.

Dados un vector $v = (v_1, \ldots, v_n)$, su proyección sobre un conjunto de coordenadas $S = \{i_1, \ldots, i_k\}$ es el vector

$$v|_S := (v_{i_1}, \ldots, v_{i_k})$$

. La proyección de un conjunto de vectores $A \subset \{0,1\}^n$ sobre S es el conjunto de vectores

$$A|_S := \{v|_S : v \in A\}.$$

Por lo tanto, A es (n,k)-denso si y sólo si $A|_S = \{0,1\}^k$ para al menos un subconjunto S de k coordenadas.

Es claro que todo subconjunto (n,k)-denso debe contener al menos 2^k vectores.

¿Qué tan grande puede ser un subconjunto de $\{0,1\}^n$ sin llegar a ser (n,k)-denso? Si A es el conjunto de todos los vectores de $\{0,1\}^n$ con menos de k unos, entonces:

- A **no** es(n,k)-denso.
- $H(n,k) := |A| = \sum_{i=0}^{k-1} \binom{n}{i}$

Por ejemplo, los $H(4,3) = \binom{4}{0} + \binom{4}{1} + \binom{4}{2} = 11$ vectores que siguen NO son (4,3)-densos

$$\begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array} \qquad \begin{array}{cccc} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{array} .$$

Pero basta que agreguemos uno más al conjunto, por ejemplo el

$$1 \quad 0 \quad 1 \quad 1 \; .$$

Para que el conjunto se vuelva (4,3)-denso, siendo las coordenadas adecuadas (en este caso) $S = \{1,3,4\}$.

Que todo subconjunto de $\{0,1\}^n$ con más de $H(n,k)$ elementos es (n,k)-denso es un hecho que ha sido redescubierto varias veces (en probabilidad, en teoría computacional del aprendizaje, etc), pues es de suma utilidad en muchos campos: lógica, teoría de conjuntos, probabilidad, etc.

Teorema 2.2. *Si $A \subset \{0,1\}^n$ y $|A| > H(n,k)$, entonces A es (n,k)-denso.*

Demostración. Por inducción sobre n y k. Si $k = 1$ entonces A tiene al menos dos vectores diferentes y por lo tanto es (n,1)-denso (hay al menos una coordenada en la que difieren). Ahora tomemos un subconjunto arbitrario $A \subset \{0,1\}^n$ de tamaño $|A| > H(n,k)$. Sea

$$B := A|_{\{1,\ldots,n-1\}}$$

la proyección de A sobre las primeras $n-1$ coordenadas. Sea C el conjunto de todos los vectores u de $\{0,1\}^{n-1}$ para los cuales **ambos** vectores $(u,0)$ y $(u,1)$ pertenecen a A. La siguiente observación es simple pero crucial:

$$|A| = |B| + |C|.$$

Ahora, si $|B| > H(n-1,k)$, entonces por hipótesis inductiva B es (n-1,k)-denso, y por lo tanto el conjunto A es (n,k)-denso (basta tomar las coordenadas que hacen que B sea (n-1,k)-denso).

Si $|B| \leq H(n-1,k)$ entonces, usando la identidad de Pascal,

$$\begin{aligned} |C| &= |A| - |B| > H(n,k) - H(n-1,k) \\ &= \sum_{i=0}^{k-1} \binom{n}{i} - \sum_{i=0}^{k-2} \binom{n-1}{i} \\ &= \sum_{i=0}^{k-2} \binom{n-1}{i} \\ &= H(n-1,k-1). \end{aligned}$$

Luego, por hipótesis inductiva el conjunto C es (n-1,k-1)-denso, y como $C \times \{0,1\} \subset A$, entonces el propio A es (n,k)-denso. □

En 1983, en forma independiente, Alon y Frankl, hicieron una observación a partir de la cual se puede deducir fácilmente el teorema que acabamos de ver: podemos restringir nuestra atención a subconjuntos de $\{0,1\}^n$ con una estructura muy particular, los conjuntos *cerrados hacia abajo o hereditarios.*

2.2. Conjuntos Hereditarios. Estamos interesados en trabajar con subconjuntos de $\{0,1\}^n$ que tengan la propiedad de ser cerrados al cambiar 1s por 0s en sus vectores. Por ejemplo si $(0,1,0,1,1) \in A$ entonces también están en A

$$\begin{matrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{matrix} .$$

Usando el orden habitual para vectores de $\{0,1\}^n$: $u \leq v$ si $u_i \leq v_i$ para todo $i \in \{1,\ldots,n\}$. Podemos dar la siguiente definición:

Definición 2.3. Un conjunto $A \subset \{0,1\}^n$ es **hereditario** o **cerrado hacia abajo** si $v \in A$ y $u \leq v$ implica que $u \in A$.

Dado un conjunto de coordenadas $S \subset \{1,\ldots,n\}$, definimos $t_S(A)$ como el número de vectores en la proyección $A|_S$.

Por ejemplo si A es el conjunto formado por los vectores

$$\begin{matrix} 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{matrix}$$

y $S = \{1, 2, 4\}$, tenemos que $A|_S$ está formado por

$$\begin{matrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{matrix}$$

por lo que $t_S(A) = 3$.

Llamaremos i-ésimo vecino de v al vector $v_{i\to 0}$ que se obtiene de v cambiando la i-ésima coordenada (o bit) por 0 (cero). Por ejemplo, si $v = (0, 1, 0, 1, 1)$, tenemos que $v_{2\to 0} = (0, 0, 0, 1, 1)$ y $v_{3\to 0} = (0, 1, 0, 1, 1) = v$.

El siguiente teorema dice que para cualquier subconjunto A del n-cubo $\{0, 1\}^n$, existe un conjunto hereditario de la misma cardinalidad cuya sombra está a la izquierda (en algún sentido) de la sombra de A.

Teorema 2.4. *Para cada conjunto A del n-cubo $\{0, 1\}^n$ existe un conjunto $B \subset \{0, 1\}^n$ tal que:*

- $|B| = |A|$,
- *Para todo conjunto de coordenadas S tenemos que $t_S(B) \leq t_S(A)$.*

Por ejemplo tomemos el conjunto A

$$\begin{matrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{matrix} .$$

Este conjunto no es hereditario, pues el 3-vecino de $v = (1, 1, 1, 1, 1)$ es el vector $(1, 1, 0, 1, 1) \notin A$. Este vector v de A tiene una coordenada *mala*:$v \in A$ pero $v_{3\to 0} \notin A$. Por otro lado, el conjunto B formado por los vectores

$$\begin{matrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{matrix}$$

si es hereditario y satisface todas las condiciones del teorema.

¿Cómo podemos obtener un subconjunto B *bueno* para cualquier subconjunto A con vectores que posean coordenadas malas?

Sustituyendo los vectores con coordenadas malas por sus i-vecinos. Para lo cual definimos la siguiente transformación. Dado $i \in \{1, \ldots, n\}$ y un subconjunto $A \in \{\{0, 1\}^n$, definimos T_i como sigue: tome un vector $v \in A$ tal que $v_i = 1$, y vea si $v_{i\to 0}$ pertenece a A. Si es así, no haga nada. De lo contrario reemplace v en A por $v_{i\to 0}$.

Por ejemplo $T_5(A)$ solo cambia el primer vector de A:

$$\begin{matrix} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{matrix} .$$

Es claro que T es inyectiva: $|T_i(A)| = |A|$.

Un poco más sutil es ver que para todo subconjunto de coordenadas $S \in \{1, \ldots, n\}$ se tiene que $t_S(T_i(A)) \leq t_S(A)$ (Ejercicio).

Si A es hereditario $T_i(A) = A$ para todo $i \in \{1, \ldots, n\}$.

Demostración. $T_n(T_{n-1}(\cdots T_1(A)\cdots))$ es conjunto buscado. Notar que si A es hereditario, $T_n(T_{n-1}(\cdots T_1(A)\cdots)) = A$. □

Este teorema implica inmediatamente al teorema 2.2 (ejercicio).

El siguiente resultado de Kleitman (1966) sobre la intersección de conjuntos hereditarios tiene numerosas aplicaciones y generalizaciones.

Teorema 2.5. *Sean A, B dos subconjuntos hereditarios de $\{0,1\}^n$. Entonces*

$$|A \cap B| \geq \frac{|A||B|}{2^n}.$$

Demostración. Aplicaremos inducción sobre n. Los casos $n = 0$ y $n = 1$ son triviales. Dado $\omega \in \{0, 1\}$, definimos

$$A_\omega := \{(a_1, \ldots, a_{n-1}) \ : \ (a_1, \ldots, a_{n-1}, \omega) \in A\}$$

y

$$B_\omega := \{(b_1, \ldots, b_{n-1}) \ : \ (b_1, \ldots, b_{n-1}, \omega) \in B\}.$$

Llamaremos $\alpha_\omega := |A_\omega|$ y $\beta_\omega := |B_\omega|$.

Entonces

$$\begin{aligned} |A \cap B| &= |A_0 \cap B_0| + |A_1 \cap B_1| \\ &\geq \frac{\alpha_0\beta_0 + \alpha_1\beta_1}{2^{n-1}} \\ &= \frac{(\alpha_0 + \alpha_1)(\beta_0 + \beta_1)}{2^n} + \frac{(\alpha_0 - \alpha_1)(\beta_0 - \beta_1)}{2^n}. \end{aligned}$$

Ya que ambos conjuntos A y B son hereditarios (o cerrados hacia abajo), tenemos que $A_1 \subset A_0$ y $B_1 \subseteq B_0$, lo que implica que

$$(\alpha_0 - \alpha_1)(\beta_0 - \beta_1) \geq 0.$$

El teorema se sigue del hecho que $|A| = \alpha_0 + \alpha_1$ y $|B| = \beta_0 + \beta_1$. □

2.3. Ortogonalidad. La independencia lineal no es la única forma de obtener cotas superiores (usando algebra lineal). Si los miembros de una familia $\mathcal{F}$ puede ser asociada con los elementos de un espacio vectorial ${\mathbb{F}_q}^m$, entonces $|\mathcal{F}| \leq q^m$. Si somos afortunados, los vectores asociados puede que sean ortogonales a algún subespacio de dimensión d, y como $\dim(U) + \dim(U^\perp) = dim(V)$, podemos mejorar nuestra cota: $|\mathcal{F}| \leq q^{m-d}$.

Dadas dos familias $\mathcal{A}$ y $\mathcal{B}$ de subconjuntos de $[n]$, una pregunta habitual en combinatoria es que tan grande puede llegar a ser $|\mathcal{A}||\mathcal{B}|$. Si no sabemos nada de las familias, este número pude ser tan grande como 2^{2n}. Si sabemos algo, por ejemplo que las familias son monótonas crecientes (o monótonas decrecientes), i.e. cerradas hacia arriba o hacia abajo, el teorema de Kleitman nos da la siguiente cota no trivial:

$$|\mathcal{A}||\mathcal{B}| \leq 2^n |\mathcal{A} \cap \mathcal{B}|.$$

Si además sabemos que la paridad de las intersecciones se mantiene constante, podemos obtener una cota aún mejor.

Teorema 2.6 (Ahlswede-El Gamal-Pang 1984)**.** *Sean $\mathcal{A}$ y $\mathcal{B}$ dos familias de subconjuntos de $[n]$ con la propiedad que $|A \cap B|$ es par para todo $A \in \mathcal{A}$ y $B \in \mathcal{B}$. Entonces $|\mathcal{A}||\mathcal{B}| \leq 2^n$.*

Vamos a dar la demostración de Delsarte-Piret de 1985.

Demostración. A cada subconjunto de $[n]$ le asociamos su vector de incidencia, estos pueden ser vistos como elementos de ${\mathbb{F}_2}^n$. Sean U y V los conjuntos de vectores de incidencias de correspondientes a las familias de subconjuntos $\mathcal{A}$ y $\mathcal{B}$. Observe dado $u \in U$ y $v \in V$, tenemos que

$$\langle u, v \rangle = 0.$$

Luego los espacios $span(U)$ y $span(V)$ son ortogonales, por lo que

$$\dim span(U) + span(V) \leq n.$$

Pues el producto escalar registra la cardinalidad de la intersección de los conjuntos representados por u y v respectivamente. Entonces

$$|\mathcal{A}||\mathcal{B}| = |U||V| \leq |span(U)||span(V)| \leq 2^{\dim(span(U))+\dim(span(V))} \leq 2^n. \tag{2.1}$$

□

2.4. Ejercicios.

1. Sean A y B son dos subconjuntos hereditarios de $\{0,1\}^n$. Cuáles de los siguientes son hereditarios:
   a) $A \cap B$,
   b) $A \cup B$,
   c) $A \triangle B$,
   d) $A \times B$.
2. Sea $A \subset \{0,1\}^n$, definamos $t_s(A) = \max t_S(A)$ sobre todos los $S \subset \{1, \ldots, n\}$ con $|S| = s$.
   a) Si $|A| \leq n$ entonces $t_{n-1}(A) = |A|$.
   b) Si $|A| \leq \lceil \frac{3}{2} n \rceil$ entonces $t_{n-1}(A) \geq |A| - 1$.
3. Probar el teorema de Ahlswede-El Gamal-Pang para intersección impar. Sean $\mathcal{A}$ y $\mathcal{B}$ dos familias de subconjuntos de $[n]$ con la propiedad que $|A \cap B|$ es impar para todo $A \in \mathcal{A}$ y $B \in \mathcal{B}$. Entonces $|\mathcal{A}||\mathcal{B}| \leq 2^n$. ¿Se puede mejorar la cota a 2^{n-1}?

3. Método Polinomial

3.1. Lema de DeMillo-Lipton-Schwartz-Zippel. El método polinomial está basado en diferentes extensiones del teorema factor para polinomios en una variable, al caso de polinomios en varias variables:

1. Todo polinomio (en una variable) de grado d tiene a lo más d raíces.
2. Para todo conjunto $S \subset \mathbb{F}$ existe un polinomio no-nulo f (con coeficientes en $\mathbb{F}$) de grado a lo más $|S|$ tal que $f(x) = 0$ para todo $x \in S$.

Estos resultados nos permiten obtener cotas para el tamaño de un conjunto S dado. Si hallamos un polinomio que se anula sobre S, entonces el grado del polinomio da una cota superior para la cardinalidad de S. Si todas las raíces de un polinomio pertenecen al conjunto S, entonces el grado del polinomio da una cota inferior para la cardinalidad de S.

Repasaremos primero algunos hechos sobre polinomios en cuerpos finitos de una y varias variables.

Dado un anillo conmutativo $\mathbb{R}$, un polinomio en una indeterminada es una expresión de la forma:

$$p(x) = a_0 + a_1 x + \cdots + a_t x^t$$

donde los coeficientes $a_0, a_1, \ldots, a_t \in \mathbb{R}$, y t es un entero no-negativo, el grado del polinomio.

$\mathbb{R}[x]$ denota el conjunto de todos los polinomios con coeficientes en el anillo conmutativo $\mathbb{R}$.

Todo polinomio $p(x)$ con coeficientes en el anillo conmutativo $\mathbb{R}$ define una función de $\mathbb{R}$ sobre $\mathbb{R}$, mapeando $r \in \mathbb{R}$ a $p(r)$ (mapeo evaluación): Así, si $p(x) = 3x + 4x^3$ en $\mathbb{Z}/8\mathbb{Z}$ tenemos que

$$p(2) = 3 \cdot 2 + 4 \cdot 2^3 = 6 + 4 \cdot 0 = 6.$$

En el anillo $\mathbb{Z}/8\mathbb{Z}$.

Pero los polinomios, no son funciones. Esto es debido a la definición de igualdad entre dos polinomios.

Dos polinomios

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n$$

y

$$q(x) = b_0 + b_1 x + \cdots + b_m x^m$$

son iguales si y sólo si los coeficientes en cada potencia de ambos son iguales:

- $n = m$,
- $a_0 = b_0, a_1 = b_1, \ldots, a_n = b_n$.

Así en $\mathbb{Z}/2\mathbb{Z}$, los polinomios

$$p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$$

y

$$q(x) = 1 + x^3$$

son iguales si (y sólo si) $a_0 = a_3 = 1$ y $a_1 = a_2 = 0$.

Por otro lado dos funciones f y g de A en B, son iguales si para todo $a \in A$ se tiene que $f(a) = g(a)$.

Dos polinomios iguales también son iguales como funciones, pero puede ocurrir que dos polinomios sean iguales como funciones, pero no como polinomios. Por ejemplo en $\mathbb{Z}/2\mathbb{Z}$, los polinomios

$$p(x) = 1 + x$$

y

$$q(x) = 1 + x^3.$$

Son claramente diferentes como polinomios, pero vistos como funciones son idénticos:

$$p(0) = 1 = q(0), \ \mathbf{y} \ p(1) = 0 = q(1).$$

Este fenómeno no ocurre si el anillo sobre el que se definen los polinomios es un cuerpo infinito (para polinomios con coeficientes sobre un cuerpo infinito igualdad coeficiente a coeficiente es equivalente a igualdad como funciones).

De ahora en más trabajaremos con un cuerpo $\mathbb{F}$ fijo. El símbolo $\mathbb{F}[x_1, \ldots, x_n]$ denota el anillo de polinomios en n indeterminadas sobre el cuerpo $\mathbb{F}$ (decir "sobre el anillo/cuerpo" es equivalente a decir "a coeficientes en el anillo/cuerpo").

Un monomio de grado t (donde t es un entero no-negativo) es una expresión de la forma

$$x_1^{t_1} x_2^{t_2} \cdots x_n^{t_n}$$

donde cada t_i es un entero no-negativo tal que

$$t_1 + t_2 + \cdots + t_n = t.$$

Observaciones:

- El único monomio de grado 0 (cero) es la constante 1.
- El polinomio **nulo** es el polinomio cuyos coeficientes son todos nulos.
- Hay $\binom{n+t-1}{t}$ monomios en n variables de grado t.
- Todo polinomio de $\mathbb{F}[x_1, \ldots, x_n]$ es una combinación lineal de monomios con coeficientes tomados de $\mathbb{F}$.
- Sea $p \in \mathbb{F}[x_1, \ldots, x_n]$, el grado de p, denotado $\deg(p)$, es el máximo grado de sus monomios con coeficientes no-nulos.
- Un polinomio se dice **homogéneo** si todos sus monomios con coeficientes no-nulos tienen el mismo grado.
- Una punto $a \in \mathbb{F}^n$ tal que $p(a) = 0$ es una **raíz** del polinomio $p \in \mathbb{F}[x_1, \ldots, x_n]$.
- Diremos que un polinomio $p \in \mathbb{F}[x_1, \ldots, x_n]$ **se anula sobre** un conjunto $E \subset \mathbb{F}^n$ si para todo elemento $e \in E$ tenemos que $p(e) = 0$.

Ocurre que con polinomios en varias variables sobre cuerpos infinitos podemos perder la relación entre el grado del polinomio y el número de raíces, por ejemplo

$$p(x, y) = x^2 + y^2 - 1$$

visto como polinomio de $\mathbb{R}[x, y]$ tiene infinitas raíces.

El siguiente ejemplo muestra que aún sobre cuerpos finitos la relación entre grado y número de raíces es no trivial para polinomios en varias indeterminadas. Consideremos el polinomio

$$x_1x_2 + x_2x_3 + x_3x_4 + x_4x_5 + x_5x_1$$

sobre $\mathbb{F}_2^5$, donde $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$. Este polinomio de grado 2 tiene 16 raíces:

0	0	0	0	0		0	1	0	1	0
0	0	0	0	1		0	1	0	0	1
0	0	0	1	0		0	0	1	0	1
0	0	1	0	0		0	0	1	1	1
0	1	0	0	0		0	1	1	1	0
1	0	0	0	0		1	1	1	0	0
1	0	1	0	0		1	1	0	0	1
1	0	0	1	0		1	0	0	1	1

mientras que el polinomio

$$x_1^2 - x_1$$

se anula sobre todo $\mathbb{F}_2^5$.

Los siguientes lemas recuperan la relación entre grado de un polinomio y número de raíces para polinomios en varias indeterminadas.

Lema 3.1. *Dado un conjunto $E \subset \mathbb{F}^n$ de tamaño*

$$|E| < \binom{n+d}{d}.$$

Existe un polinomio no-nulo $p \in \mathbb{F}[x_1, \ldots, x_n]$ de grado a lo más d que se anula sobre E.

Haremos algunas acotaciones antes de proceder a demostrar el lema. Como

$$\binom{n+d}{d} \longrightarrow \infty, \text{ cuando } d \to \infty$$

siempre podremos hallar un d (relativamente pequeño) tal que $|E| < \binom{n+d}{d}$. Por ejemplo, si tomamos los siguientes 5 puntos de $\mathbb{F}_2^5$,

$$\begin{array}{ccccc} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{array} .$$

El lema nos garantiza que existe un polinomio de grado 1 que se anula sobre estos puntos, pues

$$5 < 6 = \binom{6}{1} = \binom{5+1}{1}.$$

Por ejemplo $p(x_1, \ldots, x_5) = x_1 + x_3 + 1$.

Ahora daremos la demostración del lema 3.1.

Demostración. Denotemos por V_d al espacio vectorial de todos los polinomios de $\mathbb{F}[x_1, \ldots, x_n]$ de grado a lo más d. El conjunto de todos los monomios de grado a lo más d forma una base de V_d y como hay:

$$\sum_{k=0}^{d} \binom{n+k-1}{k} = \binom{n+d}{d}$$

monomios distintos de grado a lo más d tenemos que la dimensión de

$$\dim(V_d) = \binom{n+d}{d}.$$

Por otro lado, el espacio $\mathbb{F}_E$de todas las funciones $g : E \to \mathbb{F}$ tiene dimensión

$$|E| < \binom{n+d}{d}.$$

Por lo tanto tenemos que el mapeo (lineal) evaluación:

$$f \to (f(a))|_{a \in E}$$

de V_d en $\mathbb{F}^E$ no es inyectivo. Así al menos dos polinomios f_1 y f_2 de v_d deben ser mapeados al mismo elemento de $\mathbb{F}^E$. Así el polinomio $f = f_1 - f_2$ pertenece a V_d y es mapeado al elemento nulo de $\mathbb{F}^E$, i.e., f se anula sobre E. □

Ahora daremos la generalización al caso de polinomios en varias indeterminadas del hecho que un polinomio (en una variable) de grado d tiene a lo más d raíces.

Lema 3.2. *Todo polinomio $p \in \mathbb{F}[x_1, \ldots, x_n]$ de grado d, donde $\mathbb{F}$ es un cuerpo finito con q elementos, tiene a lo más dq^{n-1} raíces.*

Por ejemplo, el lema nos garantiza que el polinomio $x_1x_2 + x_2x_3 + x_3x_4 + x_4x_5 + x_5x_1$ sobre $\mathbb{F}_2^5$ tienen a lo más 16 raíces, y que el polinomio $x_1^2 + 2$ sobre $\mathbb{F}_3^2$ tiene a lo más 6 raíces.

La demostración que daremos sigue a Dvir (2009) y Moshkovitz (2010).

Demostración. Sea $q = |\mathbb{F}|$, sin pérdida de generalidad podemos asumir que:

- $n \geq 2$,
- $1 \leq d \leq q$.

La demostración es por reducción al caso $n = 1$. Escribamos $f = g + h$ donde

- g es homogénea de grado d,
- h solo contiene monomios de grado estrictamente menor que d.

Cómo f no es el polinomio nulo, existe algún $w \in \mathbb{F}^n$, $w \neq 0$, tal que $g(w) = 0$ (recordar que g es homogéneo). Para cada $u \in \mathbb{F}^n$ definimos la recta que pasa por u en la dirección w:

$$L_u := \{u + tw \ : \ t \in \mathbb{F}\}.$$

Un par de observaciones:

- $L_u \cap L_v = \emptyset$ si $v \notin L_u$,
- $|L_u| = q$.

Por lo que podemos particionar a $\mathbb{F}^n$ en $q^n/q = q^{n-1}$ rectas.

Ahora mostraremos que sobre cada recta L_u puede haber a lo más d raíces de f. Para cada $u \in \mathbb{F}^n$, la función

$$p_u(t) := f(u + tw)$$

es un polinomio en la variable t de grado d, de hecho el coeficiente de t^d es $g(w) \neq 0$. Cómo $p_u(t)$ tiene a lo más d raíces, tenemos que el polinomio f puede tener a lo más d raíces en cada L_u. Como $\mathbb{F}^n$ fue particionado en q^{n-1} rectas, tenemos que f no puede tener más de dq^{n-q} raíces. □

DeMillo y Lipton (1978), Zippel (1979) y Schwartz (1980) de forma independiente probaron la siguiente generalización.

Lema 3.3 (DeMillo-Lipton-Schwartz-Zippel)**.** *Para todo conjunto $S \subset \mathbb{F}$ tal que $|S| > d$, se tiene que todo polinomio $f \in \mathbb{F}[x_1, \ldots, x_n]$ de grado d puede tener a lo más $d|S|^{n-1}$ raíces en S^n.*

Demostración. Suponga que f es un polinomio no nulo. La demostración es por inducción sobre n, el número de variables de f.

Para $n = 1$ el lema es trivialmente verdadero.

Ahora analicemos el caso $n \geq 2$. Para hacerlo reescribamos f en términos de las potencias de x_n:

$$f = f_0 + f_1 x_n + f_2 x_n^2 + \cdots + f_t x_n^t$$

donde $f_0, f_1, \ldots, f_t$ son polinomios en las $n - 1$ variables $x_1, x_2, \ldots, x_{n-1}$, el polinomio f_t no es el polinomio nulo, y $t \leq d$.

Ahora vamos a estimar (por arriba) el número de puntos $(a, b) \in S^{n-1} \times S$, tales que $f(a, b) = 0$. Analizaremos dos casos

> Caso 1. $f_t(a) = 0$. Como f_t no es el polinomio nulo, y tiene grado a lo más $d - t$, por hipótesis inductiva, este polinomio se anula a lo más en $(d - t)|S|^{n-2}$ puntos de S^{n-1}. Por lo tanto, en este caso, hay a lo más $(d - t)|S|^{n-1}$ puntos de $(a, b) \in \Sigma^{n-1} \times S$ para los cuales $f(a, b) = 0$ y $f_t(a) = 0$.
> Caso 2. $f_t(a) \neq 0$. Para todo punto $a \in S^{n-1}$ tal que $f_t(a) \neq 0$, el polinomio $f(a, x_n)$ es un polinomio en una variable de grado t, y no es el polinomio nulo. Por o tanto tiene a lo más t raíces. Como a lo suma hay $|S|^{n-1}$ de tales puntos a, el número de puntos $(a, b) \in S^{n-1} \times S$ para los cuales $f(a, b) = 0$ y $f_t(a) \neq 0$, es a lo más $t|s|^{n-1}$.

Por lo tanto, a lo más hay

$$(d-t)|S|^{n-1} + t|S|^{n-1} = d|S|^{n-1}$$

puntos de $(a,b) \in S^n$ para los cuales $f(a,b) = 0$. □

3.2. Solución del problema de Kakeya en cuerpos finitos. La conjetura de Kakeya es uno de los problemas no resueltos favoritos de Terence Tao en teoría geométrica de la medida. Esta conjetura desciende del problema de la aguja, planteado por Kakeya en 1971:

¿Cuál es la menor área en el plano requerida para rotar completamente (i.e. 360° de forma continua una aguja de longitud unitaria (y grosor cero)?

Por ejemplo, es claro que podemos hacer esto en un círculo de diámetro uno, y por lo tanto de área $\pi/4$. Pero al usar una deltoide solo usamos un área de $\pi/8$. En 1928, Besicovitch demostró que de hecho se puede rotar una aguja unitaria usando sólo una cantidad arbitrariamente pequeña de área (en realidad resolvió otro problema, pero su trabajo aplica). Este hecho contraintuitivo, motivó la definición de lo que conocemos como **conjunto de Kakeya-Besicovitch**: es un conjunto que contiene un segmento unitario en cada dirección.

En 1981, Davies demostró que existen conjuntos de Kakeya-Besicovitch con medida de Lebesgue cero, pero los cuales seguían siendo objetos esencialmente bi-dimensionales: tienen dimensión de Hausdorff 2 y dimensión de Minkowski 2.Esto llevo a la siguiente conjetura en mayores dimensiones:

Conjetura 3.4 (de Kakeya:)**.** *Un conjunto de Kakeya-Besicovitch en $\mathbb{R}^n$ tiene dimensión de Minkowski y dimensión de Hausdorff igual a n.*

En 1999, Wolff propuso el análogo de la conjetura de Kakeya sobre cuerpos finitos a fin de evitar los problemas técnicos asociados con las dimensiones de Minkowski y Hausdorff. Dado un cuerpo finito $\mathbb{F}$, definimos un **conjunto de Kakeya** como cualquier subconjunto K de $\mathbb{F}^n$ que contenga una línea en cada dirección, i.e., para cada vector $w \in \mathbb{F}^n$ existe un vector $v \in \mathbb{F}^n$ tal que la línea $L_{v,w} := \{v + tw \,:\, t \in \mathbb{F}\} \subset K$.

Conjetura 3.5 (Conjetura de Kakeya en cuerpos finitos, Wolff 1999)**.** *Sea $E \subset \mathbb{F}^n$ un conjunto de Kakeya. Entonces $|K| \geq c_n|\mathbb{F}|^n$, donde c_n es una constante positiva que depende sólo de n.*

En 2009, Dvir sorprendió a toda la comunidad matemática que trabajaba en el las diferentes conjeturas de Kakeya (Tao incluido), al usar el método polinomial para probar la conjetura de Kakeya en cuerpos finitos.

Lema 3.6. *Sea $f \in \mathbb{F}[x_1, \ldots, x_n]$ un polinomio de grado a lo más $q-1$ sobre un cuerpo finito $\mathbb{F}$ con $q = |\mathbb{F}|$ elementos. Si f se anula sobre un conjunto de Kakeya K, entonces f es el polinomio nulo.*

El argumento de Dvir es el siguiente.

Demostración. Supongamos que f no es el polinomio nulo. Reescribamos f de la forma

$$f = \sum_{i=0}^{d} f_i$$

donde $0 \leq d \leq q-1$ es el grado de f y cada f_i es la componente homogénea de grado i de f, por lo que f_d es no nula. Como f es no-nulo y se anula sobre K, tenemos que $d \neq 0$.

Tomemos una dirección arbitraria w de $\mathbb{F}^n$, i.e. $0 \neq w \in \mathbb{F}^n$. Como K es un conjunto de Kakeya, existe $v \in \mathbb{F}^n$ tal que la recta $L_{v,w} = \{v + tw \ : \ t \in \mathbb{F}\} \subset K$. Por lo tanto

$$f(v + tw) = 0$$

para todo $t \in \mathbb{F}$. Ahora miremos un poco más en detalle a $f(v + tw)$. Es un polinomio en t de grado a lo más $q - 1$ sobre el cuerpo $\mathbb{F}$ (pues es del mismo grado que f). Pero se anula sobre todo $\mathbb{F}$, por lo que no le queda otra que ser el polinomio nulo. Ahora el coeficiente que acompaña a t^d es $f_d(w)$, por lo tanto $f_d(w) = 0$. Esto demuestra que f_d se anula sobre todos los puntos de $\mathbb{F}^n$. Pero como

$$dq^{n-1} \leq (q-1)q^{n-1} < q^n.$$

Tenemos que por el lema 3.2 que el polinomio f_d debe ser el polinomio nulo, lo que es una contradicción. □

Teorema 3.7 (Dvir 2009)**.** *Sea $K \subset \mathbb{F}^n$ un conjunto de Kakeya. Entonces*

$$|K| \geq \binom{|\mathbb{F}| + n - 1}{n} \geq \frac{|\mathbb{F}|^n}{n!}.$$

Demostración. Supongamos que

$$|K| < \binom{|\mathbb{F}| + n - 1}{n}.$$

Por el lema 3.1, existe un polinomio no-nulo $F \in \mathbb{F}[x_1, \ldots, x_n]$ de grado a lo más $|\mathbb{F}| - 1$ que se anula sobre K, lo cual contradice el lema 3.6. □

3.3. Nullstellensatz Combinatorio de Alon. Consideremos el siguiente "teorema" bien conocido de cursos básicos:

Teorema 3.8. *Sea $f \in \mathbb{F}[x]$ un polinomio de grado t. Si $S \subset \mathbb{F}$ tal que $|S| \geq t + 1$, entonces existe $s \in S$ tal que $f(s) \neq 0$.*

El Nullstellensatz Combinatorio de Alon (1999) generaliza este resultado para el caso de varias variables.

Teorema 3.9 (Nullstellensatz Combinatorio)**.** *Sea $f \in \mathbb{F}[x_1, \ldots, x_n]$ de grado d. Suponga que los coeficientes del monomio de f*

$$x_1^{t_1} \cdots x_n^{t_n}$$

es no nulo y que

$$t_1 + \cdots + t_n = d.$$

Si $S_1, \ldots, S_n$ son subconjuntos finitos de $\mathbb{F}$ tales que $|S_i| \geq t_i + 1$, entonces existe un punto $s \in S_1 \times \cdots \times S_n$ para el cual $f(s) \neq 0$.

El Nullstellensatz Combinatorio también es una generalización del lema de DeMillo-Lipton-Schartz-Zippel.

Este resultado sigue del un caso especial del Nullstellensatz de Hilbert:

Teorema 3.10 (Nullstellensatz)**.** *Sea $f \in \mathbb{F}[x_1, \ldots, x_n]$ y $S_1, \ldots, S_n$ son subconjuntos finitos del cuerpo $\mathbb{F}$. Si $f(s) = 0$ para todo $s \in S_1 \times \cdots \times S_n$, entonces existen polinomios $h_1, \ldots, h_n \in \mathbb{F}[x_1, \ldots, x_n]$ tales que*

$$\deg(h_i) \leq \deg(f) - |S_i|$$

y

$$f([x_1,\ldots,x_n) = \sum_{i=1}^{n} h_i([x_1,\ldots,x_n) \prod_{s\in S_i}(x_i - s).$$

Daremos la demostración de Alon (1999) del Nullstellensatz, la cual requiere del siguiente lema.

Lema 3.11. *Sea $f \in \mathbb{F}[x_1,\ldots,x_n]$. Supongamos que el grado de f como polinomio en x_i es a lo más t_i, y que tenemos subconjuntos $S_i \subset \mathbb{F}$ tales que $|S_i| \geq t_i+1$. Si $f(s) = 0$ para todo $s \in S_1 \times \cdots \times S_n$, entonces f es el polinomio nulo.*

Demostración. Es por inducción sobre n, el número de variables. Para $n = 1$ el lema establece que un polinomio en una variable de grado a lo más t_1 puede tener a lo sumo t_1 raíces distintas. Asumiendo que el lema vale para $n-1$, probaremos que vale para $n \geq 2$.

Dando un polinomio $f = f(x_1,\ldots,x_n)$ y conjuntos S_i que satisfagan las hipótesis del lema, reescribamos f como un polinomio en x_n:

$$f = \sum_{i=0}^{t_n} f_i(x_i,\ldots,x_{n-1})x_n^i$$

donde cada f_i es un polinomio en las $n-1$ variables $x_1,\ldots,x_{n-1}$, donde el grado en x_j es a lo más t_j. Para cada

$$(s_1,\ldots,s_{n-1}) \in S_1 \times \cdots \times S_{n-1}.$$

Consideremos el polinomio en una variable x_n dado por

$$f(s_1,\ldots,s_{n-1},x_n).$$

Por hipótesis tenemos que para todo $s \in S_n$ se cumple que

$$f(s_1,\ldots,s_{n-1},s) = 0.$$

Por lo tanto, $f(s_1,\ldots,s_{n-1},x)$ es el polinomio nulo. Esto implica que $f_i(s_1,\ldots,s_{n-1}) = 0$ para todo $(s_1,\ldots,s_{n-1}) \in S_1 \times \cdots \times S_{n-1}$. Luego, por hipótesis inductiva, f_i es el polinomio nulo para cada i, lo que implica que f es el polinomio nulo. □

Ahora procederemos dar la demostración de Alon del Nullstellensatz.

Demostración del Teorema 3.10, Nullstellensatz, Alon (1999). Definimos para cada i

$$d_i = |S_i| - 1$$

y consideremos los polinomios

$$g_i(x_i) := \prod_{s\in S_i}(x_i - s) = x_i^{d_i+1} - \sum_{j=0}^{d_i} a_{ij}x_i^j.$$

Observe que si $s \in S - i$, entonces $g_i(s) = 0$, lo que implica que

$$s^{d_i+1} = \sum_{j=0}^{d_i} a_{ij}x_i^j.$$

Esto sugiere utilizar la siguiente identidad para reducir el "x_i-grado"

$$x_i^{d_i+1} = \sum_{j=0}^{d_i} a_{ij}x_i^j. \tag{3.1}$$

Sea $\bar{f}$ el polinomio que se obtiene a partir de f, escribiendo a f como combinación lineal de monomios y reemplazando, las veces que haga falta, cada ocurrencias de potencial altas: $x_i^{t_i}$ con $t_i > d$ $(1 \leq i \leq n)$; por combinaciones lineales de potencias más pequeñas de x_i, usando la relación 3.1.

Miremos con atención al polinomio $\bar{f}$. El grado de $\bar{f}$ es a lo más d_i en x_i para cada $1 \leq i \leq n$. Tras meditarlo un rato, nos damos cuenta que $\bar{f}$ puede ser obtenido a partir de f restándole productos de la forma $h_i g_i$, donde

$$\deg(h_i) \leq \deg(f) - \deg(g_i) = \deg(f) - |S_i|$$

Por lo tanto

$$\bar{f}(x) = f(x) - \sum_{i=1}^{n} h_i(x) g_i(x). \tag{3.2}$$

Además como para $s \in S_1 \times \cdots \times S_n$ vale la relación 3.1, tenemos que

$$\bar{f}(s) = f(s) = 0.$$

Luego, por el lema 3.11, tenemos que $\bar{f}(s) = 0$ para todo $s \in \mathbb{F}^n$. Esto, junto con 3.2, implican que $f = \sum_{i=1}^{n} h_i g_i$. □

Ahora demostraremos el Nullstellensatz Combinatorio de Alon, 1999.

Demostración del Teorema 3.9. Sin pérdida de generalidad podemos asumir $|S_i| = t_i + 1$ para cada i. Supongamos que el resultado es falso, i.e., que para todo $t \in S_1 \times \cdots \times S_n$ tenemos que $f(t) = 0$. Definimos

$$g_i(x_i) = \prod_{s \in S_i} (x_i - s).$$

Sean $h_1, \ldots, h_n$ los polinomios que nos da el Nullstellensatz (teorema 3.10), luego:

$$\deg(h_i) \leq \deg f - \deg(g_i) = \deg(f) - t_i - 1 \tag{3.3}$$

y además

$$f(x) = \sum_{i=1}^{n} h_i(x) g_i(x_i).$$

Lo que se puede reescribir de la siguiente forma:

$$f(x) = \sum_{i=1}^{n} x_i^{t_i+1} h_i(x) + (\text{términos de grado } < \deg(f)).$$

Por hipótesis, el coeficiente de $\prod_{i=1}^{n} x_i^{t_i}$ del lado izquierdo de la ecuación anterior es no cero, pero es imposible tener tal monomio del lado derecho, lo que nos da una contradicción. □

Para finalizar estas notas daremos una aplicación del Nullstellensatz Combinatorio a la teoría de números aditiva.

Dados dos conjuntos A y B de elementos de un cuerpo $\mathbb{F}$, definimos su **conjunto suma**

$$A + B := \{a + b \mid a \in A, b \in B\}.$$

Por ejemplo es $\mathbb{Z}/5\mathbb{Z}$, si $A = \{2, 4\}$ y $B = \{2, 3\}$ tenemos que

$$A + B = \{0, 1, 2, 4\}.$$

Una pregunta relevante es si el tamaño de A y B nos dan un indicio de que del tamaño mínimo que puede tener $A + B$. El teorema de Cauchy-Davenport da una respuesta.

Teorema 3.12 (Cauchy-Davenport). *Si p es un primo, y A y B son dos subconjuntos no-vacíos de $\mathbb{Z}_p$, entonces*

$$|A+B| \geq \text{mín}\{p, |A|+|B|-1\}.$$

Usaremos el Nullstellensatz Combinatorio para realizar la prueba.

Demostración. Si $|A|+|B| > p$ es resultado es trivial, pues para todo $x \in \mathbb{Z}_p$ los conjuntos A y $x-B$ tienen intersección no vacía, lo que implica que $A+B=\mathbb{Z}_p$.

Asumamos entonces que $|A|+|B| \leq p$. La demostración procede por reducción al absurdo: supongamos que

$$|A+B| \leq |A|+|B|-2.$$

Sea $C \subset \mathbb{Z}_p$ tal que $A+B \subset C$ y $|C| = |A|+|B|-2$.

Definamos el polinomio

$$f(x,y) = \prod_{c \in C}(x+y-c).$$

Por definición,

$$f(a,b) = 0 \text{ para todo } (a,b) \in A \times B.$$

Ahora, llamando $t_1 = |A|-1$ y $t_2 = |B|-1$, notamos que el coeficiente de $x^{t_1}y^{t_2}$ en f es:

$$\binom{t_1+t_2}{t_1} = \binom{|A|+|B|-2}{|A|-1}$$

el cual es no cero en $\mathbb{Z}_p$, pues $|A|+|B|-2 < p$. Aplicando el Nullstellensatz Combinatorio obtenemos un punto $(a,b) \in A \times B$ para el cual $f(a,b) \neq 0$, contradicción. □

3.4. Ejercicios.

1. Sea p un primo impar. Consideremos un cuerpo $\mathbb{F}_p$. Entonces para todo $t \leq p-2$

$$\sum_{x \in \mathbb{F}_p} x^t = 0$$

en $\mathbb{F}_p$.

2. Demostrar el Lema de DeMillo-Lipton-Schwartz-Zippel usando el Nullstellensatz Combinatorio de Alon.
3. Probar el teorema de Olson usando el Nullstellensatz Combinatorio de Alon: Sean $(a_1,b_1),\ldots,(a_n,b_n)$ una sucesión de elementos en $\mathbb{Z}_p$ con $n \geq 2p-1$. Entonces existe un subconjunto no vacío $A \subset [n]$ tal que

$$\sum_{i \in A}(a_i,b_i) = (0,0).$$

4. Dos números distintos son escritos sobre los vértices de un 100-agono convexo. Probar que siempre podemos remover un número de cada vértice de forma en vértices adyacentes no queden números iguales.
5. Sea n un entero positivo. Consideremos el conjunto:

$$S = \{(x,y,z)\,|x,y,z \in [n], (x,y,z) \neq (0,0,0)\}.$$

Determinar el número más pequeño de planos tales que su unión contenga a S pero no al $(0,0,0)$.

6. Probar el teorema de Chevalley-Warning usando el Nullstellensatz Combinatorio: Sea p un primo, y $f_1, \ldots, f_m$ polinomios en $\mathbb{F}_p[x_1, \ldots, x_n]$. Si
$$\sum_{i=1}^{m} \deg(f_i) < n$$
entonces el número de ceros comunes de $f_1, \ldots, f_m$ es divisible por p.
7. Demostrar el teorema de Erdös-Ginzburg-Ziv (1961) a la Alon (i.e. usando el teorema de Chevalley-Warning, ver ejercicio 6): Cualquier sucesión de $2n-1$ enteros contiene una subsucesión de longitud n, tal que su suma es divisible por n.

Referencias

[1] Alon, Noga (1999): *Combinatorial Nullstellensatz*, Comb. Prob. Comput. 8, 7–29.
[2] Babai, L and Frankl, P (1992): *Linear Algebra Methods in Combinatorics*, Preliminary Version 2, University of Chicago.
[3] Brualdi, R (2010): *Introductory Combinatorics*, 5th Edition. Pretince Hall.
[4] Jukna, Stasys (2011): *Extremal Combinatorics*. 2nd Edition. Springer.
[5] Tait, M. (2013): *My favorite application using eigenvalues: Eigenvalues and the Graham-Pollak Theorem*. Personal Comunication

Departamento de Matemáticas, Facultad de Ciencias Físico Matemáticas y Naturales, Universidad Nacional de San Luis

INTRODUCCIÓN A LAS ÁLGEBRAS DE LIE

VANESA MEINARDI

Índice

Introducción

Los objetivos de este curso son introducir las nociones básicas de álgebras de Lie. El mismo consta de dos partes. En la primer parte se introducen las definiciones, nociones básicas y ejemplos de álgebras de Lie. En la segunda parte, se presenta el teorema de Weyl, que establece la completa reducibilidad de las representaciones de dimensión finita. Cabe destacar que en algunos casos sólo se hará mención de los enunciados de ciertos Teoremas sin su prueba correspondiente, con el objetivo de dejar sólo una noción de la importancia de los mismo en los alumnos. Las demostraciones de ellos podrán ser encontradas fácilmente en la bibliografía citada.

1. Nociones básicas de álgebras de Lie

1.1. Definiciones y primeros ejemplos de álgebras de Lie.

Definición 1.1. Sea k un anillo conmutativo con unidad con característica cero. Un *álgebra* sobre k es un par (V, μ) donde V es un módulo sobre k y $\mu : V \times V \to V$ es una aplicación bilineal.

(Denotaremos, de ahora en más, de la misma manera a una aplicación bilineal y a la correspondiente aplicación lineal desde el producto tensorial).

Un *ideal a izquierda* de (V, μ) (respectivamente *ideal a derecha*) es un k-submódulo W de V tal que $\mu(W \otimes V) \subseteq W$ (resp. $\mu(V \otimes W) \subseteq W$). Una *subálgebra* de (V, μ) es un k-submódulo W de V tal que $\mu(W \otimes W) \subseteq W$. Un *morfismo de álgebras* es un morfismo de k-módulos $f : (V, \mu) \to (V', \mu')$ tal que $f(\mu(v, w)) = \mu'(f(v), f(w))$ para todo $v, w \in V$.

Definición 1.2. Un *álgebra de Lie* sobre k es un álgebra $\mathfrak{g}$ con una operación $\mathfrak{g} \otimes \mathfrak{g} \to \mathfrak{g}$ denotada por $(x, y) \to [x, y]$ llamada *corchete o conmutador* de x e y, tal que satisface los siguientes axiomas:

(L_1) La operación es antisimétrica, i.e. $[x,\ y] = -[y, x]$ para todo $x, y \in \mathfrak{g}$.
(L_2) $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ para todo $x, y, z \in \mathfrak{g}$.

Date: 17 de Julio del 2014.
2010 *Mathematics Subject Classification.* 17B05.
CIEM&CONICET.

El axioma (L_2) es llamado la *identidad de Jacobi*. Notar que $(\mathfrak{g}, [\,,])$ es un álgebra. Luego las nociones de ideal, subálgebras y morfismos se siguen de las definiciones previas.

Si A, B son submódulos de $\mathfrak{g}$ denotaremos $[A, B]$ el submódulo generado por los elementos de la forma $[a, b]$ con $a \in A, b \in B$.

Definición 1.3. Sea $(\mathfrak{g}, [\,,])$ un álgebra de Lie, $\mathfrak{h}$ una subálgebra de $\mathfrak{g}$ y $\mathfrak{a}$ un submódulo de $\mathfrak{g}$, se define el *normalizador* de $\mathfrak{a}$ en $\mathfrak{h}$ como

$$\mathrm{norm}_{\mathfrak{h}}(\mathfrak{a}) = \{x \in \mathfrak{h} : [x, \mathfrak{a}] \subseteq \mathfrak{a}\}$$

y el *centralizador* de $\mathfrak{a}$ en $\mathfrak{h}$ como

$$\mathrm{cent}_{\mathfrak{h}}(\mathfrak{a}) = \{x \in \mathfrak{h} : [x, \mathfrak{a}] = 0\}.$$

Ambas son subálgebras de $\mathfrak{g}$ mientras que $\mathrm{cent}_{\mathfrak{h}}(\mathfrak{a})$ es un ideal de $\mathrm{norm}_{\mathfrak{h}}(\mathfrak{a})$.

Ejemplo 1.4. (i) Si V es un k-módulo y $[\,,]$ es la aplicación bilineal nula, $(V, [\,,])$ es un álgebra de Lie. Tales álgebras de Lie son llamadas *abelianas*.

(ii) A cada álgebra asociativa A sobre k le asociamos una estructura de álgebra de Lie definiendo $[x, y] = xy - yx$. Por ejemplo si $A = \mathrm{End}_k(V)$ es el algebra de endomorfismos de un módulo V sobre k (el producto es la composición), se obtiene un álgebra de Lie que denotaremos $\mathfrak{gl}(V)$ y se llama *algebra lineal general*. En particular denotaremos $\mathfrak{gl}(n, k)$ a $\mathfrak{gl}(k^n)$.

(iii) Sea (V, μ) un álgebra, $T \in \mathrm{End}(V)$ se dice una *derivación* de (V, μ) si

$$T(\mu(x, y)) = \mu(T(x), y) + \mu(x, T(y));$$

denotaremos $\mathrm{Der}(V, \mu) = \mathrm{Der(V)}$ al espacio vectorial de todas las derivaciones de (V, μ). $\mathrm{Der}(V)$ es una subálgebra de Lie de $\mathfrak{gl}(V)$.

(iv) Sean V un k-módulo libre de rango finito y tr la forma lineal traza en End (V). Se define

$$\mathfrak{sl}(V) = \{T \in \mathrm{End}\ (V) :\ \mathrm{Tr}(T) = 0\}$$

es un ideal de $\mathfrak{gl}(V)$. De hecho, se puede probar que $[\mathfrak{gl}(V), \mathfrak{gl}(V)] = \mathfrak{sl}(V)$, pues $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ si $A, B \in \mathrm{End}(V)$.

(v) Sean V un k-módulo y $b : V \otimes V \to k$ una forma bilineal. Se define

$$\mathfrak{g}(b) = \{T \in \mathrm{End(V)} :\ b(T(v), w) + b(v, T(w)) = 0\, \forall\, v, w \in V\}.$$

Entonces $\mathfrak{g}(V)$ es un álgebra de Lie.

(vi) Describamos las 4 familias clásicas de álgebras de Lie:
A_N : es la notación correspondiente a $\mathfrak{sl}(N + 1, k)$.
B_N : son las álgebras de Lie ortogonales correspondientes a $b : k^{2N+1} \otimes k^{2N+1} \to k$, $b(v, w) = {}^t vSw$, donde identificamos los vectores de k^{2N+1} con vectores columnas y

$$S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & \mathrm{id}_N \\ 0 & \mathrm{id}_N & 0 \end{pmatrix} \in \mathrm{Mat}_{2N+1}(k);$$

id_N es la matriz identidad de $\mathrm{Mat}_N(k)$.
En este caso denotaremos $\mathfrak{g}(b) = so(2N + 1, k)$. No es difícil ver que

$$so(2N + 1, k) = \{x \in \mathfrak{gl}(2N + 1, k) : xS + S^t x = 0\}$$

$$= \left\{ \begin{pmatrix} 0 & b_1 & b_2 \\ c_1 & m & n \\ c_2 & p & q \end{pmatrix} \in \mathfrak{gl}(2N + 1, k) : \begin{array}{lll} c_2 = -{}^t b_1, & c_1 = -{}^t b_2, & \\ q = -{}^t m, & n = -{}^t n & p = -{}^t p \end{array} \right\}.$$

C_N : son las álgebras de Lie simplécticas correspondientes a $b : k^{2N} \otimes k^{2N} \to k$, $b(v,w) = {}^tvQw$, donde

$$Q = \begin{pmatrix} 0 & \mathrm{id}_N \\ -\mathrm{id}_N & 0 \end{pmatrix} \in \mathrm{Mat}_{2N}(k).$$

Estas álgebras de Lie serán denotadas $sp(2N,k)$. Como antes no es difícil ver que

$$sp(2N,k) = \left\{ \begin{pmatrix} m & n \\ p & q \end{pmatrix} \in \mathfrak{gl}(2N,k) : q = -{}^tm,\ n \text{ y} p \text{ son simétricas} \right\}.$$

D_N : son las álgebras de Lie ortogonales correspondientes a $b : k^{2N} \otimes k^{2N} \to k$, $b(v,w) = {}^tvPw$, donde

$$P = \begin{pmatrix} 0 & \mathrm{id}_N \\ \mathrm{id}_N & 0 \end{pmatrix} \in \mathrm{Mat}_{2N}(k);$$

Estas álgebras de Lie serán denotadas $so(2N,k)$. Explícitamente

$$so(2N,k) = \left\{ \begin{pmatrix} m & n \\ p & q \end{pmatrix} \in \mathfrak{gl}(2N,k) : q = -{}^tm,\ n \text{ y} p \text{ son antisimétricas} \right\}.$$

(vii) Construiremos álgebras de Lie nuevas a partir de álgebras de Lie dadas
   (a) Sea $(\mathfrak{g}_i)_{i\in I}$ una familia de álgebras de Lie, *el producto directo* $\prod \mathfrak{g}_i$ es un álgebra de Lie con el corchete definido coordenada a coordenada (análogamente la suma directa).
   (b) Sean $\mathfrak{a}$ y $\mathfrak{b}$ álgebras de Lie y $\theta : \mathfrak{b} \to \mathrm{Der}(\mathfrak{a})$ un morfismo de álgebras de Lie. Sea $\mathfrak{g}$ el k-módulo $\mathfrak{a} \times \mathfrak{b}$. Entonces $\mathfrak{g}$ admite una única estructura de álgebras de Lie dada por $[x,a] = \theta(x)a$, con $x \in \mathfrak{b}$ y $a \in \mathfrak{a}$. Tal que $\mathfrak{a}$ es un ideal de $\mathfrak{g}$ y $\mathfrak{b}$ es una subálgebra de $\mathfrak{g}$.
   (c) Sea $(\mathfrak{g}, [\,,])$ un álgebra de Lie se define $[a,b]_{op} = -[a,b]$ entonces $(\mathfrak{g}, [\,,]_{op})$ es un álgebra de Lie llamada *álgebra de Lie opuesta.*
   (d) Sea $(\mathfrak{g}, [\,,])$ un álgebra de Lie y A una k-álgebra asociativa, entonces $(\mathfrak{g} \otimes_k A, [\,,]_A)$ es un álgebra de Lie con $[x \otimes a, y \otimes b]_A = [x,y] \otimes ab$, para todo $x, y \in \mathfrak{g},\ a, b \in A$.

1.2. Ejercicios. En los siguientes ejercicios sea $(\mathfrak{g}, [\,,])$ un álgebra de Lie, probar que;

1. *a*) Los axiomas (L_1) y la bilinealidad del corchete es equivalente a $[x,x] = 0$ para todo $x \in \mathfrak{g}$;
   b) Ideal a izquierda es ideal a derecha y viceversa.
2. *a*) Si $\mathfrak{h}$ y $\mathfrak{h}'$ son ideales de $\mathfrak{g}$, entonces $[\mathfrak{h}, \mathfrak{h}']$ también lo es;
   b) Sea $\mathfrak{h}$ un ideal $\mathfrak{g}$ entonces el cociente $\mathfrak{g}/\mathfrak{h}$ tiene una estructura de álgebra de Lie tal que la *proyección canónica* $\pi : \mathfrak{g} \to \mathfrak{g}/\mathfrak{h}$ es un morfismo de álgebras de Lie.
3. Sean $\mathfrak{h}$ una subálgebra de $\mathfrak{g}$ y $\mathfrak{a}$ un submódulo de $\mathfrak{g}$
   a) verificar que $\mathrm{norm}_{\mathfrak{h}}(\mathfrak{a})$ y el $\mathrm{cent}_{\mathfrak{h}}(\mathfrak{a})$ son subálgebras de $\mathfrak{g}$;
   b) el $\mathrm{cent}_{\mathfrak{h}}(\mathfrak{a})$ es un ideal de $\mathrm{norm}_{\mathfrak{h}}(\mathfrak{a})$. Si $\mathfrak{a} = \mathfrak{h} = \mathfrak{g}$, entonces el $\mathrm{cent}_{\mathfrak{g}}(\mathfrak{g})$ es el centro de $\mathfrak{g}$ y será denotado por $\mathrm{cent}(\mathfrak{g})$.
4. Si $\phi : A \to B$ es un morfismo de álgebras asociativas entonces ϕ es un morfismo de álgebras de Lie.
5. *a*) Sea (V, μ) un álgebra, probar que $\mathrm{Der}(V)$ es una subálgebra de Lie de $\mathfrak{gl}(V)$.

b) Sea $(\mathfrak{g}, [\,,\,])$ un álgebra de Lie, el corchete $[\,,\,]$ induce una transformación lineal $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ dada por $\mathrm{ad}(x)(y) = [x, y]$. Probar que $\mathrm{ad}(x) : \mathfrak{g} \to \mathfrak{g}$ es una derivación para cada $x \in \mathfrak{g}$.

6. Sean V un k-módulo y $b : V \otimes V \to k$ una forma bilineal probar que,
   a) el álgebra $\mathfrak{g}(b)$ asociada a la forma bilineal b definida en el ejemplo (v) es un álgebra de Lie.
   b) Si b y b' son formas bilineales conjugadas por un automorfismo de V entonces $\mathfrak{g}(b)$ y $\mathfrak{g}(b')$ son isomorfas.
7. Probar que $[\mathfrak{gl}(V), \mathfrak{gl}(V)] = \mathfrak{sl}(V)$ y que $\mathfrak{gl}(V) = \mathfrak{sl}(V) + kId$.
8. Considerar el ejemplo (vii) inciso (b), probar que $\mathfrak{a}$ es un ideal de $\mathfrak{g}$ y $\mathfrak{b}$ es una subálgebra de $\mathfrak{g}$.
9. Sea $(\mathfrak{g} \otimes_k A, [\,,\,]_A)$ como en el ejemplo (vii) inciso (d), probar que con esta estructura $\mathfrak{g} \otimes_k A$ es un álgebra de Lie.

1.3. Módulos y representaciones. En lo que sigue $\mathfrak{g}$ será un álgebra de Lie sobre k.

Definición 1.5. Un $\mathfrak{g}$-*módulo* es un par $(V, \cdot)$ donde V es un k-módulo y $\cdot : \mathfrak{g} \otimes V \to V$ es una aplicación bilineal tal que

$$[x, y] \cdot v = x \cdot (y \cdot v) - y \cdot (x \cdot v), \quad \text{para todo } x, y \in \mathfrak{g},\ v \in V.$$

Una noción equivalente a la de $\mathfrak{g}$-módulo es la de representación. Una *representación* de $\mathfrak{g}$ en V es un morfismo de álgebras de Lie $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$.

Las nociones de subrepresentación, submódulo, morfismo de módulos, representaciones o submódulos cociente, etc. se definen como es usual.

Definición 1.6. Un módulo no nulo se dice *irreducible* si no admite submódulos propios (i.e. submódulos distintos de si mismo y de 0).

Un módulo se dice *completamente reducible* si es suma de submódulos simples.

Ejemplo 1.7. (i) Se deduce de la identidad de Jacobi que $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ es una representación que llamaremos la *representación adjunta*.

(ii) Llamaremos *representaciones naturales de las álgebras de Lie clásicas* a las respectivas inclusiones que aparecen en su definición: por ejemplo la representación natural de C_N es en k^{2N}.

(iii) Sea $\epsilon : \mathfrak{g} \to k$ la aplicación nula, es un morfismo de álgebras de Lie respecto del cual k se dice el $\mathfrak{g}$ módulo trivial.

(iv) Sean V, W $\mathfrak{g}$-módulos $V \oplus W$ es un $\mathfrak{g}$-módulo con la acción

$$x(v \oplus w) = xv \oplus xw, \quad x \in \mathfrak{g},\ v \in V,\ w \in W.$$

(v) Sean V, W $\mathfrak{g}$-módulos $V \otimes W$ es un $\mathfrak{g}$-módulo con la acción

$$x(v \otimes w) = xv \otimes w + v \otimes xw, \quad x \in \mathfrak{g},\ v \in V,\ w \in W.$$

(vi) Si V es un $\mathfrak{g}$-módulo, entonces $V^* = \mathrm{Hom}_k(V, k)$ el dual de V es un $\mathfrak{g}$-módulo con la acción

$$(xT)(v) = -T(xv), \quad T \in V^*,\ x \in \mathfrak{g},\ v \in V.$$

(vii) Sean V, W $\mathfrak{g}$-módulos $\mathrm{Hom}(V, W)$ es un $\mathfrak{g}$-módulo con la acción

$$(xT)(v) = x(T(v)) - T(xv), \quad T \in \mathrm{Hom}(V, W),\ x \in \mathfrak{g},\ v \in V.$$

1.4. Ejercicios.

1. Probar que módulo sobre una álgebra de Lie es equivalente a tener una representación.
2. Dar las definiciones precisas de subrepresentación, submódulo, morfismo de módulos, representaciones o submódulos cociente.
3. Probar que las subrepresentaciones de la representación adjunta son los ideales de $\mathfrak{g}$.
4. Verificar que los ejemplos (iii)-(vii) definen representaciones.

2. Álgebras de Lie nilpotentes y solubles

En esta sección k denota un cuerpo. Solo consideraremos álgebras de Lie $\mathfrak{g}$ de dimensión finita

Definición 2.1. Introduciremos las siguientes series de ideales de $\mathfrak{g}$

Serie derivada:

$$D^0(\mathfrak{g}) = \mathfrak{g}, \quad D^1\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}], \quad D^n(\mathfrak{g}) = [D^{n-1}(\mathfrak{g}), D^{n-1}(\mathfrak{g})].$$

Serie central descendente:

$$C^0(\mathfrak{g}) = \mathfrak{g}, \quad C^1\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}], \quad C^n(\mathfrak{g}) = [\mathfrak{g}, C^{n-1}(\mathfrak{g})].$$

Definición 2.2. Un álgebra de Lie $\mathfrak{g}$ se dice *soluble* (respectivamente *nilpotente*) si existe un entero i tal que $D^i(\mathfrak{g}) = 0$ (resp. $C^i(\mathfrak{g}) = 0$)).

Lema 2.3. *Los siguientes enunciados son equivalentes:*

(i) $\mathfrak{g}$ *es nilpotente*

(ii) *Existe un natural* n, *tal que para toda* n*-tupla* $(x_1, \cdots, x_n)$ *de elementos de* $\mathfrak{g}$, *se satisface*

$$[x_1, [x_2, [\cdots, x_n] \cdots]] = adx_1 adx_2 \cdots adx_{n-1} x_n = 0$$

(iii) $\mathfrak{g}$ *es una sucesión de extensiones centrales de álgebras de Lie abelianas. En otras palabras, existe una cadena de ideales* $\mathfrak{g} = \mathfrak{a}_1 \supset \mathfrak{a}_2 \supset \cdots \supset \mathfrak{a}_n = 0$, *tal que* $\mathfrak{a}_i/\mathfrak{a}_{i+1}$ *es el centro de* $\mathfrak{g}/\mathfrak{a}_{i+1}$.

La demostración se deja como ejercicio al lector.

Ejemplo 2.4. (i) Toda álgebra de Lie abeliana es nilpotente y soluble.

(ii) $t(n,k) = \{(a_{ij})_{1\le i,j\le n} \in \mathfrak{gl}(n,k) : a_{ij} = 0, i > j\} \subseteq \mathfrak{gl}(n,k)$ *subálgebras de matrices triangulares superiores estrictas* de $\mathfrak{gl}(n,k)$ es nilpotente.

(iii) $\eta(n,k) = \{(a_{ij})_{1\le i,j\le n} \in \mathfrak{gl}(n,k) : a_{ij} = 0, i \ge j\} \subseteq \mathfrak{gl}(n,k)$ *subálgebras de matrices triangulares* de $\mathfrak{gl}(n,k)$ es soluble.

(iv) $\mathfrak{F} = (V_i)$ una *bandera* en un espacio vectorial V, es decir una cadena ascendente de subespacios $0 = V_0 \subset V_1 \subset \cdots \subset V_n = V$ tal que la $\dim V_i = i$. Sea

$$\mathfrak{u}(\mathfrak{F}) = \{T \in \mathrm{End}(V) : T(V_i) \subset V_{i-1}\}.$$

Entonces $\mathfrak{u}(\mathfrak{F})$ es una subálgebra de Lie nilpotente de $\mathfrak{gl}(V)$. En efecto sea $\mathfrak{u}(\mathfrak{F})_k = \{T \in \mathrm{End}(V) : T(V_i) \subset V_{i-k}\}$; es fácil ver que es un ideal de $\mathfrak{u}(\mathfrak{F})$. Más aún $[\mathfrak{u}(\mathfrak{F}), \mathfrak{u}(\mathfrak{F})_k] = \mathfrak{u}(\mathfrak{F})_{k+1}$. Como $\mathfrak{u}(\mathfrak{F})_k = 0$ para $k \gg 0$, se sigue que $\mathfrak{u}(\mathfrak{F})$ es nilpotente.

(iv) Sea $\mathfrak{F}$ una bandera en V. Sea $\mathfrak{b}(\mathfrak{F}) = \{T \in \mathrm{End}(V) : T(V_i) \subset V_i\}$ es soluble como subálgebra de Lie de $\mathfrak{gl}(V)$. Pues $\mathfrak{u}(\mathfrak{F}) \subset \mathfrak{b}(\mathfrak{F})$ es un ideal de $\mathfrak{b}(\mathfrak{F})$ y $\mathfrak{b}(\mathfrak{F})/\mathfrak{u}(\mathfrak{F})$ es abeliana dado que si $T, S \in \mathfrak{b}(\mathfrak{F})$, entonces $[T, S] \in \mathfrak{U}(\mathfrak{F})$. Por lo tanto $[\mathfrak{b}(\mathfrak{F}), \mathfrak{b}(\mathfrak{F})] \subset \mathfrak{u}(\mathfrak{F})$ y $\mathfrak{u}(\mathfrak{F})$ es nilpotente i.e. $\mathfrak{b}(\mathfrak{F})$ es soluble.

2.1. Ejercicios.

1. *a*) Probar que toda álgebra nilpotente es soluble. Es cierta la recíproca?
   b) Si $\mathfrak{g}$ es un álgebra de Lie tal que $[\mathfrak{g},\mathfrak{g}]$ es nilpotente, entonces $\mathfrak{g}$ es soluble.
   c) Si $\mathfrak{g}/\mathrm{Cent}(\mathfrak{g})$ es nilpotente, entonces $\mathfrak{g}$ es nilpotente.

2.1.1. Teoremas básicos para álgebras nilpotentes.

Teorema 2.5. *(ENGEL)* $\mathfrak{g}$ *es nilpotente si y solo si* $adx \in End(\mathfrak{g})$ *es una transformación nilpotente, para todo* $x \in \mathfrak{g}$.

Para probar el teorema de Engel, consideramos el enunciado siguiente:

Teorema 2.6. *Sea* $\rho : \mathfrak{g} \to End(V)$ *una representación tal que* $\rho(x) \in End(V)$ *es una transformación nilpotente para todo* $x \in \mathfrak{g}$. *Entonces existe una bandera* $\mathfrak{F}$ *en* V *tal que* $\rho(\mathfrak{g}) \subseteq \mathfrak{u}(\mathfrak{F})$.

El teorema 2.6 es equivalente a

Teorema 2.7. *Sea* $\rho : \mathfrak{g} \to End(V)$ *una representación tal que* $\rho(x) \in End(V)$ *es una transformación nilpotente para todo* $x \in \mathfrak{g}$. *Entonces existe* $v \in V, v \neq 0$ *tal que* $\rho(x)v = 0$ *para cada* $x \in \mathfrak{g}$.

Demostración. En primer lugar, podemos suponer, reemplazando $\mathfrak{g}$ por su imagen, que $\mathfrak{g} \subseteq \mathfrak{gl}(V)$.

Se sigue que adx es nilpotente para cada $x \in \mathfrak{g}$, (se deja como ejercicio al lector).

Probaremos el teorema por inducción en la dimensión de $\mathfrak{g}$. Sea $\mathfrak{h}$ una subálgebra de $\mathfrak{g}$, $\dim\mathfrak{h} < \dim\mathfrak{g}$. Entonces $\mathrm{norm}_{\mathfrak{g}}(\mathfrak{h}) \supset \mathfrak{h}$. En efecto, se considera la representación de $\mathfrak{h}$ en $\mathfrak{g}/\mathfrak{h}$: por el teorema para $\mathfrak{h}$, existe $\bar{v} \in \mathfrak{g}/\mathfrak{h}$ tal que $x \cdot \bar{v} = 0$, para todo $x \in \mathfrak{h}$. Así, todo representante v de $\bar{v}$ pertenece a $\mathrm{norm}_{\mathfrak{g}}(\mathfrak{h}) - \mathfrak{h}$.

Deducimos que $\mathfrak{g}$ tiene un ideal de codimensión 1. Sea $\mathfrak{h}$ una subálgebra de $\mathfrak{g}$, maximal entre las subálgebras distintas de $\mathfrak{g}$. Como $\mathrm{norm}_{\mathfrak{g}}(\mathfrak{h})$ contiene propiamente a $\mathfrak{h}$, es igual a $\mathfrak{g}$, es decir $\mathfrak{h}$ es un ideal de $\mathfrak{g}$. Si $y \in \mathfrak{g} - \mathfrak{h}$, se sigue que $\mathfrak{h} \oplus k \cdot y$ es una subálgebra: luego $\mathfrak{h}$ tiene codimensión 1.

Fijemos entonces $\mathfrak{h}$ un ideal de codimensión 1, $y \in \mathfrak{g} - \mathfrak{h}$. Sea $W = \{v \in V : x \cdot v = 0, \text{ para todo } x \in \mathfrak{h}\}$. Por hipótesis inductiva, $W \neq 0$. Siendo $\mathfrak{h}$ un ideal, W es un $\mathfrak{g}$-submódulo de V. En particular, como y es nilpotente, existe $v \in W - 0$ tal que $y \cdot v = 0$, v es entonces anulado por todos los elementos de $\mathfrak{g}$. □

Demostración del Teorema 2.5. : Si $\mathfrak{g}$ es nilpotente, se sigue de (ii) en el Lema que adx es nilpotente para todo $x \in \mathfrak{g}$. Recíprocamente, si adx es nilpotente para todo $x \in \mathfrak{g}$, por el Teorema 2.6, existe una bandera $\mathfrak{F}$ en $\mathfrak{g}$ estable por la representación adjunta de $\mathfrak{g}$. Por (iii) del Lema, $\mathfrak{g}$ es nilpotente. □

2.2. Ejercicios.

1. Probar que el Teorema 2.6 es equivalente al Teorema 2.7
2. Probar que adx es nilpotente para cada $x \in \mathfrak{g}$.
3. Sea $\mathfrak{g}$ nilpotente, $\mathfrak{h} \subset \mathfrak{g}$ un ideal de $\mathfrak{g}$, entonces si $\mathfrak{h} \neq 0$, entonces $\mathfrak{h} \cap \mathrm{cent}(\mathfrak{g}) \neq 0$.

2.2.1. Teoremas básicos sobre álgebras solubles. En lo que sigue supondremos k algebraicamente cerrado y de característica cero.

Teorema 2.8. *(LIE) Si* $\mathfrak{g}$ *es soluble y* $\rho : \mathfrak{g} \to End(V)$ *es una representación de dimensión finita entonces existe una bandera* $\mathfrak{F}$ *en* V *tal que* $\rho(\mathfrak{g}) \subseteq \mathfrak{b}(\mathfrak{F})$.

El Teorema 2.8 es equivalente a

Teorema 2.9. *Si $\mathfrak{g}$ es soluble y $\rho : \mathfrak{g} \to End(V)$ es una representación irreducible de dimensión finita entonces V contiene una subrepresentación de dimensión 1, por lo tanto existe $v \neq 0$, $v \in V$ que es autovector simultaneo de $\rho(x)$ para todo x.*

2.3. Ejercicios.

1. Probar que el Teorema 2.8 es equivalente al Teorema 2.9.
2. *a)* Probar que si $\mathfrak{g}$ es soluble y $\rho : \mathfrak{g} \to \mathrm{End}(V)$ es una representación irreducible de dimensión finita entonces la $\dim V = 1$.
   b) Sea $\mathfrak{g}$ soluble entonces existe una bandera de ideales de $\mathfrak{g}$.

2.3.1. Criterio de Cartan. Sea V un espacio vectorial de dimensión finita. Sea $\mathfrak{g}$ una subálgebra de Lie de $\mathfrak{gl}(V)$. Supongamos $\mathrm{tr}(XY) = 0$ para todo $x \in \mathfrak{g}$, $y \in [\mathfrak{g}, \mathfrak{g}]$, entonces $\mathfrak{g}$ es soluble.

2.4. Ejercicios.

1. Probar la recíproca del criterio de Cartan.

3. Álgebras de Lie semisimple

3.1. Caracterización de las álgebras de Lie semisimples. Es esta sección $\mathfrak{g}$ será un álgebra de Lie de dimensión finita.

Definición 3.1. $\mathfrak{g}$ se dice *simple* si tiene exactamente dos ideales: $\mathfrak{g}$ y el 0, y si $[\mathfrak{g}, \mathfrak{g}] \neq 0$. Se dice *semisimple* si es suma directa de una familia de ideales simples.

Sea $\mathfrak{g}$ un álgebra de Lie. Si $x, y \in \mathfrak{g}$ se define $K_{\mathfrak{g}}(x, y) = \mathrm{tr}(\mathrm{ad}x\,\mathrm{ad}y)$. Entonces $K_{\mathfrak{g}}$ es una forma bilineal simétrica en $\mathfrak{g}$, llamada la *forma de Killing* de $\mathfrak{g}$, $K_{\mathfrak{g}}$ es también asociativa en el siguiente sentido $K_{\mathfrak{g}}([x, y], x) = K_{\mathfrak{g}}(x, [y, z])$.

Si $\mathfrak{a} \subset \mathfrak{g}$ se define el ortogonal de $\mathfrak{a}$ respecto de la forma de Killing como

$$\mathfrak{a}^{\perp} = \{x \in \mathfrak{g} : k(x, y) = 0, \forall y \in \mathfrak{a}\}.$$

Lema 3.2. *Si $\mathfrak{h}$ es un ideal de $\mathfrak{g}$ entonces $K_{\mathfrak{h}} = K_{\mathfrak{g}}|_{\mathfrak{h}\times\mathfrak{h}}$.*

Demostración. Se sabe que si $W \subset V$ es un subespacio vectorial de dimensión finita y ϕ es un endomorfismo de V tal que mapea V en W, entonces $\mathrm{tr}\phi = \mathrm{tr}\phi\,|_W$. Luego si $x, y \in \mathfrak{h}$ $(\mathrm{ad}x\,\mathrm{ad}y) : \mathfrak{g} \to \mathfrak{h}$, por lo tanto $K_{\mathfrak{g}}\,|_{\mathfrak{h}\times\mathfrak{h}} = \mathrm{tr}(\mathrm{ad}x\,\mathrm{ad}y) = \mathrm{tr}(\mathrm{ad}x\,\mathrm{ad}y)|_{\mathfrak{h}} = \mathrm{tr}(\mathrm{ad}_{\mathfrak{h}}x\,\mathrm{ad}_{\mathfrak{h}}y) = K_{\mathfrak{h}}(x, y)$. □

Teorema 3.3. *Son equivalentes*

(i) $\mathfrak{g}$ *es semisimple.*
(ii) $rad\,(\mathfrak{g}) = 0$, *donde* $rad\,(\mathfrak{g})$ *es el ideal soluble maximal de* $\mathfrak{g}$.
(iii) *La forma de Killing es no degenerada.*

Demostración. (i)⇒(ii): Supongamos que $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$, donde $\mathfrak{a}$ y $\mathfrak{b}$ son ideales de $\mathfrak{g}$. Es fácil ver que $\mathrm{rad}\,(\mathfrak{g}) = \mathrm{rad}\,(\mathfrak{a}) \oplus \mathrm{rad}\,(\mathfrak{b})$. De modo que basta ver que si $\mathfrak{g}$ es simple entonces $\mathrm{rad}\,(\mathfrak{g}) = 0$. Observemos que $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ pues $\mathfrak{g}$ es no abeliana. Por ende $\mathfrak{g}$ es no soluble y así su radical debe ser 0. (ii)⇒(iii): Sea $\mathfrak{s} = \mathfrak{g}^{\perp}$; es un ideal de $\mathfrak{g}$ pues es el núcleo de la aplicación de $\mathfrak{g}$ en $\mathfrak{g}^*$ inducida por la forma de Killing. Ahora $K_{\mathfrak{s}}(\mathfrak{s}, [\mathfrak{s}, \mathfrak{s}]) = K_{\mathfrak{g}}(\mathfrak{s}, [\mathfrak{s}, \mathfrak{s}]) = 0$. Sea $\tilde{\mathfrak{s}}$ la imagen de $\mathfrak{s}$ en $\mathrm{ad}(\mathfrak{g}) \subset \mathrm{End}(\mathfrak{g})$: $\tilde{\mathfrak{s}} = \mathfrak{s}/\mathrm{cent}(\mathfrak{g})$. Por el criterio de Cartan, $\tilde{\mathfrak{s}}$ (y por ende $\mathfrak{s}$) es soluble. De modo que $\mathfrak{s} \subset \mathrm{rad}\,(\mathfrak{g}) = 0$ y K es no degenerada. (iii)⇒ (i): Veamos primero que (iii)⇒(ii): si $\mathrm{rad}\,(\mathfrak{g}) \neq 0$, existe un ideal abeliano no nulo $\mathfrak{h}$ de $\mathfrak{g}$. Si $x \in \mathfrak{h}$, $y \in \mathfrak{g}$, es fácil ver que $(\mathrm{ad}x\,\mathrm{ad}y)^2 = 0$ por ende $\mathrm{tr}(\mathrm{ad}x\,\mathrm{ad}y) = 0$, i.e. $\mathfrak{h} \subset \mathfrak{g}^{\perp} = 0$, absurdo. Ahora probemos que (ii)⇒ (i): Si $\mathfrak{g}$

es simple no hay nada que probar. Sino sea $\mathfrak{h}$ un ideal propio de $\mathfrak{g}$, $\mathfrak{h} \cap \mathfrak{h}^{\perp}$ es un ideal de $\mathfrak{g}$, soluble por el criterio de Cartan, por lo tanto nulo. De modo que $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^{\perp}$ y podemos aplicar la hipótesis inductiva en la dimensión de $\mathfrak{g}$. □

3.1.1. Completa reducibilidad de representaciones.

Definición 3.4. Sea V un $\mathfrak{g}$-módulo, V se dice completamente reducible si es suma directa de $\mathfrak{g}$-submódulos irreducibles.

Teorema 3.5 (WEYL). *Si $\mathfrak{g}$ es semisimple, entonces todo $\mathfrak{g}$-módulo de dimensión finita es completamente reducible.*

Usando el Teorema de Weyl se prueba:

Teorema 3.6 (LEVI). *Toda álgebra de Lie de dimensión finita se descompone en suma directa de su radical y de una subálgebra de Lie semisimple.*

3.2. Ejercicios.

1. Probar que V es un $\mathfrak{g}$-módulo completamente reducible si cada $\mathfrak{g}$-submódulo $W \subset V$ tiene un complemento W', i.e. $V = W \oplus W'$.
2. *a*) Si $\mathfrak{h}$ es un ideal semisimple de un álgebra de Lie de dimensión finita entonces existe un único ideal $\mathfrak{b}$ de $\mathfrak{g}$ tal que $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{b}$.
   b) Si $\mathfrak{g}$ es semisimple entonces la aplicación adjunta es inyectiva.

Referencias

[1] J. Humphreys, *Introduction to Lie algebras and representation theory*, 3rd. printing, Springer-Verlag, (1980).
[2] N. Jacobson, *Lie algebras*, Wile Intersciences, New York-London, (1962).
[3] H. Samelson, *Notes on Lie algebras*, Springer-Verlag.
[4] J.-P. Serre, *Algébres de Lie semisimples complexes*, Benjamin, New York-Amsterdam, (1966).
[5] Varadarajan, *Algébres de Lie semisimples complexes*, Springer-Verlag, (1966).
[6] N. Bourbaki, *Groupes et algébres de Lie*, Chapitres 1, 2, 4, 5 et 6, Hermann, Paris, (1968).
[7] N. Andruskiewitsch, *Álgebras de Lie Semisimples y Representaciones de dimensión finita*, (1994).

FaMAF- Universidad Nacional de Córdoba, Medina Allende s/n-Ciudad Universitaria, CP:X5000HUA-Córdoba, Argentina.

E-mail address: vanemeinardi@gmail.com

TEORÍA COMBINATORIA DE NUDOS

LEANDRO VENDRAMIN

Resumen. Estas notas corresponden a un minicurso dictado en el Encuentro Nacional de Álgebra, elENA VII, La Falda, Córdoba, 2014.

Índice

Introducción

Robert Graves[1] cuenta una leyenda griega en la que un oráculo anunció a los habitantes de Frigia que reconocerían a su futuro rey al verlo llegar en una carreta de bueyes. Tiempo después, el pueblo reconoció en un campesino de nombre Gordias a su nuevo rey. En agradecimiento, Gordias ofreció a Zeus su carro y el yugo, que había atado con un nudo tan complicado que nadie podría desatar. Se dijo que el que fuera capaz de desatar el nudo de Gordias conquistaría Asia. Siglos después, Alejandro Magno tuvo que enfrentarse al reto de desatar ese nudo y, sin vacilación, deshizo el nudo al cortar la cuerda con su espada. Según se dice, esa misma noche, Zeus, con una fuerte tormenta, mostró su aprobación a la solución encontrada por Alejandro Magno.

La teoría de nudos, en principio, intenta entender los nudos que podríamos encontrarnos en la vida real.

El objetivo final de la teoría es obtener una clasificación completa de nudos a menos de deformaciones continuas. A simple vista, uno podría pensar entonces que la teoría de nudos es una divertida rama de la topología. Si bien esto es cierto, es necesario destacar que el estudio de los nudos se lleva a cabo gracias al uso de técnicas muy profundas que provienen de distintas ramas de la matemática como

[1] R. Graves, Los mitos griegos, Alianza Editorial, 1996.

la geometría, el álgebra y el análisis. La teoría de nudos tiene además muchas aplicaciones en otras ciencias como la biología, la física y la criptografía.

Invitamos al lector a que tome un pedazo de cuerda y haga un nudo tal como el que vemos a la izquierda en la figura 1. Si pegamos los extremos de esa cuerda obtendremos una cuerda anudada que no tiene extremos tal como la que vemos a la derecha en la figura 1.

FIGURA 1. Nudos.

¿Puede desatarse ese nudo? Después de varios minutos de experimentación se hace más o menos evidente que ese nudo podrá deshacerse solamente si nos permitimos cortar la cuerda. Sin embargo, si aceptamos soluciones como la que encontró Alejandro Magno, entonces todo nudo puede desatarse y el problema de clasificar nudos –que resulta ser muy poco interesante– queda resuelto: todo nudo es trivial.

¿Qué pasa si no está permitido cortar cuerdas? ¿Cómo podríamos demostrar matemáticamente que un nudo no puede desatarse? Para responder esta pregunta, primero es necesario describir matemáticamente un nudo de forma tal que la definición permita modelar con cierta fidelidad el fenómeno real de anudar una cuerda. Necesitamos además que nuestra definición excluya patologías matemáticas desagradables tales como hacer desaparecer un nudo al tirar indefinidamente de los extremos de la cuerda. Por último, necesitamos una definición precisa y acertada de lo que significa que dos nudos sean *equivalentes*, es decir iguales aunque se vean distintos. Una vez que tengamos estas cosas, habremos formulado matemáticamente el problema de estudiar nudos, y entonces, tal como se hace en muchas ramas de la matemática, podremos concentrarnos en estudiar nudos mediante el uso de invariantes.

Agradecimientos. Le agradezco a Edwin Clark, por haber leído estas notas y por los muchos comentarios que me ayudaron a mejorar este curso. Agradezco también a Jonathan Barmak, Marco Farinati, César Galindo, Juliana García Galofre, y Masahico Saito.

1. NUDOS

Esta primera sección está dedicada a los conceptos básicos de la teoría de nudos y al teorema de Reidemeister, que nos permite traducir el problema topológico de distinguir nudos al lenguaje de la combinatoria.

FIGURA 2. Kurt Reidemeister (1893–1971)

1.1. Un nudo (en $\mathbb{R}^3$) es una función inyectiva y continua $S^1 \to \mathbb{R}^3$, donde

$$S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}.$$

En el conjunto de nudos definimos la relación de equivalencia dada por **isotopía**. Diremos que los nudos dados por las funciones α y β son **equivalentes** si y sólo si existe una función continua $H\colon S^1 \times [0,1] \to \mathbb{R}^3$ tal que la función $H_t\colon z \mapsto H(z,t)$ es un nudo para todo $t \in [0,1]$, $H_0 = \alpha$ y $H_1 = \beta$.

1.2. Los nudos pueden ser dotados de una orientación. A lo largo de este trabajo siempre trabajaremos con nudos orientados.

1.3. Para evitar patologías desagradables, en este curso consideraremos únicamente nudos equivalentes a nudos dados por poligonales. Estos nudos se llaman **nudos mansos**. En la figura 3 mostramos un ejemplo de nudo salvaje (no manso). Para nosotros, un **nudo** siempre será un nudo manso.

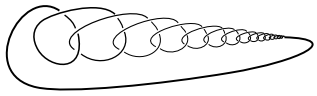


FIGURA 3. Un ejemplo de nudo salvaje.

1.4. En la práctica, nuestros nudos poligonales tendrán tantos segmentos que será casi imposible diferenciar a nuestra curva de una curva suave. En relación con esta observación, mencionamos el siguiente teorema:

Teorema. *Un nudo parametrizado por longitud de arco y de clase C^1 es manso.*

Demostración. La prueba de este resultado es muy técnica pero sólo utiliza conceptos básicos de cálculo avanzado. Para una demostración completa referimos a [6, Apéndice I]. □

1.5. Sea K un nudo. Consideremos la proyección de K en el plano dada por $\pi\colon (x,y,z) \mapsto (x,y,0)$. Un punto $p \in \pi(K)$ es un **punto múltiple** si $\pi^{-1}(p)$ contiene más de un punto de K. La **multiplicidad** de p se define como el cardinal del conjunto $\pi^{-1}(p) \cap K$. Una proyección de K en el plano se dice **genérica** tiene las siguientes propiedades: 1) hay finitos puntos de multiplicidad mayor a uno, 2) no hay puntos de multiplicidad mayor a dos, y 3) no hay puntos dobles donde uno de los puntos es un vértice. La figura 4 muestra algunos ejemplos de cruces no admitidos en una proyección genérica.

FIGURA 4. Cruces no admitidos en el diagrama de un nudo dado por una poligonal.

Un **diagrama** de K es una proyección genérica de K en el plano donde en cada **cruce** (punto de multiplicidad dos) se puede distinguir qué segmento pasa por arriba y qué segmento pasa por debajo. Para ilustrar esta situación, el segmento que pasa por debajo se dibuja cortado. Las componentes conexas del diagrama se llaman entonces **arcos**. Para diagramas de nudos orientados, agregaremos una flecha al diagrama que indique la orientación.

1.6. En virtud de simplificar la notación, seremos un poco imprecisos a la hora de hablar de nudos. Para nosotros, un nudo será una función inyectiva y continua $S^1 \to \mathbb{R}^3$ o una clase de equivalencia de tales funciones. Por más extraño que parezca, esto no causará confusión alguna.

1.7. Cualquier nudo equivalente a S^1 será considerado como el **nudo trivial**. En la práctica, no siempre es fácil reconocer la trivialidad de un nudo. En la figura 5 vemos tres proyecciones distintas del nudo trivial. Una proyección aún más curiosa del nudo trivial puede verse en la figura 6.

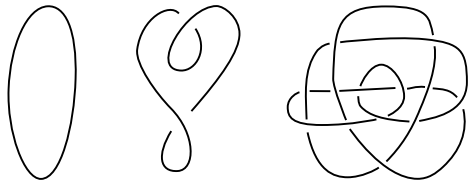


FIGURA 5. Tres proyecciones del nudo trivial.

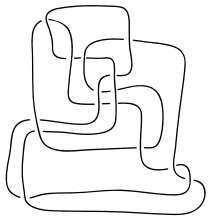


FIGURA 6. Una extraña proyección del nudo trivial descubierta por Morwen Thistlethwaite.

1.8. En 1926 Reidemeister vislumbró una forma combinatoria de chequear si dos nudos son equivalentes. Básicamente, dos diagramas representarán al mismo nudo si y sólo si puede pasarse de un diagrama al otro mediante una sucesión finita de ciertas transformaciones, $\mathcal{R}_1$, $\mathcal{R}_2$ y $\mathcal{R}_3$, llamadas **movimientos de Reidemeister**. Hay tres de estos movimientos: el primero se muestra en la figura 7, el segundo en la figura 8 y el tercero en la figura 9.

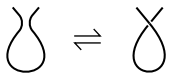


FIGURA 7. $\mathcal{R}_1$

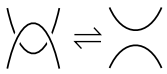


FIGURA 8. $\mathcal{R}_2$

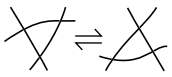


FIGURA 9. $\mathcal{R}_3$

Teorema (Reidemeister)**.** *Dos nudos son equivalentes si y sólo si sus diagramas están conectados por una sucesión finita de movimientos de Reidemeister.*

Demostración. Para la demostración referimos por ejemplo a [3, 1.14]. □

1.9. El teorema de Reidemeister también puede utilizarse para nudos orientados si se consideran todas las orientaciones posibles para los diagramas de las figuras 7–9.

1.10. El teorema de Reidemeister es el núcleo de la teoría combinatoria de nudos, ya que nos permite, por ejemplo, pensar que un nudo es una clase de equivalencia de diagramas, donde dos diagramas son equivalentes si y sólo si están conectados por una sucesión finita de movimientos de Reidemeister.

1.11. El **nudo trébol** es quizá el nudo más famoso. Una representación paramétrica de la curva que da este nudo es

$$\begin{aligned} x &= \sin(t) + 2\sin(2t) \\ y &= \cos(t) - 2\cos(2t) \\ z &= -\sin(3t) \end{aligned}$$

El nudo trébol, tal como lo vemos en la figura 10, se denota por el símbolo 3_1. Queda como ejercicio demostrar que los nudos de la figura 10 son equivalentes.

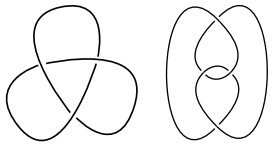


FIGURA 10. Dos proyecciones del nudo 3_1.

1.12. Otro nudo famoso es el nudo 4_1 o **nudo ocho**. Una representación paramétrica para la curva que da este nudo es

$$\begin{aligned} x &= (2 + \cos(2t))\cos(3t) \\ y &= (2 + \cos(2t))\sin(3t) \\ z &= \sin(4t) \end{aligned}$$

y una proyección puede verse en la figura 11.

FIGURA 11. El nudo 4_1.

1.13. La **imagen especular** de un nudo se obtiene al aplicarle al nudo la transformación $(x, y, z) \mapsto (x, y, -z)$. En la figura 12 vemos el nudo 3_1 y su imagen especular $m(3_1)$. En 8.7 demostraremos que los nudos 3_1 y $m(3_1)$ no son equivalentes.

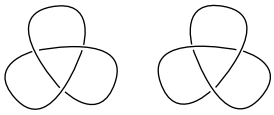


FIGURA 12. El nudo 3_1 (derecha) y su imagen especular $m(3_1)$ (izquierda).

1.14. Si K es un nudo, el **reverso** $r(K)$ de K es K como conjunto pero con la orientación opuesta. Los operadores r y m son involuciones en el espacio de nudos y generan un grupo isomorfo a $\mathbb{Z}_2 \times \mathbb{Z}_2$.

1.15. El nudo 3_1 es equivalente al nudo $r(3_1)$. El nudo 4_1 es **totalmente simétrico**, es decir: los nudos 4_1, $m(4_1)$, $r(4_1)$ y $rm(4_1)$ son todos equivalentes. El 9_{32} de la figura 13 es **totalmente asimétrico**, es decir: los nudos 9_{32}, $m(9_{32})$, $r(9_{32})$ y $rm(9_{32})$ son todos no equivalentes.

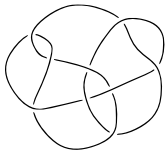
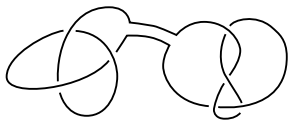


FIGURA 13. El nudo 9_{32} es totalmente asimétrico.

2. Composición de nudos y nudos primos

En esta sección definiremos una cierta forma de componer nudos de forma tal que el nudo trivial sea el neutro con respecto a esta operación.

2.1. Dados dos nudos K y L orientados podemos obtener un nuevo nudo con el siguiente procedimiento: Quitamos un pedacito de arco de cada una de las proyecciones de nuestros nudos y luego unimos los cuatro puntos finales obtenidos con dos nuevos arcos (¡es importante que hagamos esto sin agregar nuevos cruces!) tal como muestra la figura siguiente:



Esta operación se denomina **composición** de nudos. La composición de los nudos K y L se denota por $K\#L$. No es difícil demostrar que la composición de nudos es una operación asociativa y conmutativa y que el nudo trivial es el neutro de esta operación.

2.2. *Observación.* La composición de nudos solamente tiene sentido si se hace sobre nudos orientados.

2.3. Un nudo no trivial es **primo** si no puede descomponerse como la composición de otros nudos no triviales. Un nudo es **compuesto** si no es primo.

2.4. El problema de determinar si un nudo dado es primo es extremadamente difícil. La figura 14 contiene las proyecciones de los primeros nudos primos (salvo reverso e imagen especular) donde cada nudo tiene a lo sumo siete cruces.

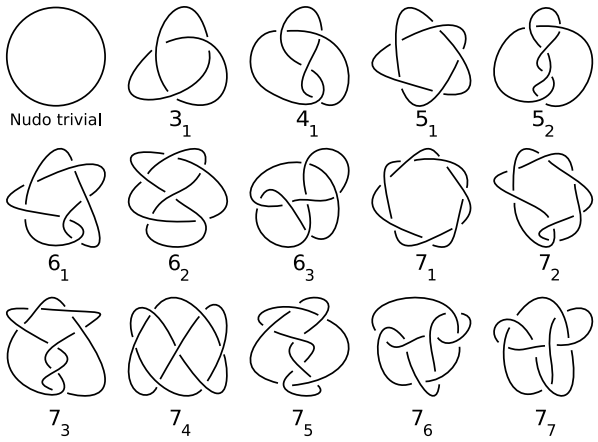


FIGURA 14. Algunos nudos primos.

2.5. **Ejemplo.** Consideremos el nudo que se forma al componer dos nudos 3_1. Este nudo se conoce como el **nudo de la abuela** (o *granny knot*, en inglés), se denota por $3_1\#3_1$, y se muestra a la izquierda en la figura 15.

El nudo compuesto formado por 3_1 y su imagen especular $m(3_1)$ se conoce como el **nudo cuadrado** (o *square knot*, en inglés), se denota por $3_1\#m(3_1)$, y se muestra a la derecha en la figura 15.

En 7.24 y 8.8 demostraremos que el nudo de la abuela y el nudo cuadrado no son equivalentes.

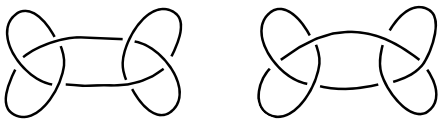


FIGURA 15. El nudo de la abuela (izquierda) y el nudo cuadrado (derecha) no son equivalentes.

2.6. Un teorema de H. Schubert establece que todo nudo puede expresarse en forma única como la composición de nudos primos [3, Cap. VII].

2.7. El género de un nudo y las superficies de Seifert permiten demostrar que el nudo trivial no puede escribirse como la composición de dos nudos no triviales, ver por ejemplo [1, §4.3]. ¡Este resultado nos dice que no es posible hacer dos nudos consecutivos en un pedacito de cuerda de forma tal que estos nudos se cancelen mutuamente!

2.8. Como la demostración del resultado mencionado en 2.7 es bastante difícil, nos gustaría tener a mano una prueba más sencilla. Es por eso que formulamos el siguiente problema:

Problema. ¿Existe algún invariante sencillo que permita demostrar que el nudo trivial no puede escribirse como la composición de dos nudos no triviales?

2.9. Puede construirse un invariante de nudos a partir de la cantidad de cruces que tienen los diagramas de un nudo. Para ser más precisos, definiremos el **número de cruces** $c(K)$ de un nudo K como el menor número de cruces que aparece en cualquier diagrama del nudo K. La siguiente conjetura lleva abierta más de cien años y nos recuerda lo poco que sabemos del número $c(K)$.

Conjetura. $c(K\#L) = c(K) + c(L)$.

3. COLOREOS

En esta sección definiremos algunos invariantes elementales y probaremos la no trivialidad de algunos nudos.

3.1. Supongamos que queremos mostrar que un cierto nudo no es trivial. ¿Qué invariante sencillo podríamos obtener a partir de los tres movimientos de Reidemeister? Responderemos esta pregunta al introducir el **coloreo con tres colores**. Fijemos un conjunto de tres colores, digamos $\{\mathsf{rojo}, \mathsf{verde}, \mathsf{azul}\}$. Una proyección de un nudo es **coloreable con tres colores** si cada arco de la proyección puede colorearse con uno de los tres colores de tal forma que en cada cruce se ven los tres colores elegidos o únicamente uno de los tres.

3.2. La cantidad de coloreos con tres colores da un invariante de nudos. Este resultado es un caso particular del teorema 3.8 que veremos más adelante.

FIGURA 16. Ralph Fox (1913–1973)

3.3. **Ejemplo.** La figura 17 nos muestra dos caras del mismo fenómeno: el nudo 3_1 tiene coloreos no triviales con tres colores y el nudo 4_1 no. Vemos entonces que el coloreo con tres colores nos permite distinguir el nudo 3_1 del nudo trivial y del nudo 4_1. Sin embargo, no nos permite determinar si el nudo 4_1 es trivial.

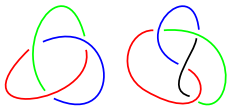


FIGURA 17. El núdo 3_1 coloreado con tres colores. El nudo 4_1 puede colorearse con tres colores solamente de forma trivial.

3.4. Vamos a profundizar un poco en la idea de colorear con tres colores. Supongamos que nuestros colores son los elementos de $\mathbb{Z}_3 = \{0, 1, 2\}$ y que K es un nudo con n cruces. Si etiquetamos los arcos del nudo K con los elementos de $\mathbb{Z}_3$ vemos que la condición 3.1 que define coloreos por tres colores puede traducirse en términos de la compatibilidad de un sistema de ecuaciones lineales que tiene una ecuación por cada cruce del diagrama. Para ser más precisos, la ecuación que corresponde al cruce de la figura 18 es $a + b + c = 0$, donde $a, b, c \in \mathbb{Z}_3$. Observemos que esta ecuación puede reescribirse como

$$2a - b - c = 0. \tag{3.5}$$

Cada coloreo del nudo K será una solución del sistema de ecuaciones formado por las ecuaciones (3.5). En particular, un coloreo no trivial será una solución que involucre todos los elementos de $\mathbb{Z}_3$.

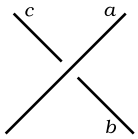


FIGURA 18. En este cruce la ecuación de coloreo es $2a - b - c = 0$.

3.6. **Ejemplo.** Si etiquetamos con $a, b, c \in \mathbb{Z}_3$ los arcos del diagrama del nudo 3_1 que vemos a la izquierda en la figura 17, el sistema de ecuaciones que resuelve el problema del coloreo con tres colores es el siguiente:

$$2a - b - c = 0, \qquad -a - b + 2c = 0, \qquad -a + 2b - c = 0,$$

La cantidad de coloreos con tres colores del nudo 3_1 es entonces la cantidad de vectores que tiene el núcleo de la **matriz de coloreos** del nudo:

$$C(3_1) = \begin{pmatrix} 2 & -1 & -1 \\ -1 & -1 & 2 \\ -1 & 2 & -1 \end{pmatrix} \in \mathbb{Z}_3^{3\times 3}.$$

Como esta matriz tiene rango uno, el núcleo es un espacio vectorial (sobre $\mathbb{Z}_3$) de dimensión dos. Esto implica que el núcleo tiene nueve elementos, y por lo tanto el nudo 3_1 tiene nueve coloreos con tres colores (de los cuales seis son no triviales).

3.7. En 1956 Fox definió una generalización del coloreo con tres colores. Sea $p > 2$ un número primo. Diremos que un nudo admite un **coloreo de Fox con p colores** si cada arco puede etiquetarse con un número de $\mathbb{Z}_p = \{0, \ldots, p-1\}$ de forma tal que en cada cruce como el que vemos en la figura 18 se cumple la equación

$$2a - b - c = 0,$$

donde $a, b, c \in \mathbb{Z}_p$. Tal como se hizo en el ejemplo 3.6, estudiar coloreos de Fox con p colores es equivalente a estudiar el núcleo de la matriz de coloreos vista como matriz con coeficientes en $\mathbb{Z}_p$.

3.8. **Teorema.** *Sea p un número primo. La cantidad de coloreos de Fox con p colores es un invariante de nudos.*

Demostración. Fijemos un diagrama y un coloreo de Fox del diagrama. Como vemos en la figura 19, el coloreo no se altera al aplicar el primer movimiento de Reidemeister ya que $b = 2a - a = a$.

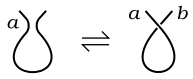


FIGURA 19. El coloreo de Fox es invariante bajo el primer movimiento de Reidemeister pues $b = 2a - a = a$.

La figura 20 nos muestra que el coloreo tampoco se altera al aplicar el segundo movimiento de Reidemeister pues $c = 2a - (2a - b) = b$.

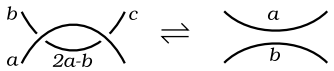


FIGURA 20. El coloreo de Fox es invariante bajo el segundo movimiento de Reidemeister pues $c = 2a - (2a - b) = b$.

Finalmente, la figura 21 prueba que el coloreo de Fox es invariante bajo el tercer movimiento de Reidemeister pues, como $d = 2a-(2b-c)$ y $d' = 2(2a-b)-(2b-c)$, se tiene que $d = d'$.

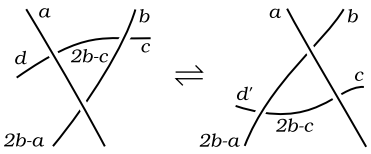


FIGURA 21. El coloreo de Fox es invariante bajo el tercer movimiento de Reidemeister pues $d = d'$.

Hemos probado entonces que existe una biyección entre los coloreos antes y después de aplicar los movimientos de Reidemeister. Luego, la cantidad de coloreos de Fox es un invariante de nudos. □

3.9. **Ejemplo.** Estudiemos algunos coloreos del nudo 4_1. Fijemos un número primo $p > 2$. Si suponemos que los arcos del nudo 4_1 están etiquetados con $a, b, c, d \in \mathbb{Z}_p$ tal como vemos en la figura 22, el sistema de ecuaciones que resuelve el problema del coloreo de Fox con p colores es

$$-a + 2b - d = 0, \quad -a - b + 2c = 0, \quad 2a - c - d = 0, \quad -b - c + 2d = 0. \tag{3.10}$$

Como vimos en el ejemplo 3.3, la matriz asociada al sistema (3.10) es lo que denominamos la matriz de coloreo del nudo 4_1:

$$C(4_1) = \begin{pmatrix} -1 & 2 & 0 & -1 \\ -1 & -1 & 2 & 0 \\ 2 & 0 & -1 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix} \in \mathbb{Z}_p^{4\times 4}.$$

Un cálculo elemental nos muestra que

$$\dim \ker C(4_1) = \begin{cases} 1 & \text{si } p = 3, \\ 2 & \text{si } p = 5. \end{cases}$$

Esto nos dice dos cosas: primero, que 4_1 no puede colorearse de forma no trivial con tres colores; y segundo, que 4_1 admite al menos un coloreo de Fox no trivial con cinco colores. Luego, el nudo 4_1 no es equivalente al nudo trivial.

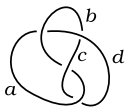


FIGURA 22. El nudo 4_1.

3.11. **Ejercicio.** Sean $p > 2$ un número primo y K un nudo con n cruces. Pruebe que la cantidad de coloreos de Fox con p colores que tiene K es p^m para algún $m \leq n$.

3.12. **Ejercicio.** El cuadro 1 muestra cuáles de los nudos de la figura 14 admiten coloreos de Fox no triviales para $p \in \{3, 4, 7, 11, 13, 17\}$. ¿Qué conclusiones puede obtener?

CUADRO 1. Algunos coloreos de Fox para los nudos de la figura 14.

	3	5	7	11	13	17
3_1	√					
4_1		√				
5_1		√				
5_2			√			
6_1	√					
6_2				√		
6_3					√	
7_1			√			
7_2				√		
7_3					√	
7_4	√	√				
7_5						√

4. El grupo fundamental de un nudo

En esta sección definiremos el grupo fundamental de un nudo y mostraremos que esta construcción da un buen invariante. Para una exposición detallada sobre las nociones básicas respecto del grupo fundamental de un nudo referimos a [6].

4.1. Se define el **grupo fundamental** $\pi_1(K)$ del nudo K como $\pi_1(\mathbb{R}^3 \setminus K)$.

4.2. Veamos cómo calcular el grupo fundamental de un nudo. Supongamos que K es un nudo con n cruces. Etiquetamos los arcos con las variables $a_1, a_2, a_3, \ldots$ Como se ve en la figura 23, el diagrama tendrá dos tipos de cruce: cruces positivos y cruces negativos. Por cada cruce χ del diagrama como el que vemos en la figura 23, consideramos la **relación de Wirtinger** $r_\chi = 1$, donde

$$r_\chi = a_i a_j a_i^{-1} a_k^{-1}. \tag{4.3}$$

A principios del siglo XX Wirtinger demostró que el grupo fundamental del nudo K es isomorfo al grupo dado por los generadores $a_1, a_2, a_3 \ldots$ y las relaciones de Wirtinger. Esta presentación del grupo fundamental se conoce como la **presentación de Wirtinger**.

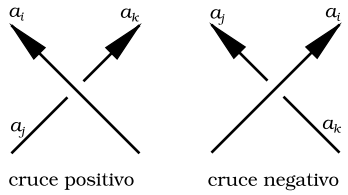


Figura 23. Una orientación en el nudo da dos tipos de cruce. En estos casos, la relación de Wirtinger es $a_i a_j a_i^{-1} = a_k$.

Teorema (Wirtinger)**.** *Sea K un nudo y supongamos que K tiene una proyección con n arcos y m cruces. Entonces*

$$\pi_1(K) \simeq \langle a_1, a_2, \ldots, a_n : r_1, r_2, \ldots, r_m \rangle, \tag{4.4}$$

donde las relaciones $r_1, \ldots, r_m$ están dadas por las fórmulas (4.3)*. La ecuación* (4.4) *simboliza el cociente grupo libre en $a_1, \ldots, a_n$ por el menor subgrupo normal que contiene a los elementos $r_1, \ldots, r_m$.*

Figura 24. Wilhelm Wirtinger (1865–1945)

4.5. **Ejercicio.** Pruebe que el grupo fundamental del nudo trivial es isomorfo a $\mathbb{Z}$.

FIGURA 25. Max Dehn (1878–1952)

4.6. En 1915 Dehn demostró que el grupo fundamental permite detectar la trivialidad de un nudo. Más precisamente, el teorema de Dehn establece que un nudo es trivial si y sólo si el grupo fundamental del nudo es isomorfo a $\mathbb{Z}$. Es importante remarcar que, en general, es muy difícil determinar si un grupo finitamente presentado es isomorfo al grupo trivial.

4.7. **Ejemplo.** Vamos a calcular el grupo fundamental del nudo 3_1 que vemos en la figura 26. Las relaciones de Wirtinger son

$$a_1 a_2 a_1^{-1} = a_3, \quad a_2 a_3 a_2^{-1} = a_1, \quad a_3 a_1 a_3^{-1} = a_2, \tag{4.8}$$

y entonces,

$$\pi_1(3_1) \simeq \langle a_1, a_2, a_3 : \ a_1 a_2 a_1^{-1} = a_3, \ a_2 a_3 a_2^{-1} = a_1, \ a_3 a_1 a_3^{-1} = a_2 \rangle.$$

Vamos a utilizar el grupo $\pi_1(3_1)$ para dar otra demostración de la no trivialidad de 3_1. Como existe un morfismo de grupos $\pi_1(3_1) \to \mathbb{S}_3$ tal que

$$a_1 \mapsto (12), \quad a_2 \mapsto (23), \quad a_3 \mapsto (13),$$

y el grupo simétrico $\mathbb{S}_3$ es un grupo no abeliano, se sigue que $\pi_1(3_1)$ es un grupo no abeliano. En particular $\pi_1(3_1) \not\simeq \mathbb{Z}$ y entonces el nudo 3_1 no es equivalente al nudo trivial.

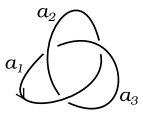


FIGURA 26. El grupo fundamental de 3_1 es isomorfo al grupo de trenzas $\mathbb{B}_3$.

4.9. **Ejercicio.** Sea $n \in \mathbb{N}_{\geq 2}$. Recordemos que el **grupo de trenzas** $\mathbb{B}_n$ se define como el grupo dado por los generadores $\sigma_1, \ldots, \sigma_{n-1}$ y las relaciones

$$\sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1} \qquad \text{para todo } i \in \{1, \ldots, n-2\},$$
$$\sigma_i \sigma_j = \sigma_j \sigma_i \qquad \text{para todo par } i, j \text{ tal que } |i-j| > 1.$$

Pruebe que $\pi_1(3_1) \simeq \mathbb{B}_3$.

4.10. Tietze fue el primero en calcular el grupo fundamental del nudo 3_1 en 1908. Ese mismo año conjeturó que dos nudos son equivalentes si y sólo si sus complementos en $\mathbb{R}_3$ son homeomorfos. Muchos años después, en 1986, Gordon y Luecke probaron esta afirmación [7]. Como consecuencia del teorema de Gordon y Luecke puede probarse que dos nudos primos son equivalentes si y sólo si sus grupos fundamentales son isomorfos.

FIGURA 27. Heinrich Tietze (1880–1964)

4.11. **Ejemplo.** Calculemos el grupo fundamental del nudo 4_1 de la figura 28.

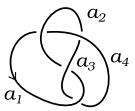


FIGURA 28. El grupo fundamental de 4_1 no es un grupo abeliano.

El diagrama tiene entonces dos cruces positivos y dos negativos. Las relaciones de Wirtinger son:

$$a_4 = a_1 a_3 a_1^{-1}, \quad a_2 = a_3 a_1 a_3^{-1}, \quad a_1 = a_2^{-1} a_4 a_2, \quad a_3 = a_4^{-1} a_2 a_4.$$

Si sustituimos $a_2 = a_3 a_1 a_3^{-1}$ en $a_4 = a_2 a_1 a_2^{-1}$, obtenemos $a_4 = a_3 a_1 a_3^{-1} a_1 a_3 a_1^{-1} a_3^{-1}$. Por lo tanto $a_4^{-1} a_3 a_4 = a_2$ es equivalente a $a_1 a_3 a_1^{-1} a_3 a_1 = a_3 a_1 a_3^{-1} a_1 a_3$. Como el resto de las relaciones de Wirtinger es consecuencia de $a_1 a_3 a_1^{-1} a_3 a_1 = a_3 a_1 a_3^{-1} a_1 a_3$,

$$\pi_1(4_1) \simeq \langle x, y \mid xyx^{-1}yx = yxy^{-1}xy \rangle.$$

Vamos a utilizar el grupo $\pi_1(4_1)$ para dar otra demostración de la no trivialidad del nudo 4_1. Consideremos el morfismo $\pi_1(4_1) \to \mathbf{SL}(2, \mathbb{Z}_3)$ dado por

$$x \mapsto \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix}, \quad y \mapsto \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}.$$

Como el grupo $\mathbf{SL}(2, \mathbb{Z}_3)$ no es abeliano, se sigue que $\pi_1(4_1)$ no es abeliano. En particular $\pi_1(4_1) \not\simeq \mathbb{Z}$ y entonces 4_1 no es el nudo trivial.

Ahora vamos a utilizar el $\pi_1(4_1)$ para demostrar que los nudos 3_1 y 4_1 no son equivalentes. Por lo visto en el ejemplo 4.7, si los grupos $\pi_1(4_1)$ y $\pi_1(3_1)$ fueran isomorfos, existiría un epimorfismo $\pi_1(4_1) \to \mathbb{S}_3$. Sin embargo, un cálculo directo muestra que ningún morfismo $\pi_1(4_1) \to \mathbb{S}_3$ es sobreyectivo.

4.12. **Ejercicio.** En la figura 15 vimos dos nudos compuestos: el nudo de la abuela y el nudo cuadrado. Pruebe que estos nudos tienen grupos fundamentales isomorfos.

4.13. *Observación.* Como veremos más adelante, el nudo de la abuela y el nudo cuadrado no son equivalentes. Luego, el grupo fundamental de un nudo es un buen invariante pero no es infalible, es decir: existen nudos no equivalentes con grupos fundamentales isomorfos.

5. QUANDLES

Sabemos que gracias a los movimientos de Reidermeister el problema de distinguir nudos puede formularse en términos de combinatoria. Hemos visto además dos invariantes de nudos: el grupo fundamental y el coloreo. En esta sección vamos a definir el quandle fundamental de un nudo y vamos a probar que es un invariante que generaliza al grupo fundamental y a los invariantes por coloreo.

5.1. Los quandles son estructuras algebraicas que modelan la conjugación en un grupo. La primera aparición de cierta familia de quandles fue en 1943 cuando Mituhisa Takasaki introdujo los quandles involutivos –los llamó *keis*– con el fin de entender reflexiones. En 1959, los matemáticos ingleses John Conway y Gavil Wraith, después de un interesante intercambio de cartas e ideas, definieron los *wracks*. La idea de Conway y Wraith es que un wrack (hoy llamado simplemente *rack*) es esencialmente lo que queda de un grupo una vez que uno olvida la multiplicación y se preocupa únicamente por la conjugación. En 1982, Joyce introdujo los *quandles* con el fin de producir invariantes de nudos. Inspirado en la presentación de Wirtinger del grupo fundamental de un nudo, Joyce construyó el *quandle fundamental* de un nudo y probó que esta construcción da un buen invariante de nudos no orientados. Ese mismo año, y en forma independiente, Sergei Matveev también introdujo los quandles –los llamó *grupoides distributivos*– y probó un resultado similar al de Joyce. En 1988, para estudiar ciertos aspectos de la teoría de singularidades de curvas, el matemático alemán Egbert Brieskorn introdujo una estructura equivalente a la de Conway y Wraith: los *automorphic sets*.

FIGURA 29. Algunos de los padres de la teoría de quandles. De izquierda a derecha: John Conway, Gavin Wraith, David Joyce, Egbert Brieskorn.

5.2. Un **quandle** es un par $(X, \triangleright)$, donde X es un conjunto no vacío con una operación binaria $\triangleright\colon X \times X \to X$ tal que

(5.3) la función $\varphi_x\colon X \to X,\ y \mapsto x \triangleright y$, es biyectiva para todo $x \in X$,

(5.4) $x \triangleright (y \triangleright z) = (x \triangleright y) \triangleright (x \triangleright z)$ para todo $x, y, z \in X$,

(5.5) $x \triangleright x = x$ para todo $x \in X$.

5.6. Sean X e Y dos quandles. Una función $f\colon X \to Y$ es un **morfismo** de quandles si $f(x \triangleright x') = f(x) \triangleright f(x')$ para todo $x, x' \in X$.

5.7. **Ejemplo.** Sea X un conjunto no vacío. Entonces X es un quandle con $x \triangleright y = y$ para todo $x, y \in X$. Este quandle se denomina **quandle trivial** sobre X.

5.8. **Ejemplo.** Sea G un grupo y X una clase de conjugación de G. Entonces X es un quandle con $x \triangleright y = xyx^{-1}$ para todo $x, y \in X$. El quandle asociado a la clase de conjugación de g en G se llama **quandle de conjugación** y se denota por g^G.

5.9. **Ejercicio.** Si X es un quandle de conjugación entonces

(5.10) $$x \triangleright y = y \Leftrightarrow y \triangleright x = x \quad \text{para todo } x, y \in X.$$

Encuentre un quandle de tres elementos que no cumpla con la condición (5.10).

5.11. **Ejercicio.** Sea $n \in \mathbb{N}$. Pruebe que $\mathbb{Z}_n$ es un quandle con la acción dada por $x \triangleright y = 2x - y$ para todo $x, y \in \mathbb{Z}_n$. Este quandle se denomina **quandle diedral** y se denota por $\mathbb{D}_n$.

5.12. Sea M un $\mathbb{Z}[t,t^{-1}]$-módulo a izquierda. Definimos el **quandle de Alexander** sobre M como el quandle dado por

$$x \triangleright y = (1-t)x + ty \qquad \text{para todo } x, y \in M. \tag{5.13}$$

Demostremos que la acción (5.13) define una estructura de quandle sobre M. Es evidente que para cada $x \in M$ la función $\varphi_x \colon y \mapsto (1-t)x + ty$ es inversible y la inversa φ_x^{-1} está dada por $y \mapsto (1-t^{-1})x + t^{-1}y$. Además $x \triangleright x = x$ para todo $x \in X$. Para demostrar la distributividad, tomamos $x, y, z \in M$ y calculamos

$$\begin{aligned}(x \triangleright y) \triangleright (x \triangleright z) &= ((1-t)x + ty) \triangleright ((1-t)x + tz) \\ &= (1-t)\left((1-t)x + ty\right) + t\left((1-t)x + tz\right) \\ &= (1-t)x + t(1-t)y + t^2 z \\ &= (1-t)x + t(y \triangleright z) \\ &= x \triangleright (y \triangleright z).\end{aligned}$$

FIGURA 30. James Alexander (1888–1971)

5.14. Veamos un caso particular de la contrucción que vimos en 5.12. Sea $\mathbb{F}_q$ el cuerpo de q elementos, donde q es una potencia de un número primo. Para cada $\alpha \in \mathbb{F}_q \setminus \{0\}$ definimos el **quandle de Alexander** de tipo (q, α) como el quandle sobre el cuerpo $\mathbb{F}_q$ dado por $x \triangleright y = (1-\alpha)x + \alpha y$ para todo $x, y \in \mathbb{F}_q$.

5.15. **Ejercicio.** Pruebe que el quandle $(123)^{\mathbb{A}_4}$ es un quandle de Alexander.

5.16. Así como puede calcularse el grupo fundamental de un nudo gracias a la presentación de Wirtinger, es posible considerar el quandle fundamental de un nudo. Supongamos que K es un nudo con n arcos y m cruces. Como hicimos en 4.2, etiquetamos los arcos de la proyección con las variables $a_1, a_2, a_3 \ldots$ En cada cruce χ como el que vemos en la figura 23 consideramos la relación

$$r_\chi : a_i \triangleright a_j = a_k. \tag{5.17}$$

El **quandle fundamental** del nudo K es el quandle

$$Q(K) = \langle a_1, a_2, \ldots, a_n : r_1, \ldots, r_m \rangle,$$

donde las relaciones $r_1, \ldots, r_m$ están dadas por las fórmulas (5.17).

5.18. **Ejemplo.** El quandle fundamental del nudo 3_1 de la figura 26 es

$$Q(3_1) = \langle a_1, a_2, a_3 : a_1 \triangleright a_2 = a_3,\ a_2 \triangleright a_3 = a_1,\ a_3 \triangleright a_1 = a_2 \rangle. \tag{5.19}$$

5.20. **Ejemplo.** En el ejemplo 4.11 calculamos el grupo fundamental del nudo 4_1. El quandle fundamental del nudo 4_1 de la figura 28, es

$$\begin{aligned}Q(4_1) = \langle a_1, \ldots, a_4 : a_1 \triangleright a_3 = a_4,\ a_3 \triangleright a_1 = a_2,\ a_2 \triangleright a_1 = a_4, \\ a_4 \triangleright a_3 = a_2 \rangle.\end{aligned} \tag{5.21}$$

Observemos que las dos primeras relaciones corresponden a cruces positivos y las dos últimas a cruces negativos.

5.22. *Observación.* El quandle fundamental de un nudo no necesariamente es un quandle finito.

5.23. No es difícil demostrar que el quandle fundamental de un nudo queda invariante por los movimientos de Reidemeister. La demostración es apenas más complicada que la del teorema 3.8. La diferencia está en que ahora habrá más diagramas de Reidemeister para chequear pues deben considerarse todas las orientaciones posibles.

Para el primer movimiento de Reidemeister deben chequearse los diagramas de la figura 31.

FIGURA 31. Las cuatro orientaciones posibles en el primer movimiento de Reidemeister.

Las posibles orientaciones que pueden aparecer en el segundo movimiento de Reidemeister se muestran en la figura 32.

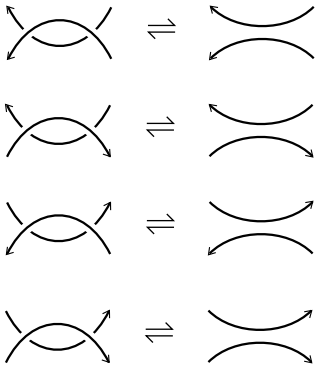


FIGURA 32. Las cuatro orientaciones posibles en el segundo movimiento de Reidemeister.

Finalmente, para el tercer movimiento hay ocho posibles diagramas. Por ejemplo, la condición a verificar en la figura 33 es

$$a \triangleright^{-1} (b \triangleright c) = (a \triangleright^{-1} b) \triangleright (a \triangleright^{-1} c),$$

que es consecuencia de la definición 5.2.

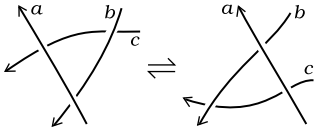


FIGURA 33. Una de las orientaciones posibles para el tercer movimiento de Reidemeister.

5.24. En 1982 Matveev probó el siguiente resultado: Si K y L son dos nudos, entonces $Q(K) \simeq Q(L)$ si y sólo si $K = L$ o $L = rm(K)$. Para la demostración referimos a [8].

5.25. Sea X un quandle. El **grupo envolvente** de X es el grupo

$$G_X = F_X/\langle xy = (x \triangleright y),\ x, y \in X\rangle,$$

donde F_X es el grupo libre con base en los elementos de X. En 1982 Joyce demostró que el grupo envolvente del quandle fundamental satisface $G_{Q(K)} \simeq \pi_1(K)$ para todo nudo K.

6. Ejemplos de quandle fundamentales

6.1. Consideremos el nudo 5_1 de la figura 34. Todos los cruces del diagrama son positivos y el quandle fundamental de 5_1 es

$$Q(5_1) = \langle a_1, \dots a_5 : a_1 \triangleright a_3 = a_4,\ a_4 \triangleright a_1 = a_2,\ a_2 \triangleright a_4 = a_5,\ a_5 \triangleright a_2 = a_3,\ a_3 \triangleright a_5 = a_1\rangle. \tag{6.2}$$

Consideremos ahora el nudo 5_2 de la figura 35. Todos los cruces del diagrama son positivos y entonces el quandle fundamental $Q(5_2)$ es

$$Q(5_2) = \langle a_1, \dots a_5 : a_1 \triangleright a_4 = a_5,\ a_4 \triangleright a_1 = a_2,\ a_2 \triangleright a_3 = a_4,\ a_5 \triangleright a_2 = a_3,\ a_3 \triangleright a_5 = a_1\rangle. \tag{6.3}$$

6.4. El fundamental $Q(6_1)$ del nudo 6_1 de la figura 36 es

$$Q(6_1) = \langle a_1, \dots, a_6 : a_1 \triangleright a_4 = a_3,\ a_3 \triangleright a_2 = a_1,\ a_2 \triangleright a_5 = a_6,\ a_5 \triangleright a_2 = a_3,\ a_4 \triangleright a_1 = a_6,\ a_6 \triangleright a_5 = a_4\rangle, \tag{6.5}$$

donde las ecuaciones $a_2 \triangleright a_5 = a_6$ y $a_5 \triangleright a_2 = a_3$ corresponden a los únicos cruces positivos. El quandle fundamental del nudo 6_2 tal como lo vemos en la figura 37 es

$$Q(6_2) = \langle a_1, \dots, a_6 : a_1 \triangleright a_4 = a_5,\ a_5 \triangleright a_3 = a_2,\ a_2 \triangleright a_6 = a_5,\ a_6 \triangleright a_3 = a_4,\ a_3 \triangleright a_1 = a_2,\ a_4 \triangleright a_6 = a_1\rangle, \tag{6.6}$$

donde las ecuaciones $a_5 \triangleright a_3 = a_2$ y $a_2 \triangleright a_6 = a_5$ corresponden a los únicos cruces positivos del diagrama. Por último, el quandle fundamental del nudo 6_3 de la figura 38 es

$$Q(6_3) = \langle a_1, \dots, a_6 : a_1 \triangleright a_5 = a_4,\ a_5 \triangleright a_2 = a_1,\ a_2 \triangleright a_3 = a_4,\ a_6 \triangleright a_2 = a_3,\ a_3 \triangleright a_6 = a_1,\ a_4 \triangleright a_6 = a_5\rangle, \tag{6.7}$$

donde las ecuaciones $a_2 \triangleright a_3 = a_4$, $a_6 \triangleright a_2 = a_3$ y $a_3 \triangleright a_6 = a_1$ corresponden a los únicos cruces positivos del diagrama.

6.8. Presentemos los quandles fundamentales de los nudos de la figuras 39–43. Un cálculo directo muestra que

$$Q(7_1) = \langle a_1, \dots, a_7 : a_1 \triangleright a_4 = a_5,\ a_5 \triangleright a_1 = a_2,\ a_2 \triangleright a_5 = a_6,\ a_6 \triangleright a_2 = a_3,\ a_3 \triangleright a_6 = a_7,\ a_7 \triangleright a_3 = a_4,\ a_4 \triangleright a_7 = a_1\rangle, \tag{6.9}$$

donde todos los cruces involucrados son positivos. Similarmente,

$$Q(7_2) = \langle a_1, \dots, a_7 : a_1 \triangleright a_4 = a_3,\ a_3 \triangleright a_2 = a_1,\ a_2 \triangleright a_6 = a_5,\ a_5 \triangleright a_7 = a_6,\ a_7 \triangleright a_5 = a_4,\ a_4 \triangleright a_1 = a_7,\ a_6 \triangleright a_3 = a_2\rangle, \tag{6.10}$$

donde todos los cruces son negativos. Para el nudo 7_3 tenemos

$$Q(7_3) = \langle a_1, \dots, a_7 : a_1 \triangleright a_5 = a_6,\ a_5 \triangleright a_1 = a_2,\ a_2 \triangleright a_4 = a_5,\ a_6 \triangleright a_2 = a_3,\ a_3 \triangleright a_6 = a_7,\ a_7 \triangleright a_3 = a_4,\ a_4 \triangleright a_7 = a_1\rangle, \tag{6.11}$$

donde todos los cruces son positivos. Para el nudo 7_4 se tiene

$$Q(7_4) = \langle a_1, \ldots, a_7 : a_5 \triangleright a_1 = a_2,\ a_2 \triangleright a_4 = a_5,\ a_4 \triangleright a_1 = a_7,\ a_6 \triangleright a_3 = a_4,\ a_3 \triangleright a_6 = a_7,\ a_7 \triangleright a_2 = a_3,\ a_1 \triangleright a_5 = a_6 \rangle, \tag{6.12}$$

donde la ecuación $a_4 \triangleright a_1 = a_7$ corresponde al único cruce negativo que tiene el diagrama. Para el nudo 7_5 tenemos

$$Q(7_5) = \langle a_1, \ldots, a_7 : a_1 \triangleright a_4 = a_3,\ a_5 \triangleright a_2 = a_1,\ a_2 \triangleright a_6 = a_5,\ a_7 \triangleright a_3 = a_2,\ a_3 \triangleright a_1 = a_7,\ a_4 \triangleright a_7 = a_6,\ a_6 \triangleright a_4 = a_5 \rangle, \tag{6.13}$$

donde la ecuación $a_6 \triangleright a_4 = a_5$ corresponde al único cruce negastivo del diagrama. Para el nudo 7_6 tenemos

$$Q(7_6) = \langle a_1, \ldots, a_7 : a_6 \triangleright a_2 = a_1,\ a_4 \triangleright a_3 = a_2,\ a_2 \triangleright a_6 = a_5,\ a_3 \triangleright a_1 = a_7,\ a_1 \triangleright a_2 = a_3,\ a_7 \triangleright a_4 = a_5,\ a_5 \triangleright a_6 = a_7 \rangle, \tag{6.14}$$

donde las relaciones $a_1 \triangleright a_2 = a_3$, $a_7 \triangleright a_4 = a_5$ y $a_5 \triangleright a_6 = a_7$ corresponden a los cruces positivos del diagrama. Finalmente, el quandle fundamental del nudo 7_7 es

$$Q(7_7) = \langle a_1, \ldots, a_7 : a_1 \triangleright a_4 = a_5,\ a_6 \triangleright a_2 = a_1,\ a_2 \triangleright a_4 = a_3,\ a_7 \triangleright a_2 = a_3,\ a_3 \triangleright a_6 = a_7,\ a_5 \triangleright a_7 = a_1,\ a_4 \triangleright a_6 = a_5 \rangle, \tag{6.15}$$

donde las relaciones $a_6 \triangleright a_2 = a_1$, $a_2 \triangleright a_4 = a_3$ y $a_4 \triangleright a_6 = a_5$ son las que corresponden a los cruces negativos del diagrama.

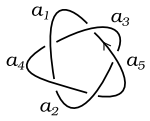


FIGURA 34. 5_1

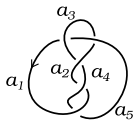


FIGURA 35. 5_2

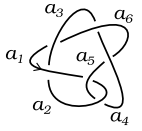


FIGURA 36. 6_1

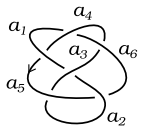


FIGURA 37. 6_2

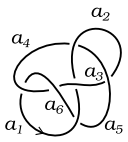


FIGURA 38. 6_3

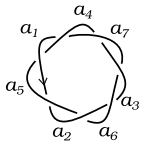


FIGURA 39. 7_1

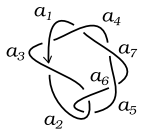


FIGURA 40. 7_2

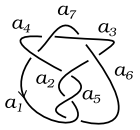


FIGURA 41. 7_3

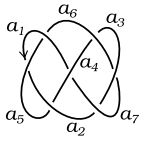


FIGURA 42. 7_4

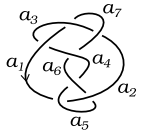


FIGURA 43. 7_5

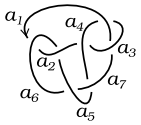


FIGURA 44. 7_6

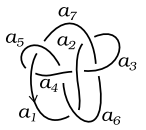


FIGURA 45. 7_7

6.16. Presentemos el quandle fundamental de la imagen especular del nudo 3_1 de la figura 46. El diagrama tiene tres cruces negativos y el quandle fundamental es

$$Q(m(3_1)) = \langle a_1, a_2, a_3 : a_3 \triangleright a_2 = a_1,\ a_2 \triangleright a_1 = a_3,\ a_1 \triangleright a_3 = a_2 \rangle.$$

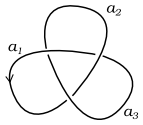


FIGURA 46. El nudo $m(3_1)$.

6.17. La figura 47 muestra el nudo de la abuela y la figura 48 el nudo cuadrado. Un cálculo sencillo muestra que el quandle fundamental del nudo de la abuela es

$$Q(3_1\#3_1) = \langle a_1, \ldots, a_6 : a_1 \triangleright a_5 = a_6,\ a_6 \triangleright a_1 = a_2,\ a_2 \triangleright a_3 = a_4, \\ a_4 \triangleright a_2 = a_3,\ a_3 \triangleright a_4 = a_5,\ a_5 \triangleright a_6 = a_1 \rangle, \tag{6.18}$$

donde todas las ecuaciones corresponden a cruces positivos, y que el quandle fundamental del nudo cuadrado es

$$Q(3_1\#m(3_1)) = \langle a_1, \ldots, a_6 : a_5 \triangleright a_6 = a_1,\ a_1 \triangleright a_5 = a_6,\ a_6 \triangleright a_1 = a_2, \\ a_4 \triangleright a_3 = a_2,\ a_3 \triangleright a_5 = a_4,\ a_5 \triangleright a_4 = a_3 \rangle, \tag{6.19}$$

donde las tres primeras ecuaciones corresponden a los cruces positivos y las tres últimas a los cruces negativos.

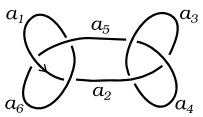


FIGURA 47. $3_1\#3_1$

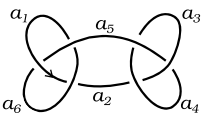


FIGURA 48. $3_1\#m(3_1)$

7. Coloreos generalizados

En esta sección utilizaremos el quandle fundamental para dar una generalización de los invariantes por coloreo.

7.1. Si etiquetamos los arcos de un nudo con los elementos de un quandle X de forma tal que en cada cruce se cumplan las relaciones que mencionamos en 5.16, o equivalentemente, en la figura 49, la cantidad de formas en que pueden ponerse esas etiquetas quedará invariante después de aplicar movimientos de Reidemeister. Esto nos permite "colorear" nudos de forma abstracta, donde ahora los "colores" son en realidad los elementos del quandle X.

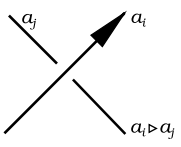


FIGURA 49. La regla para colorear con un quandle.

7.2. Un **coloreo** del nudo K con el quandle X es entonces un morfismo de quandles $Q(K) \to X$. Por lo dicho anteriormente, la cantidad de morfismos $Q(K) \to X$ es un invariante de nudos. Este invariante se denota por $\mathrm{Col}_X(K)$. Observemos que siempre existirán al menos $|X|$ coloreos del nudo K con el quandle X (los coloreos triviales).

7.3. **Ejemplo.** El coloreo por tres colores es en realidad el invariante asociado al quandle $\mathbb{D}_3$. El coloreo de Fox con p colores que vimos en 3.7 es en realidad el coloreo asociado al quandle $\mathbb{D}_p$.

7.4. **Ejemplo.** Veamos que el nudo 4_1 puede colorearse de forma no trivial con un quandle de Alexander. Consideremos el cuerpo

$$\mathbb{F}_4 = \mathbb{F}_2[\alpha]/(\alpha^2+\alpha+1) = \{0, 1, \alpha, \alpha+1\},$$

y sea X el quandle de Alexander de tipo $(4, \alpha)$. Vimos en el ejemplo 3.9 que los coloreos con X del diagrama de la figura 28 son las soluciones $(a, b, c, d) \in \mathbb{F}_4$ del sistema de ecuaciones

$$\begin{aligned} (1-\alpha)a + \alpha c = d, \quad (1-\alpha)b + \alpha a = d, \\ (1-\alpha)c + \alpha a = b, \quad (1-\alpha)d + \alpha c = b. \end{aligned} \tag{7.5}$$

Como $(a, b, c, d) = (0, 1, \alpha, \alpha+1)$ es una solución de (7.5), el nudo 4_1 admite entonces al menos un coloreo con X no trivial. Tenemos así otra demostración de la no trivialidad del nudo 4_1.

7.6. Los coloreos con quandles de Alexander se llaman **coloreos de Alexander**.

7.7. **Ejercicio.** Pruebe que el quandle de conjugación asociado a la matriz $\begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix}$ de $\mathbf{SL}(2, \mathbb{Z}_3)$ es isomorfo al quandle de Alexander de tipo $(4, \alpha)$ visto en el ejemplo 7.4. De alguna forma, habíamos utilizamos este quandle en 4.11.

7.8. **Ejemplo.** En 6.1 presentamos el quandle fundamental $Q(5_1)$. Un cálculo sencillo muestra que la función $Q(5_1) \to (12345)^{\mathbb{A}_5}$ definida por

$$a \mapsto (15432), \quad b \mapsto (12453), \quad c \mapsto (14352), \quad d \mapsto (15324), \quad e \mapsto (14523),$$

es un morfismo de quandles. Esta función nos permite "colorear" el nudo 5_1 con los 5-ciclos del grupo alternado $\mathbb{A}_5$.

7.9. **Ejemplo.** En 6.4 mostramos las relaciones que definen el quandle fundamental del nudo 6_3. Estas relaciones nos permiten demostrar, por ejemplo, que este nudo puede colorearse de forma no trivial con el quandle de Alexander de tipo $(7, 2)$. En efecto, si traducimos (6.7) a un sistema de ecuaciones obtenemos:

$$\begin{aligned} -a_1 + 2a_5 = a_4, \quad -a_5 + 2a_2 = a_1, \quad -a_2 + 2a_3 = a_4, \\ -a_6 + 2a_2 = a_3, \quad -a_3 + 2a_6 = a_1, \quad -a_4 + 2a_6 = a_5. \end{aligned} \tag{7.10}$$

Como $(a_1, a_2, a_3, a_4, a_5, a_6) = (1, 2, 0, 5, 3, 4)$ es una solución de (7.10), el nudo 6_3 puede colorearse de forma no trivial con el quandle de Alexander de tipo $(7, 2)$. Observemos que si X es el quandle de Alexander de tipo $(7, 2)$, entonces la función $Q(K) \to X$ dada por

$$a_1 \mapsto 1, \qquad a_2 \mapsto 2, \qquad a_3 \mapsto 0, \qquad a_4 \mapsto 5, \qquad a_5 \mapsto 3, \qquad a_6 \mapsto 6$$

es un morfismo de quandles.

7.11. **Ejercicio.** Pruebe que el quandle de Alexander de tipo $(4, \alpha)$ definido sobre el cuerpo $\mathbb{F}_4 = \mathbb{F}_2[\alpha]/(\alpha^2+\alpha+1) = \{0, 1, \alpha, \alpha+1\}$ permite distinguir los siguientes pares de nudos: a) 3_1 y 6_1; b) 4_1 y 5_1; c) 6_2 y 7_2; d) 6_3 y 7_3.

7.12. **Ejercicio.** Pruebe que con una clase de conjugación del grupo alternado $\mathbb{A}_5$ es posible distinguir los nudos 5_2 y 7_1.

7.13. **Ejercicio.** Utilice los resultados de los ejercicios 3.12, 7.11 y 7.12 y demuestre que todos los nudos de la figura 14 son no triviales y distintos.

7.14. Sean A un grupo abeliano (escrito multiplicativamente), X un quandle, y $f\colon X \times X \to A$ una función. Sobre el conjunto $X \times A$ definimos la operación

$$(x,a) \triangleright (y,b) = (x \triangleright y, bf(x,y)) \quad \text{para todo } (x,a),(y,b) \in X \times A. \tag{7.15}$$

Es fácil demostrar que la operación (7.15) define una estructura de quandle sobre $X \times A$ si y sólo si:

$$f(x,x) = 1, \qquad \text{para todo } x \in X, \tag{7.16}$$

$$f(x,z)f(x \triangleright y, x \triangleright z) = f(y,z)f(x, y \triangleright z) \qquad \text{para todo } x,y,z \in X. \tag{7.17}$$

Como ejemplo, demostremos la distributividad. Sean $x,y,z \in X$ y $a,b,c \in A$. Un cálculo directo nos dice que

$$\begin{aligned}(x,a) \triangleright ((y,b) \triangleright (c,z)) &= (x,a) \triangleright (y \triangleright z, cf(y,z)) \\ &= (x \triangleright (y \triangleright z), cf(y,z)f(x, y \triangleright z)),\end{aligned}$$

y, por otro lado,

$$\begin{aligned}&((x,a) \triangleright (y,b)) \triangleright ((x,a) \triangleright (c,z)) \\ &\qquad = (x \triangleright y, bf(x,y)) \triangleright (x \triangleright z, cf(x,z)) \\ &\qquad = ((x \triangleright y) \triangleright (x \triangleright z)), cf(x,z)f(x \triangleright y, x \triangleright z)).\end{aligned}$$

Como X es un quandle, tenemos entonces que la condición (7.17) es equivalente a la distributividad de la operación binaria en $X \times A$.

7.18. Una función $f : X \times X \to A$ que satisface las condiciones (7.16) y (7.17) se llama 2-**cociclo del quandle** X con coeficientes en A.

7.19. El quandle obtenido en 7.14 se llama **extensión abeliana** de X por el grupo abeliano A y el 2-cociclo f, y se denota por $X \times_f A$.

7.20. **Ejemplo.** Supongamos que $\mathbb{F}_4 = \mathbb{F}_2[\alpha]/(\alpha^2+\alpha+1) = \{0,1,\alpha,\alpha+1\}$ y sea X el quandle de Alexander de tipo $(4,\alpha)$. Sea $A = \langle\sigma\rangle = \{1,\sigma\}$ el grupo cíclico de orden dos. La función $f\colon X \times X \to A$ dada por

$$f(x,y) = \begin{cases} 1 & \text{si } x = y \text{ o } x = 1 \text{ o } y = 1, \\ \sigma & \text{en otro caso.} \end{cases}$$

es un 2-cociclo de X con coeficientes en el grupo abeliano A.

7.21. **Ejemplo.** Sea X el quandle $(1234)^{\mathbb{S}_4}$ y sea $A = \langle\sigma\rangle = \{1,\sigma,\sigma^2,\sigma^3\}$ el grupo cíclico de orden cuatro (escrito multiplicativamente). La función $f\colon X \times X \to A$ dada por la tabla

f	(1234)	(1432)	(1342)	(1243)	(1324)	(1423)
(1234)	1	σ	σ^2	σ^2	σ	σ^3
(1432)	σ	1	σ^2	1	σ^3	σ^3
(1342)	σ^2	σ	1	σ	σ^2	σ^3
(1243)	σ^3	σ^2	σ	1	1	σ^3
(1324)	σ	σ	σ	σ	1	σ
(1423)	1	1	1	1	σ	1

es un 2-cociclo de X con coeficientes en A.

7.22. El 2-cociclo $f\colon X \times X \to A$ es un **coborde** si existe $\gamma\colon X \to A$ tal que $f(x,y) = \gamma(x \triangleright y)\gamma(y)^{-1}$ para todo $x, y \in X$. Dos 2-cociclos f y g son **cohomólogos** (o equivalentes) si existe $\gamma\colon X \to A$ tal que

$$f(x,y) = \gamma(x \triangleright y)g(x,y)\gamma(y)^{-1}$$

para todo $x, y \in X$.

7.23. Como vimos en 7.14, cada 2-cociclo de un quandle X nos permite definir una extensión abeliana de X. Estos 2-cociclos son en realidad 2-cociclos en una teoría de cohomología de quandles [5]. Tal como pasa en la teoría de grupos, las clases de equivalencia de extensiones abelianas del quandle X por el grupo abeliano A están en correspondencia biyectiva con las clases de equivalencia de 2-cociclos de X con coeficientes en A. La teoría de extensiones abelianas de quandles tiene además aplicaciones a la teoría de nudos [4]. Para más información sobre la teorías de extensiones y (co)homologías de quandles referimos a [2].

7.24. **Ejemplo.** Recordemos el nudo de la abuela y el nudo cuadrado de la figura 15. En el ejercicio 4.12 vimos que estos nudos tienen grupos fundamentales isomorfos. Sin embargo, como veremos a continuación, estos nudos no son equivalentes.

Sean X el quandle $(1234)^{\mathbb{S}_4}$ y f el 2-cociclo de X que vimos en el ejemplo 7.21. Teniendo en mente 7.14, consideremos la extensión $X \times_f A$ dada por

$$(x, \sigma^i) \triangleright (y, \sigma^j) = (x \triangleright y, \sigma^j f(x,y)) \quad \text{para todo } x, y \in X,\ i, j \in \{0, \dots 3\}.$$

Gracias a una sugerencia de Edwin Clark, usaremos el quandle $X \times_f A$ para distinguir el nudo de la abuela del nudo cuadrado. Un cálculo computacional nos muestra que para el nudo de la abuela se tiene

$$\mathrm{Col}_{X \times_f A}(3_1 \# 3_1) = 24, \tag{7.25}$$

y que para el nudo cuadrado, en cambio, se tiene

$$\mathrm{Col}_{X \times_f A}(3_1 \# m(3_1)) = 408. \tag{7.26}$$

Vimos en el ejercicio 4.12 que el nudo de la abuela es trivial si y sólo si el nudo cuadrado lo es. Esta observación y las fórmulas (7.25) y (7.26) implican que estos nudos son no triviales y distintos.

8. Invariantes por 2-cociclos

A fines del siglo XX, S. Carter, D. Jelsovsky, S. Kamada, L. Langford y M. Saito anunciaron la construcción de un nuevo invariante de nudos: el invariante por 2-cociclos. En esta sección introduciremos los invariantes dados por 2-cociclos y calcularemos algunos ejemplos.

8.1. Fijemos un grupo abeliano A (escrito multiplicativamente) y un nudo K. Sean X un quandle finito, $\mathcal{C}\colon Q(K) \to X$ un coloreo de K y $f\colon X \times X \to A$ un 2-cociclo de X con coeficientes en A. En cada cruce como el que vemos en la figura 23 se define el **peso de Boltzmann** $\omega_f(\mathcal{C}, \chi)$ (con respecto al coloreo $\mathcal{C}$, al 2-cociclo f y al cruce χ) como el elemento de A dado por la expresión

$$\omega_f(\mathcal{C}, \chi) = f(a_i, a_j)^{\mathrm{signo}(\chi)}.$$

La **función de partición** $\Phi_{X,f}(K)$ del nudo K (asociada al quandle X y al 2-cociclo f) es la expresión

$$\Phi_{X,f}(K) = \sum_{\mathcal{C}} \prod_{\chi} \omega_f(\mathcal{C}, \chi), \tag{8.2}$$

donde el producto se toma sobre todos los cruces χ que tiene el diagrama del nudo K y la suma se toma sobre todos los coloreos $\mathcal{C}$ de K dados por el quandle X. La fórmula (8.2) define un elemento de $\mathbb{Z}[A]$, el anillo de grupo de A.

8.3. **Teorema.** *La función de partición $\Phi_{X,f}$ es un invariante de nudos.*

Demostración. Tenemos que demostrar que el producto de los pesos de Boltzmann es invariante bajo las versiones orientadas de los movimientos de Reidemeister.

Consideremos el primer movimiento de Reidemeister. Tal como muestra la figura 31, hay dos orientaciones posibles para tener en cuenta. En ambos casos, si suponemos que estas cuerdas llevan la etiqueta $a \in X$, entonces en el único cruce χ que tiene el diagrama tendremos el valor $f(a,a)^{\text{signo}(\chi)}$. Como f es un 2-cociclo, $f(a,a) = 1$. Luego, el primer movimiento de Reidemeister deja invariante al producto de los pesos de Boltzmann.

Consideremos ahora el segundo movimiento. Aquí tenemos cuatro posibles diagramas orientados, similares a los que se ve en la figura 32. Si etiquetamos la cuerda que pasa por arriba con $a \in X$ y la cuerda entrante que pasa por debajo con $b \in X$ entonces, como para ambos diagramas tenemos un cruce positivo y uno negativo, el producto de los pesos de Boltzmann es $f(a,b)f(a,b)^{-1} = 1$. Luego, el segundo movimiento de Reidemeister también deja invariante al producto de los pesos de Boltzmann.

Para finalizar, tenemos que demostrar que el tercer movimiento de Reidemeister deja invariante al producto de los pesos de Boltzmann. Como vimos en 5.23, hay ocho casos para chequear. Hagamos como ejemplo el caso que se corresponde con la figura 50 y dejemos el resto como ejercicio.

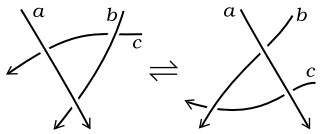


FIGURA 50. Otra de las orientaciones posibles para el tercer movimiento de Reidemeister.

Si calculamos el producto de los pesos de Boltzmann sobre los tres cruces que tienen los diagramas de la figura 50 vemos que este producto es invariante por el tercer movimiento de Reidemeister si y sólo si

$$f(a, b \triangleright c) f(b,c) f(a,b) = f(a,b) f(a \triangleright b, a \triangleright c) f(a,c).$$

Como A es un grupo abeliano, al cancelar $f(a,b)$ en ambos miembros, obtenemos el resultado deseado. □

8.4. Los invariantes por quandles y 2-cociclos extienden a los invariantes por coloreo con quandles. Más precisamente, si f es un coborde entonces $\Phi_{X,f}(K) = \text{Col}_X(K)$ para todo nudo K.

8.5. La afirmación hecha en 8.4 puede generalizarse: si f y g son cohomólogos entonces $\Phi_{X,f} = \Phi_{X,g}$. Para las demostración referimos a [5, Proposición 4.5].

8.6. **Ejemplo.** Sea X el quandle de Alexander de tipo $(4,\alpha)$ y f el 2-cociclo que vimos en el ejemplo 7.20. Un cálculo computacional nos permite calcular $\Phi_{X,f}$ para los nudos de la figura 14:

$$\Phi_{X,f}(K) = \begin{cases} 4 + 12\sigma & \text{si } K \in \{3_1, 4_1, 7_2, 7_3\}, \\ 4 & \text{en otro caso.} \end{cases}$$

8.7. **Ejemplo.** En este ejemplo vamos a distinguir el nudo 3_1 de su imagen especular $m(3_1)$, ver figura 12. Consideremos el quandle X y f el 2-cociclo de X que vimos en el ejemplo 7.21. Un cálculo directo muestra que

$$\Phi_{X,f}(3_1) = 6 + 24\sigma^3, \quad \Phi_{X,f}(m(3_1)) = 6 + 24\sigma.$$

Esto nos dice que los nudos 3_1 y $m(3_1)$ no son equivalentes.

8.8. **Ejemplo.** Como hicimos en el ejemplo anterior, vamos a utilizar el quandle X y f el 2-cociclo de X que vimos en el ejemplo 7.21. El invariante dado por X y el 2-cociclo f para el nudo de la abuela es

$$\Phi_{X,f}(3_1\#3_1) = 6 + 48\sigma + 96\sigma^2,$$

mientras que para el nudo cuadrado es

$$\Phi_{X,f}(3_1\#m(3_1)) = 102 + 24\sigma + 24\sigma^3.$$

Esto nos muestra que el nudo de la abuela no es equivalente al nudo cuadrado.

Referencias

[1] C. C. Adams. *The knot book*. American Mathematical Society, Providence, RI, 2004. An elementary introduction to the mathematical theory of knots, Revised reprint of the 1994 original.

[2] N. Andruskiewitsch and M. Graña. From racks to pointed Hopf algebras. *Adv. Math.*, 178(2):177–243, 2003.

[3] G. Burde, H. Zieschang, and M. Heusener. *Knots*, volume 5 of *De Gruyter Studies in Mathematics*. De Gruyter, Berlin, extended edition, 2014.

[4] J. S. Carter, M. Elhamdadi, M. A. Nikiforou, and M. Saito. Extensions of quandles and cocycle knot invariants. *J. Knot Theory Ramifications*, 12(6):725–738, 2003.

[5] J. S. Carter, D. Jelsovsky, S. Kamada, L. Langford, and M. Saito. Quandle cohomology and state-sum invariants of knotted curves and surfaces. *Trans. Amer. Math. Soc.*, 355(10):3947–3989, 2003.

[6] R. H. Crowell and R. H. Fox. *Introduction to knot theory*. Springer-Verlag, New York-Heidelberg, 1977. Reprint of the 1963 original, Graduate Texts in Mathematics, No. 57.

[7] C. M. Gordon and J. Luecke. Knots are determined by their complements. *J. Amer. Math. Soc.*, 2(2):371–415, 1989.

[8] S. V. Matveev. Distributive groupoids in knot theory. *Mat. Sb. (N.S.)*, 119(161)(1):78–88, 160, 1982.

INTRODUCCIÓN A LA TEORÍA DE FORMAS MODULARES

MARTÍN MEREB

Resumen. Tomando como motivación el problema de la representación de un número como suma de cuadrados, introduciremos las formas modulares y la geometría de las superficies de Riemann subyacentes. Probaremos varias identidades entre funciones aritméticas provistos de resultados de dimensión finita y ciertos operadores lineales diagonalizables. Serán necesarios algunos resultados de análisis complejo y topología.

Índice

Introducción.

0.1. Problema: Sumas de cuadrados. Sea $r_k(n)$ la cantidad de soluciones enteras de la ecuación

$$n = x_1^2 + \ldots + x_k^2$$

con n y k enteros positivos fijos. Buscamos expresiones sencillas para $r_k(n)$. Para simplificar esta exposición nos restringiremos a los k pares.

Date: 25 de Julio de 2014.

Veremos cómo encontrar de manera sistemática fórmulas como

$$r_2(n) = 2\left(1+\left(\frac{-1}{n}\right)\right)\sum_{d|n}\left(\frac{-1}{d}\right) \tag{0.1}$$

$$r_4(n) = 8\sum_{d|n,4\nmid d} d \tag{0.2}$$

$$r_6(n) = \left(\left(\frac{-1}{n}\right)2^{2\nu+4}-4\right)\sum_{d|n}\left(\frac{-1}{d}\right)d^2 \tag{0.3}$$

y

$$r_8(n) = \begin{cases} \sum_{d|n} d^3 & \text{si } n \text{ es par} \\ \sum_{d|n} d^3 - 2\sum_{d|g} d^3 & \text{si } n \text{ es impar} \end{cases} \tag{0.4}$$

donde $n = 2^\nu g$ con g impar y $\nu \geq 0$, y el símbolo de Jacobi $\left(\frac{-1}{n}\right) = (-1)^{(n-1)/2}$ si n es impar y $\left(\frac{-1}{n}\right) = 0$ para n par.

Considerando funciones generatrices, para $\tau \in \mathbb{C}$ definimos

$$\Theta(\tau) = \sum_{n=-\infty}^{\infty} q^{n^2} \tag{0.5}$$

donde $q = \exp(2\pi i\tau)$. Esta serie converge uniforme y absolutamente sobre compactos del semiplano de Poincaré

$$\mathcal{H} = \{\tau \in \mathbb{C}/\mathrm{Im}(\tau) > 0\}.$$

Se tiene pues que

$$(\Theta(\tau))^k = \sum_{n\geq 0} r_k(n) q^n$$

transformando el problema a hallar nuevas expresiones para Θ y extraer luego los coeficientes.

Sucede que las funciones Θ^k satisfacen ciertas ecuaciones funcionales que limitan enormemente la dimensión del espacio donde viven, permitiendo escribirlas como combinación lineal de otras cuyos coeficientes son fáciles de calcular.

Ejercicios.

1. Probar que $\Theta(\tau) + \Theta(\tau + 1/2) = 2\Theta(4\tau)$.
2. Probar
$$\sum_{n=0}^{\infty} p(n)q^n = \prod_{k=1}^{\infty}\left(\frac{1}{1-q^k}\right)$$
donde $p(n)$ es el número de maneras de escribir a n como suma de enteros positivos, sin importar el orden de los mismos.

0.2. Un poco de Fourier. Para $f \in L^1(\mathbb{R})$ se define la *transformada de Fourier* como

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)\, e^{-2\pi i x\xi}\, \mathrm{d}\, x$$

y la transformada inversa de f como

$$\check{f}(x) = \int_{-\infty}^{\infty} f(\xi)\, e^{2\pi i \xi x}\, \mathrm{d}\, \xi.$$

Cuando f y $\hat{f}$ pertenecen a $L^1(\mathbb{R})$ se tiene que $\check{\hat{f}} = f$ en casi todo punto.

Afirmación 0.1. *Entre las propiedades de la transformada podemos mencionar*

- *Si* $g(x) = f(x+a)$, $\hat{g}(\xi) = e^{2\pi i a\xi}\hat{f}(\xi)$.
- *Si* $g(x) = f(bx)$ *con* $b > 0$, *entonces* $\hat{g}(\xi) = \frac{1}{b}\hat{f}(\xi/b)$.
- *Sea* $f(x) = \exp(-\pi x^2)$, *entonces* $\hat{f} = f$.

La fórmula de Poisson (ver ejercicio 0.2 abajo)

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(n) \tag{0.6}$$

aplicada a $g(x) = f(\sqrt{t}x)$, nos da

$$\Theta\left(\frac{-1}{4\tau}\right) = \sqrt{\frac{2\tau}{i}}\Theta(\tau)$$

para todo $\tau \in H$. Es esta propiedad, junto a $\Theta(\tau) = \Theta(\tau+1)$, la que limitará las dimensiones.

Para más detalles sobre transformadas y series de Fourier, ver [11].

Ejercicios.

- Sea $f \in L^1(\mathbb{R})$. Probar que $g(x) := \sum_{n\in\mathbb{Z}} f(n+x)$ es periódica y pertenece a $L^1([0,1])$.
- Sea $\sum_{n\in\mathbb{Z}} c_n e^{2\pi i n x}$ el desarrollo en serie de Fourier de $g(x)$. Probar que $c_n = \hat{f}(n)$.
- Probar que si $f(x) = O(|x|^2)$ y $\hat{f}(\xi) = O(|\xi|^2)$ entonces

$$\sum_{n\in\mathbb{Z}} f(n+x) = \sum_{n\in\mathbb{Z}} \hat{f}(n) e^{2\pi i n x}.$$

0.3. La curva modular $X(1)$. Estudiemos una ecuación funcional mas simple

$$f(\tau) = f(\tau+1) \text{ y } \ f(\tau) = f(-1/\tau). \tag{0.7}$$

Para $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{Gl}_2(\mathbb{R})$ notamos

$$\gamma\tau := \frac{a\tau + b}{c\tau + d}.$$

Dado que $\mathrm{Sl}_2(\mathbb{Z})$ está generado por $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ y por $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, las ecuaciones (0.7) equivalen a

$$f(\gamma\tau) = f(\tau),\ \forall\, \gamma \in \mathrm{Sl}_2(\mathbb{Z}). \tag{0.8}$$

Lo que nos lleva a estudiar el espacio $Y(1) := \mathrm{Sl}_2(\mathbb{Z})\backslash\mathcal{H}$, junto con sus funciones meromorfas.

Un dominio fundamental estándar para la acción de $\mathrm{Sl}_2(\mathbb{Z})$ en $\mathcal{H}$ es

$$D = \left\{\tau \in \mathcal{H}/\ \frac{-1}{2} \le \Re(\tau) \le \frac{1}{2},\ |\tau| \ge 1\right\}. \tag{0.9}$$

Es conveniente considerar una compactificación $X(1)$ de $Y(1)$ dada por

$$X(1) := \mathrm{Sl}_2(\mathbb{Z})\backslash\overline{\mathcal{H}}$$

donde $\overline{\mathcal{H}} := \mathcal{H} \cup \mathbb{Q} \cup \{\infty\}$, denominada *curva modular*. Consideramos en $\overline{\mathcal{H}}$ la topología de las bolas tangentes, donde $\mathcal{H} \subseteq \overline{\mathcal{H}}$ es abierto, una base de entornos para $x \in \mathbb{Q}$ es

$$U_x(r) = \left\{\tau \in \overline{\mathcal{H}}/\ |x + ir - \tau| \le r\right\}$$

y una para ∞ es

$$U_\infty(R) = \left\{\tau \in \mathcal{H}/\ \mathrm{Im}(\tau) \ge R\right\} \cup \{\infty\}.$$

Observemos que con esta topología $\overline{D} = D \cup \{\infty\}$ resulta compacto y por lo tanto $X(1)$ también.

Ejercicios.

1. Probar que $\mathrm{Im}(\gamma\tau) = \frac{\mathrm{Im}(\tau)}{|c\tau+d|^2}$ con $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{Sl}_2(\mathbb{R})$.
2. Probar que $\mathrm{Sl}_2(\mathbb{R})$ actúa transitivamente en $\mathcal{H}$. Hallar el estabilizador de $i \in \mathcal{H}$.
3. Probar que la medida $\frac{\mathrm{d}\,x\,\mathrm{d}\,y}{y^2}$ es invariante por $\mathrm{Sl}_2(\mathbb{R})$.
4. Probar que en $\mathcal{H}$, la métrica $(\mathrm{d}\,s)^2 = \frac{(\mathrm{d}\,x)^2+(\mathrm{d}\,y)^2}{y^2}$ es invariante por la acción de $\mathrm{Sl}_2(\mathbb{R})$.
5. Para la acción de $\mathrm{Sl}_2(\mathbb{Z})$ en $\overline{\mathcal{H}}$, probar que
   a) el estabilizador de ∞ es $\left\{\pm T^k,\ k \in \mathbb{Z}\right\}$,
   b) el estabilizador de i es $\left\{S^k,\ k = 0, 1, 2, 3\right\}$,
   c) el estabilizador de ω es $\left\{\left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)^k,\ k = 0, 1, 2, 3, 4, 5\right\}$,
   d) cualquier punto de $\overline{\mathcal{H}}$ con estabilizador más grande que $\pm\mathrm{Id}$ es equivalente a uno de los anteriores.

0.4. Superficies de Riemann. Una *superficie de Riemann* es un espacio topológico Hausdorff X con una estructura compleja, es decir, tal que para todo punto $x \in X$ existe un entorno abierto $x \in U \subseteq X$, un $V \subseteq \mathbb{C}$ abierto y un homeomorfismo $z : U \to V$, de manera tal que los entornos coordenados (U, z) sean compatibles. Entendemos aquí por "compatibles" que los cambios de coordenadas

$$z_2(z_1)^{-1} : z_1(U_1 \cap U_2) \to z_2(U_1 \cap U_2)$$

son funciones holomorfas, para todo par de entornos coordenados $(U_1, z_1 : U_1 \to V_1)$, y $(U_2, z_2 : U_2 \to V_2)$.

La función $f : X \to \mathbb{C}$ es holomorfa si lo es localmente, es decir, si las composiciones $fz : V \to \mathbb{C}$ son holomorfas para todos los $(U, z : U \to V)$ entornos coordenados.

También tiene sentido hablar de funciones meromorfas, órdenes de anulación y de polos y aplicaciones holomorfas $F : X \to Y$ entre dos superficies de Riemann, tomando entornos coordenados de la manera usual.

La superficie $Y(1)$ hereda de manera natural la estructura holomorfa de $\mathcal{H}$.

Sea $f : \mathcal{H} \to \mathbb{C}$ una función meromorfa que satisface la primera de las ecuaciones (0.7), es decir $f(\tau) = f(\tau + 1)$. Por ser periódica admite un desarrollo en series de Fourier

$$f(\tau) = \sum_{n\in\mathbb{Z}} a_n q^n.$$

Diremos que f es *meromorfa en* ∞ si hay a lo sumo un número finito de $a_n \neq 0$ con $n < 0$, y que es *holomorfa en* ∞ si $a_n = 0$ para todo $n < 0$.

Observación 0.1. $X(1)$ es topológicamente una esfera.

Para más detalles sobre superficies de Riemann, recomendamos [9], [6] o [10].

Ejercicios.

1. Sea X una superficie de Riemann. Probar que las funciones meromorfas de X se corresponden con aplicaciones holomorfas de X en la esfera de Riemann que no son constantemente ∞.
2. Sea $F : X \to Y$ una aplicación holomorfa no constante entre superficies de Riemann con X conexa. Probar que F es abierta.
3. Sea $F : X \to Y$ una aplicación holomorfa no constante entre superficies de Riemann con X compacta e Y conexa. Probar que F es sobreyectiva.
4. Sea X una superficie de Riemann compacta. Probar que las únicas funciones holomorfas de X en $\mathbb{C}$ son las constantes. Comparar con Teorema de Liouville.
5. Sean X una superficie de Riemann compacta y $x_1, \ldots, x_k \in X$ puntos diferentes. Sean $n_1, \ldots, n_k \in \mathbb{Z}_{\geq 0}$. Probar que el espacio de funciones meromorfas de X con polos únicamente en x_i de orden a lo sumo n_i tiene dimensión a lo sumo $1 + n_1 + \ldots + n_k$. (Nota: hay que agregar la función 0 para que formen un espacio vectorial.)

6. Sea X la esfera de Riemann. Probar que las funciones meromorfas de X son funciones racionales. Probar además que si $f : X \to \mathbb{C}$ sólo tiene un polo en ∞ entonces f es un polinomio cuyo grado coincide con el orden del polo. Comparar con el ejercicio anterior.
7. ¿Cómo deberían definirse los entornos coordenados de la clase de ∞ en $X(1)$?

1. Formas Modulares.

Definición 1.1. Sea f meromorfa en $\mathcal{H}$ y $k \in \mathbb{Z}$. Supongamos que f satisface

$$f(\gamma\tau) = (c\tau + d)^k f(\tau), \ \forall\, \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Sl}_2(\mathbb{Z}) \tag{1.1}$$

y meromorfa en ∞. Entonces, se dice que f es una *función modular de peso k para* $\mathrm{Sl}_2(\mathbb{Z})$.

Las funciones modulares se corresponden con formas diferenciales en $X(1)$ (ver ejercicio 16 de la sección 2 más abajo).

Proposición 1.2. *Sea $f(\tau)$ una función modular no nula de peso k para $\mathrm{Sl}_2(\mathbb{Z})$. Para cada $P \in \mathcal{H}$ notamos $\nu_P(f)$ al orden de anulación (o menos el orden del polo) de $f(\tau)$ en P. Sea $\nu_\infty(f)$ el subíndice del primer término no nulo en la q-expansión de f. Entonces*

$$\nu_\infty(f) + \frac{1}{2}\nu_i(f) + \frac{1}{3}\nu_\omega(f) + \sum_{\substack{P \in Y(1) \\ P \neq i,\omega}} \nu_P(f) = \frac{k}{12} \tag{1.2}$$

donde la suma se hace tomando un punto $P \in \mathcal{H}$ por cada clase de equivalencia, salvo la de $\omega = \frac{1+i\sqrt{3}}{2}$ y la de i.

Demostración. Para simplificar la demostración, supongamos que f no tiene ni ceros ni polos en el borde del dominio fundamental D, salvo quizás en i y en ω que es equivalente a $-\overline{\omega}$.

Integrando $\frac{1}{2\pi i}\frac{f'(\tau)}{f(\tau)}\,\mathrm{d}\,\tau$ a lo largo de un lazo que encierre a todos los ceros y polos de f del interior de D, recorrido en sentido antihorario, se obtiene la sumatoria del miembro izquierdo de (1.2).

Observar que $\nu_i(f)$ es la integral de $\frac{1}{2\pi i}\frac{f'(\tau)}{f(\tau)}\,\mathrm{d}\,\tau$ a lo largo de una circunferencia suficientemente pequeña alrededor de i recorrida en sentido antihorario. Por (1.3) del siguiente ejercicio 4 (con $\gamma = S = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$) la integral a lo largo del arco de la curva contenido en D recorrido en sentido horario tiende a $\frac{-1}{2}\nu_i(f)$ cuando el radio tiende a 0.

Análogamente, las integrales de los arcos dentro de D de circunferencias centradas en ω y $-\overline{\omega}$ tienden a $\frac{-1}{6}\nu_\omega(f) = \frac{-1}{6}\nu_{-\overline{\omega}}(f)$ (tomar γ recorriendo $\mathrm{Id}, TS, (TS)^2, \ldots, (TS)^5$, el estabilizador de ω).

La integral de $\frac{1}{2\pi i}\frac{f'(\tau)}{f(\tau)}\,\mathrm{d}\,\tau$ a lo largo de

$$C_N(t) = \frac{1}{2} - t + iN, \quad t \in [0,1]$$

tiende a $-\nu_\infty(f)$ cuando $N \to \infty$.

Integrando sobre el borde de:

$$D \cap \{\mathrm{Im}(\tau) \leq N\} - (\{|\tau - i| < \epsilon\} \cup \{|\tau - \omega| < \epsilon\} \cup \{|\tau + \overline{\omega}| < \epsilon\})$$

con $N \to \infty$ y $\epsilon \to 0$, obtenemos (gracias a (1.3) del ejercicio 4)

$$\sum_{\substack{P \in Y(1) \\ P \neq i,\omega}} \nu_P(f) = \frac{-1}{2}\nu_i(f) + \frac{-1}{6}\nu_\omega(f) + \frac{-1}{6}\nu_{-\overline{\omega}}(f) - \nu_\infty(f) - k\frac{1}{2\pi i}\int_{-\overline{\omega}}^{i} \frac{\mathrm{d}\,\tau}{\tau}$$

donde la última integral se recorre sobre $\{|\tau| = 1\}$. Calculando la integral (ver ejercicio 5) y reagrupando términos se obtiene (1.2) $\square$

Definición 1.3. Sea f una función modular de peso k para $\mathrm{Sl}_2(\mathbb{Z})$. Si además f es holomorfa en $\mathcal{H}$ y en ∞, decimos que es una *forma modular de peso k para* $\mathrm{Sl}_2(\mathbb{Z})$. Notamos al conjunto de tales funciones como $\mathrm{M}_k(\mathrm{Sl}_2(\mathbb{Z}))$.

Ejercicios.

1. Probar que $\mathrm{M}_0(\mathrm{Sl}_2(\mathbb{Z})) = \mathbb{C}$.
2. Probar que si k es impar $\mathrm{M}_k(\mathrm{Sl}_2(\mathbb{Z})) = 0$.
3. Probar que si k es negativo o 2 entonces $\mathrm{M}_k(\mathrm{Sl}_2(\mathbb{Z})) = 0$.
4. Sea $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{Sl}_2(\mathbb{Z})$ y sea $f(\tau)$ meromorfa en $\mathcal{H}$ sin ceros ni polos en una curva $C \subseteq \mathcal{H}$. Supongamos que $f(\gamma\tau) = (c\tau + d)^k f(\tau)$. Probar que

$$\int_C \frac{f'(\tau)}{f(\tau)}\,\mathrm{d}\,\tau - \int_{\gamma C} \frac{f'(\tau)}{f(\tau)}\,\mathrm{d}\,\tau = -k\int_C c\frac{\mathrm{d}\,\tau}{c\tau + d}. \tag{1.3}$$

5. Sea $C(t) = \exp(-2\pi i t), t \in [-1/3, -1/4]$ el arco de la circunferencia $\{|\tau| = 1\}$ recorrida en sentido horario desde $-\overline{\omega}$ hasta i. Probar que

$$\frac{1}{2\pi i}\int_C \frac{\mathrm{d}\,\tau}{\tau} = -\frac{1}{12}. \tag{1.4}$$

6. Sea $f \in \mathrm{M}_k(\mathrm{Sl}_2(\mathbb{Z}))$ no nula. Probar que
   - Si $k = 4$, entonces $\nu_\omega(f) = 1$, y $\nu_P(f) = 0$ para los demás P.
   - Si $k = 6$, entonces $\nu_i(f) = 1$, y $\nu_P(f) = 0$ para los demás P.
   - Si $k = 8$, entonces $\nu_\omega(f) = 2$, y $\nu_P(f) = 0$ para los demás P.
   - Si $k = 10$, entonces $\nu_\omega(f) = \nu_i(f) = 1$, y $\nu_P(f) = 0$ para los demás P.
   - Si $k = 14$, entonces $\nu_\omega(f) = 2$, $\nu_i(f) = 1$, y $\nu_P(f) = 0$ para los demás P.

1.1. Series de Eisenstein. A continuación definiremos una familia clásica de formas modulares de gran utilidad.

Sea $k \geq 4$ un entero par. La función

$$G_k(\tau) := \sum_{\substack{m,n\in\mathbb{Z} \\ (m,n)\neq(0,0)}} \frac{1}{(m\tau + n)^k} \tag{1.5}$$

es una forma modular de peso k para $\mathrm{Sl}_2(\mathbb{Z})$, llamada *serie de Eisenstein*, que verifica

$$\lim_{\tau\to i\infty} G_k(\tau) = 2\zeta(k), \tag{1.6}$$

donde $\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s}$, es la función zeta de Riemann.

Se define la *serie normalizada de Eisenstein* como

$$E_k(\tau) := \frac{G_k(\tau)}{2\zeta(k)}.$$

Proposición 1.4. *La expansión de Fourier de $E_k(\tau)$ es*

$$E_k(\tau) = 1 - \frac{2k}{B_k}\sum_{n=1}^{\infty} \sigma_{k-1}(n)q^n \tag{1.7}$$

donde los B_k son los números de Bernoulli, definidos por

$$\frac{t}{e^t - 1} = \sum_{m=0}^{\infty} B_m \frac{t^m}{m} \tag{1.8}$$

y

$$\sigma_k(n) := \sum_{d|n} d^k.$$

Demostración. Se tiene que

$$\pi\cot(\pi\tau) = \frac{1}{\tau} + \sum_{n=1}^{\infty}\left(\frac{1}{\tau+n} + \frac{1}{\tau-n}\right), \quad \tau \in \mathcal{H}. \tag{1.9}$$

También vale que

$$\pi\cot(\pi\tau) = \pi\frac{\cos(\pi\tau)}{\sin(\pi\tau)} = i\pi - \frac{2i\pi}{1-q} = i\pi - 2i\pi\sum_{n=0}^{\infty} q^n.$$

Igualando ambas expresiones y derivando $k-1$ veces término a término obtenemos la *fórmula de Lipschitz*

$$\sum_{n\in\mathbb{Z}} \frac{1}{(n+\tau)^k} = \frac{(-2\pi i)^k}{(k-1)!}\sum_{n=1}^{\infty} n^{k-1}q^n. \tag{1.10}$$

El resultado se obtiene sumando esta fórmula en $m\tau$ con $m \in \mathbb{Z}$ no nulo, más

$$\sum_{0\neq n\in\mathbb{Z}} \frac{1}{n^k} = 2\zeta(k),$$

junto con

$$\zeta(k) = -\frac{(2\pi i)^k B_k}{k!2} \tag{1.11}$$

para $k > 0$ par. □

Ejercicios.

1. Probar que para $k \in \{4, 6, 8, 10, 14\}$ el espacio $\mathrm{M}_k(\mathrm{Sl}_2(\mathbb{Z}))$ está generado por E_k.
2. Usando que $\dim(\mathrm{M}_8(\mathrm{Sl}_2(\mathbb{Z}))) = 1$ probar que $E_4^2 = E_8$. Concluir que
$$\sigma_7(n) = \sigma_3(n) + 120\sum_{i=1}^{n-1}\sigma_3(i)\sigma_3(n-i).$$
3. Usando que $\dim(\mathrm{M}_{10}(\mathrm{Sl}_2(\mathbb{Z}))) = 1$ probar que
$$11\sigma_9(n) = 21\sigma_5(n) - 10\sigma_3(n) + 5040\sum_{i=1}^{n-1}\sigma_3(i)\sigma_5(n-i).$$
4. Probar que $E_6E_8 = E_4E_{10} = E_{14}$.
5. Probar (1.6).
6. Sea $\Lambda \subseteq \mathbb{C}$ un reticulado (i.e.: un $\mathbb{Z}-$módulo discreto de rango 2). Probar que la serie
$$\sum_{0\neq\rho\in\Lambda} \frac{1}{|\rho|^t}$$
es absolutamente convergente para $t > 2$.
7. Probar que
$$E_k(\tau) := \frac{1}{2}\sum_{\substack{m,n\in\mathbb{Z}\\ \gcd(m,n)=1}} \frac{1}{(m\tau+n)^k}.$$
8. Usando (1.8), probar (1.11).
9. Probar (1.9) a partir de la fórmula de Euler:
$$\sin(\tau) = \tau\prod_{n=1}^{\infty}\left(1 - \frac{\tau^2}{(n\pi)^2}\right). \tag{1.12}$$
10. Deducir (1.9) de la fórmula de Poisson (0.6), usando que si $f(x) = e^{-|x|}$, la transformada $\hat{f}(\xi) = \dfrac{2}{1+4\pi^2\xi^2}$.

1.2. Peso 2. Cabe preguntarse qué sucede con la serie de Eisenstein en el caso $k = 2$. Aquí hace falta algo de cuidado con el orden de las series, puesto que dejan de converger de manera absoluta. Definimos

$$G_2(\tau) := \sum_{m\in\mathbb{Z}} \sum_{n} \frac{1}{(m\tau+n)^2} \tag{1.13}$$

donde el subíndice de la segunda suma recorre todos los enteros, salvo cuando $m = 0$ en cuyo caso $n \in \mathbb{Z} - \{0\}$.

Con este orden de los sumandos se tiene, al igual que antes

$$E_2(\tau) := \frac{G_2(\tau)}{2\zeta(2)} = 1 - 24\sum_{n=1}^{\infty} \sigma_1(n) q^n.$$

Proposición 1.5. *La serie de Eisenstein G_2 satisface la ecuación funcional*

$$\tau^{-2} G_2\left(-1/\tau\right) = G_2(\tau) - \frac{2\pi i}{\tau}. \tag{1.14}$$

Demostración. De (1.13) se obtiene

$$\tau^{-2} G_2(-1/\tau) = \sum_{n\in\mathbb{Z}} \sum_{m} \frac{1}{(m\tau+n)^2} = 2\zeta(2) + \sum_{n\in\mathbb{Z}} \sum_{m\neq 0} \frac{1}{(m\tau+n)^2} \tag{1.15}$$

donde la suma del medio recorre todo $\mathbb{Z}^2$ salvo el término $(m, n) = (0, 0)$.

Restando

$$\sum_{m\neq 0} \sum_{n\in\mathbb{Z}} \frac{1}{(m\tau+n)(m\tau+n+1)} = 0 \tag{1.16}$$

a $G_2 = 2\zeta(2) + \sum_{m\neq 0} \sum_{n\in\mathbb{Z}} (m\tau+n)^{-2}$ obtenemos

$$G_2(\tau) = 2\zeta(2) + \sum_{m\neq 0} \sum_{n\in\mathbb{Z}} \frac{1}{(m\tau+n)^2(m\tau+n+1)}.$$

Observemos que ahora la serie del miembro derecho es absolutamente convergente. Invirtiendo el orden de estas sumatorias y restandoselo a (1.15) tenemos

$$G_2(\tau) = \tau^{-2} G_2(-1/\tau) - \sum_{n\in\mathbb{Z}} \sum_{m\neq 0} \frac{1}{(m\tau+n)(m\tau+n+1)}.$$

La demostración termina usando

$$\lim_{N\to\infty} \sum_{n=-N}^{N-1} \sum_{m\neq 0} \left(\frac{1}{(m\tau+n)} - \frac{1}{(m\tau+n+1)} \right) = \lim_{N\to\infty} \frac{2}{\tau}\pi \cot\left(\frac{\pi N}{\tau}\right) - \frac{2}{\tau} = \frac{2\pi i}{\tau} \tag{1.17}$$

para $\tau \in \mathcal{H}$. □

Ejercicios.

1. Probar (1.16).
2. Completar los detalles de (1.17).
3. Probar que la función $G_2(\tau) - \pi/\mathrm{Im}(\tau)$ satisface (1.1) pero no es holomorfa.
4. Sea $f \in \mathrm{M}_k(\mathrm{Sl}_2(\mathbb{Z}))$. Sean

$$g(\tau) = \frac{1}{2\pi i} f'(\tau) - \frac{k}{12} E_2(\tau) f(\tau).$$

   Probar que $g \in \mathrm{M}_{k+2}(\mathrm{Sl}_2(\mathbb{Z}))$.
5. Probar que $E_6 = E_4E_2 - \frac{3}{2\pi i}E_4'$ y $E_8 = E_6E_2 - \frac{1}{\pi i}E_6'$. Deducir las correspondientes relaciones de σ_5 en términos de σ_1 y σ_3, y σ_7 en términos de σ_1 y σ_5.

6. Probar que
$$E_2(\tau+1/2)-E_2(\tau)=48\sum_{\substack{n>0\\ n \text{ impar}}}\sigma_1(n)q^n.$$

7. Sean $k\geq 2$ un entero y p un número primo. Probar que
$$E_k(\tau)-(1+p^{k-1})E_k(p\tau)+p^{k-1}E_k(p^2\tau)=-\frac{2k}{B_k}\sum_{p\nmid n}\sigma_{k-1}(n)q^n.$$

8. Probar que $E_2(\tau)-3E_2(2\tau)+2E_2(4\tau)=\frac{1}{2}\left(E_2(\tau)-E_2(\tau+1/2)\right)$.

1.3. Función eta de Dedekind. La ecuación funcional (1.14) sugiere la siguiente

Definición 1.6. Para $\tau\in\mathcal{H}$ el producto infinito
$$\eta(\tau):=e^{2\pi i\tau/24}\prod_{n=1}^{\infty}(1-e^{2\pi in\tau}) \tag{1.18}$$
converge a una función holomorfa, llamada *función eta de Dedekind.*

Proposición 1.7. *La eta de Dedekind satisface*
$$\eta(-1/\tau)=\sqrt{-iz}\eta(\tau) \tag{1.19}$$
donde la raíz cuadrada corresponde a la rama que toma valores con parte real no negativa.

Demostración. Ambos miembros de (1.19) coinciden en $\tau=i$. Tomando derivada logarítmica el problema se reduce a probar
$$\frac{\eta'(-1/\tau)}{\eta(-1\tau)}\tau^{-2}=\frac{1}{2\tau}+\frac{\eta'(\tau)}{\eta(\tau)}.$$

La derivada logarítmica de (1.18) nos da
$$\begin{aligned}\frac{\eta'(\tau)}{\eta(\tau)} &= \frac{2\pi i}{24}\left(1-24\sum_{n=1}^{\infty}\frac{nq^n}{1-q^n}\right)=\frac{2\pi i}{24}\left(1-24\sum_{n=1}^{\infty}\sigma_1(n)q^n\right)\\ &= \frac{2\pi i}{24}E_2(\tau).\end{aligned}$$

Por lo que el resultado se obtiene de (1.14). □

Observación 1.8. La función eta de Dedekind y (1.19) juegan un rol crucial en la demostración de la fórmula de Rademacher para el cálculo de la función $p(n)$ del ejercicio 2 de la Introducción. Para más detalles ver [1].

Ejercicios.

1. Probar que $\eta(\tau+1/2)=e^{2\pi i/48}\eta^3(2\tau)/\eta(\tau)\eta(4\tau)$.

1.4. Formas cuspidales.

Definición 1.9. Sea $f\in \mathrm{M}_k(\mathrm{Sl}_2(\mathbb{Z}))$ tal que se anula en ∞, es decir $a_0=0$ en la q-expansión $f(\tau)=\sum_{n\geq 0}a_nq^n$. Entonces f se dice *cuspidal,* y al espacio de formas cuspidales lo notamos $\mathrm{S}_k(\mathrm{Sl}_2(\mathbb{Z}))$.

Ejemplo: (*Función discriminante*). Dado que $E_4(\tau)=1+240q+O(q^2)\in \mathrm{M}_4(\mathrm{Sl}_2(\mathbb{Z}))$ y $E_6(\tau)=1-504q+O(q^2)\in \mathrm{M}_6(\mathrm{Sl}_2(\mathbb{Z}))$, la función
$$\Delta(\tau):=\frac{E_4(\tau)^3-E_6(\tau)^2}{1728}=q-24q^2+252q^3-1472q^4+\ldots \tag{1.20}$$
resulta ser una forma cuspidal de peso 12.

Ejemplo: La función

$$\eta^{24}(\tau) = q\prod_{n=1}^{\infty}(1-q^n)^{24} = q - 24q^2 + \dots$$

también es una forma cuspidal de peso 12, gracias a (1.19). Por tratarse de un producto absolutamente convergente se tiene que η^{24} no se anula en $\mathcal{H}$.

Entonces el cociente Δ/η^{24} es una forma modular de peso 0 que en ∞ vale 1. Las formas modulares de peso 0 son funciones holomorfas de $X(1)$, o sea constantes. Concluimos pues que $\Delta = \eta^{24}$, es decir

$$\frac{E_4(\tau)^3 - E_6(\tau)^2}{1728} = q\prod_{n=1}^{\infty}(1-q^n)^{24}.$$

Definición 1.10. Al cociente

$$j(\tau) := \frac{E_4(\tau)^3}{\Delta(\tau)} = q^{-1} + 744 + 196884q + 21493760q^2 + \dots \tag{1.21}$$

se lo conoce como *invariante modular j.*

Proposición 1.11. *El invariante modular j induce una biyección entre $X(1)$ y la esfera de Riemann.*

Demostración. Sea $c \in \mathbb{C}$ arbitrario. La función $j - c$ es función modular de peso 0, con un único polo simple en ∞. Por la proposición 1.2 se tiene que $\nu_P(j-c) > 0$ para un único P, es decir, j es biyectiva. □

Observación 1.12. Por propiedades de aplicaciones holomorfas se puede ver que j induce un isomorfismo entre ambas superficies de Riemann.

Proposición 1.13. *Sea f una forma cuspidal de peso k con expansión de Fourier $f(\tau) = \sum_{n=1}^{\infty} a_n q^n$. Entonces existe $C > 0$ (dependiente de f) tal que $|a_n| \le Cn^{k/2}$ para todo n.*

Demostración. La función $\tau \mapsto y^{k/2}|f(\tau)|$ es $\mathrm{Sl}_2(\mathbb{Z})$-invariante y tiende rápidamente a 0 cuando $y \to \infty$, por lo que permanece acotada en el dominio fundamental D. Se tiene pues que

$$|f(\tau)| \le cy^{-k/2}$$

para cierta $c > 0$. La representación integral

$$a_n = e^{2\pi ny}\int_0^1 f(x+iy)e^{-2\pi inx}\,\mathrm{d}\,x$$

váilda para $y > 0$, prueba que $|a_n| \le cy^{-k/2}e^{2\pi ny}$. Tomando $y = 1/n$ se obtiene el resultado. □

Ejercicios.

1. Probar que $\mathrm{S}_k(\mathrm{Sl}_2(\mathbb{Z})) = 0$ si $k < 12$ o $k = 14$.
2. Probar que $\mathrm{S}_{12}(\mathrm{Sl}_2(\mathbb{Z})) = \mathbb{C}\Delta$.
3. Probar que $\mathrm{S}_k(\mathrm{Sl}_2(\mathbb{Z})) = \Delta\,\mathrm{M}_{k-12}(\mathrm{Sl}_2(\mathbb{Z}))$ para $k > 14$.
4. Probar que $\mathrm{M}_k(\mathrm{Sl}_2(\mathbb{Z})) = \mathrm{S}_k(\mathrm{Sl}_2(\mathbb{Z})) \oplus \mathbb{C}E_k$ para $k > 2$.
5. Sea $k > 0$ un entero par. Probar que
$$\dim(\mathrm{M}_k(\mathrm{Sl}_2(\mathbb{Z}))) = \begin{cases} \lfloor \frac{k}{12} \rfloor + 1 & k \not\equiv 2 \bmod(12) \\ \lfloor \frac{k}{12} \rfloor & k \equiv 2 \bmod(12). \end{cases}$$
6. Sea $f \in \mathrm{M}_k(\mathrm{Sl}_2(\mathbb{Z}))$ una forma modular. Probar que existen $\alpha_{a,b} \in \mathbb{C}$ tales que
$$f = \sum_{\substack{a,b\in\mathbb{Z}_{\ge 0}\\ 4a+6b=k}} \alpha_{a,b}E_4^aE_6^b.$$

7. Probar que las funciones modulares de peso 0 para $\mathrm{Sl}_2(\mathbb{Z})$ son precisamente las funciones racionales en j.
8. Probar el enunciado de 1.12.
9. Probar que en el ejercicio 4 de sección 1.2, la g construida resulta cuspidal si y sólo si la f lo es.
10. Probar que los coeficientes de la q-expansión (1.21) de j son enteros.

1.5. Interpretación modular. Las formas modulares pueden pensarse como funciones en el espacio de reticulados de $\mathbb{C}$. Dado un reticulado $\Lambda \subseteq \mathbb{C}$, se puede hallar una base $z_1, z_2 \in \Lambda$ tales que $z_1/z_2 = \tau \in \mathcal{H}$, por lo que Λ resulta homotético a otro reticulado $\Lambda_\tau = \mathbb{Z} \oplus \tau\mathbb{Z}$. De ahí que el cociente del espacio de reticulados por la relación de equivalencia

$$\Lambda_1 \sim \Lambda_2 \quad \Longleftrightarrow \quad \exists \lambda \in \mathbb{C}^*,\ \Lambda_1 = \lambda\Lambda_2$$

resulta ser $Y(1) = \mathrm{Sl}_2(\mathbb{Z})\backslash\mathcal{H}$.

Se definen los *operadores de Hecke* T_n como

$$T_n(F)(\Lambda) := \sum_{[\Lambda:\Lambda']=n} F(\Lambda')$$

donde la suma recorre los subreticulados de Λ de índice n.

Las funciones $f : \mathcal{H} \to \mathbb{C}$ invariantes por los operadores $[\gamma]_k$ pueden pensarse como funciones *de peso* k en el espacio de reticulados, es decir, funciones F que verifican

$$F(\lambda\Lambda) = \lambda^{-k} F(\Lambda),$$

via $F(z_1\mathbb{Z} \oplus z_2\mathbb{Z}) = z_2^{-k} f(z_1/z_2)$.

Con esta interpretación, los T_n definen operadores en los $\mathrm{M}_k(\mathrm{Sl}_2(\mathbb{Z}))$, preservan los $\mathrm{S}_k(\mathrm{Sl}_2(\mathbb{Z}))$ (ver ejercicio 8 de la sección 2.4) y satisfacen

1. $T_n T_m = T_{mn}$ si m y n son coprimos,
2. $T_{p^{r+1}} = T_p T_{p^r} - p^{k-1} T_{p^{r-1}}$,
3. son autoadjuntos para un producto interno (ver sección 2.3 más adelante),
4. conmutan entre sí.

Resulta entonces que los T_n son simultáneamente diagonalizables. Llamamos *autoforma* a un autovector común a todos los operadores de Hecke. Si $f = \sum a_n q^n \in \mathrm{S}_k(\mathrm{Sl}_2(\mathbb{Z}))$ es una autoforma, sus coeficientes verifican

$$a_n = \lambda_n a_1 \tag{1.22}$$

donde $T_n(f) = \lambda_n f$ (ver proposición 2.23 más adelante), de donde se deduce que $a_1 \neq 0$.

Decimos que tal f está *normalizada* si $a_1 = 1$.

Proposición 1.14. *Sea $f = \sum a_n q^n \in \mathrm{S}_k(\mathrm{Sl}_2(\mathbb{Z}))$ una autoforma normalizada. Entonces, los coeficientes verifican*

1. *$a_n a_m = a_{mn}$ si m y n son coprimos,*
2. *$a_{p^{r+1}} = a_p a_{p^r} - p^{k-1} a_{p^{r-1}}$,*

y la función $L(f,s) := \sum_{n\geq 1} \frac{a_n}{n^s}$ admite un producto de Euler

$$L(f,s) = \prod_p \frac{1}{1 - a_p p^{-s} + p^{k-1-2s}} \tag{1.23}$$

donde p recorre los números primos positivos.

Demostración. Las identidades entre los coeficientes se deducen de las análogas para operadores de Hecke, observando que para una f normalizada los coeficientes coinciden con los autovalores. Para el producto de Euler, observemos primero que por la propiedad multiplicativa de los a_n se tiene que

$$L(f,s) = \prod_p \left(\sum_{n=0}^{\infty} a_p p^{-ns} \right)$$

y distribuyendo

$$\left(\sum_{n=0}^{\infty} a_p p^{-ns}\right)\left(1 - a_p p^{-s} + p^{k-1-2s}\right) = 1$$

se obtiene (1.23). □

Observación 1.15. Hecke demostró que $L(f,s)$ se extiende de manera analítica a una función meromorfa de $\mathbb{C}$, y que

$$\Lambda_f(s) := (2\pi)^{-s}\Gamma(s)L(f,s)$$

verifica una ecuación funcional

$$\Lambda_f(s) = (-1)^{k/2}\Lambda_f(k-s).$$

(aquí $\Gamma(s) = \int_0^\infty x^{s-1}e^{-x}\,\mathrm{d}x$ es la función gamma de Euler).

Ejemplo 1.16. Como $\dim(\mathrm{S}_{12}(\mathrm{Sl}_2(\mathbb{Z}))) = 1$, la función $\Delta = q\prod(1-q^n)^{24}$ resulta ser una autoforma normalizada. Luego sus coeficientes satisfacen las identidades de la proposición 1.14. Dicha propiedad fue observada originalmente por Ramanujan (ver [13]).

Observación 1.17. La conjetura de Ramanujan-Petersson, probada por Deligne como consecuencia de su demostración de las conjeturas de Weil, establece que si f es una autoforma cuspidal normalizada de peso k para $\Gamma_1(N)$ entonces $|a_n| \leq n^{(k-1)/2}\sigma_0(n)$, resultado mucho más fuerte que la proposición 1.13 (debida al mismo Hecke). De donde también se tiene que los coeficientes a_n de Δ satisfacen $|a_p| \leq 2p^{11/2}$ con p primo.

Ejercicios.

1. Probar que $\phi(s) = \sum a_n n^{-s}$ converge en algún semiplano si y sólo si $a_n = O(n^c)$, para alguna constante c.
2. Supongamos
$$f(\tau) = \sum_{n=0}^{\infty} a_n q^n$$
es analítica en $\mathcal{H}$.
   a) Si $a_n = O(n^c)$, entonces $f(\tau) = O(y^{-c-1})$, uniformemente x, cuando $y \to 0^+$.
   b) Si $f(\tau) = O(y^{-c})$, uniformemente en x, entonces $a_n = O(n^c)$.
3. Sea $c > 0$. Probar que
$$e^{-x} = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \Gamma(s)x^{-s}\,\mathrm{d}\,s$$
para $x > 0$.
4. (Transformada de Mellin) Sean $a_n = O(n^M)$ para algún M, números complejos. Sea $f(x) = \sum_{n=1}^{\infty} a_n e^{-nx}$ y $\phi(s) = \sum_{n=1}^{\infty} a_n n^{-s}$. Probar que
$$\Gamma(s)\phi(s) = \int_0^\infty f(x)x^{s-1}\,\mathrm{d}\,x$$
para $\Re(s) > \max\{0, M+1\}$, y
$$f(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \phi(s)\Gamma(s)x^{-s}\,\mathrm{d}\,s$$
para $c > \max\{0, M+1\}$ y $\Re(s) > 0$.

2. Formas modulares para subgrupos de congruencia.

Sea N un entero positivo. Se define el *grupo de congruencia principal de nivel* N como

$$\Gamma(N) := \left\{ \gamma \in \mathrm{Sl}_2(\mathbb{Z}) : \ \gamma \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mod(N) \right\}.$$

Un subgrupo de $\mathrm{Sl}_2(\mathbb{Z})$ se dice *de congruencia* si contiene a algún $\Gamma(N)$. Llamamos *cúspides* a las clases de equivalencia de $\mathbb{Q} \cup \{\infty\}$.

Definición 2.1. Sean $k \in \mathbb{Z}$ y $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{Sl}_2(\mathbb{Z})$. Definimos $[\gamma]_k$, el *operador de peso* k, en el espacio de funciones meromorfas de $\mathcal{H}$ a $\mathbb{C}$ como

$$(f[\gamma]_k)(\tau) := (c\tau + d)^k f(\gamma\tau). \tag{2.1}$$

Observación 2.2. Notar que

$$f[\gamma_1\gamma_2]_k = (f[\gamma_1]_k)[\gamma_2]_k \tag{2.2}$$

para todo par de matrices $\gamma_1, \gamma_2 \in \mathrm{Sl}_2(\mathbb{Z})$. Basta escribir $(c\tau + d)^{-k}$ como $(d\gamma\tau/d\tau)^{k/2}$ y usar regla de la cadena.

Definición 2.3. Sea Γ un subgrupo de congruencia conteniendo a $\Gamma(N)$ y $k \in \mathbb{Z}$. Decimos que una función $f : \mathcal{H} \to \mathbb{C}$ es una *función modular de peso* k *para* Γ si

1. f es meromorfa en $\mathcal{H}$,
2. $f[\gamma]_k = f$ para toda $\gamma \in \Gamma$,
3. $f[\gamma]_k$ es meromorfa en ∞ para toda $\gamma \in \mathrm{Sl}_2(\mathbb{Z})$.

Análogamente, diremos que una tal f es una *forma modular* para Γ si es holomorfa en $\mathcal{H}$, y $f[\gamma]_k$ es holomorfa en ∞ para toda $\gamma \in \mathrm{Sl}_2(\mathbb{Z})$, y que es *cuspidal* si además las $f[\gamma]_k$ se anulan en ∞. A los espacios de formas modulares y cuspidales de peso k para Γ, los notamos $\mathrm{M}_k(\Gamma)$ y $\mathrm{S}_k(\Gamma)$, respectivamente.

Aquí entendemos por *meromorfa en* ∞ que aparecen a lo sumo finitos términos negativos en el desarrollo en series de potencias de $q_N := \exp(2\pi i\tau/N)$, que es *holomorfa en* ∞ si no aparecen términos negativos, y que *se anula en* ∞ si sólo aparecen términos positivos.

Tal expansión existe dado que al tomar $\left(\begin{smallmatrix} 1 & N \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma(N)$ en la condición 2 se tiene que

$$f(\tau + N) = f(\tau),$$

y por lo tanto f admite un desarrollo en serie de potencias de $q_N = \exp(2\pi i/N)$.

Observación 2.4. Como $\Gamma(N)$ es normal en $\mathrm{Sl}_2(\mathbb{Z})$, $f[\gamma]_k$ es también N-periódica, para toda $\gamma \in \mathrm{Sl}_2(\mathbb{Z})$.

Observación 2.5. A diferencia de lo que sucede en el caso de nivel 1, para que una forma modular sea cuspidal no alcanza con chequear un sólo coeficiente. Existen formas no cuspidales f que se anulan en ∞ pero alguna $f[\gamma]_k$ no (ver ejercicio 17).

Observación 2.6. Como Γ tiene índice finito en $\mathrm{Sl}_2(\mathbb{Z})$, sólo hace falta verificar la condición 3 para un conjunto finito de γ's, más precisamente, basta con un conjunto de representantes de las Γ-coclases.

Proposición 2.7. *La condición 3 y sus análogas para formas modulares y cuspidales sólo dependen de la* Γ*-clase de equivalencia de* $\gamma\infty$*. Más precisamente, si* $\gamma_1\infty = \gamma'\gamma_2\infty$ *para* $\gamma' \in \Gamma$*, entonces la menor potencia de* q_N *que aparece en la expansión de Fourier de* $f[\gamma_1]_k$ *coincide con la de* $f[\gamma_2]_k$*.*

Demostración. Como $\gamma_1^{-1}\gamma'\gamma_2$ pertenece al estabilizador de ∞, entonces es de la forma $\pm T^j$ (ver ejercicio 5 de la sección 0.3).

Entonces $\gamma_2 = \pm\gamma'^{-1}\gamma_1 T^j$ y

$$f[\gamma_2]_k = f[\pm \mathrm{Id}]_k[\gamma'^{-1}]_k[\gamma_1]_k[T^j]_k = (\pm 1)^k f[\gamma'^{-1}]_k[\gamma_1]_k[T^j]_k = (\pm 1)^k f[\gamma_1]_k[T^j]_k.$$

Escribiendo $g(\tau) = f[\gamma_1]_k = \sum_n a_n q_N^n$ tenemos que

$$f[\gamma_2]_k = (\pm 1)^k g(\tau + j) = (\pm)^k \sum_n a_n e^{2\pi i n j/N} q_N^n,$$

de donde se deducen las afirmaciones del enunciado. □

El siguiente lema suele ser de utilidad

Lema 2.8. *Sea f holomorfa en $\mathcal{H}$ que verifica la condición 2 de la definición 2.3 para Γ un subgrupo de congruencia de nivel N. Supongamos que la expansión de Fourier $f(\tau) = \sum_{n=0}^{\infty} a_n q_N^n$ satisface*

$$|a_n| \leq C n^r, \quad n > 0$$

para ciertas constantes positivas C y r. Entonces f satisface también la condición 3.

Demostración. La función $f[\gamma]_k$ es invariante por la acción de los operadores $[\gamma']_k$ con $\gamma' \in \gamma^{-1}\Gamma\gamma$ que también es un subgrupo de congruencia de nivel N. Para ver que la serie $f[\gamma]_k(\tau) = \sum_{n\in\mathbb{Z}} a'_n q_N^n$ no tiene términos negativos alcanza con probar que

$$\lim_{q_N \to 0} (f[\gamma]_k(\tau) q_N) = 0. \tag{2.3}$$

Si γ fija al ∞ esto es inmediato (ver proposición 2.7).

Supongamos que γ no fija al ∞.

Como $|a_n| \leq C n^r$, escribiendo $\tau = x + iy$ se tiene que

$$|f(\tau)| \leq C_1 + \frac{C_2}{y^r} \tag{2.4}$$

cuando $y \to \infty$, para ciertas constantes $C_1, C_2 > 0$ (ver ejercicio 15).

De donde

$$\lim_{q_N \to 0} |f[\gamma]_k(\tau) q_N| \leq C_3 \lim_{q \to 0} y^{r-k} |q_N| = 0 \tag{2.5}$$

dado que $|q_N|$ es exponencial en y. □

Afirmación 2.1. *Sea $\Gamma \subseteq \mathrm{Sl}_2(\mathbb{Z})$ de índice finito n, y*

$$\mathrm{Sl}_2(\mathbb{Z}) = \bigcup_{i=1}^{n} \alpha_i \Gamma$$

una descomposición en coclases disjuntas. Si D es un dominio fundamental para $\mathrm{Sl}_2(\mathbb{Z})\backslash\mathcal{H}$, entonces $D' = \cup_{i=1}^n \alpha_i^{-1} D$ es un dominio fundamental para $\Gamma\backslash\mathcal{H}$.

Al igual que en el caso de nivel 1, el espacio $\Gamma\backslash\overline{\mathcal{H}}$ tiene estructura de superficie de Riemann compacta, las únicas formas modulares de peso 0 son las constantes, y los espacios $\mathrm{M}_k(\Gamma)$ resultan de dimensión finita (ver ejercicio 5 de la sección 0.4).

La siguiente *cota de Sturm* resulta muy útil para acotar dimensiones de espacios de formas modulares.

Proposición 2.9. *Sea Γ un subgrupo de congruencia de índice M y sea $f \in \mathrm{M}_k(\Gamma)$ una forma modular. Si*

$$\nu_\infty(f) > M.\frac{k}{12} \tag{2.6}$$

entonces $f = 0$.

Demostración. Para $\Gamma = \mathrm{Sl}_2(\mathbb{Z})$ alcanza con recordar que en (1.2) todos los términos son negativos para concluir que f es nula.

En el caso general escribimos

$$\mathrm{Sl}_2(\mathbb{Z}) = \bigcup_{i=1}^{M} \Gamma\gamma_i$$

para γ_i apropiados, con $\gamma_1 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. La función

$$F := f \prod_{i=2}^{M} f[\gamma_i]_k$$

resulta ser una forma modular de peso Mk para $\mathrm{Sl}_2(\mathbb{Z})$ que satisface el enunciado de la proposición y el resultado se deduce del caso ya probado. □

Observación 2.10. El orden de anulación $\nu_\infty(f)$ de (2.6) puede no ser entero, f podría necesitar términos que sean potencias de q_N pero no de q.

Ejercicios.

1. Sea Γ un subgrupo de congruencia. Probar que $\mathrm{M}_0(\Gamma) = \mathbb{C}$.
2. Sea Γ un grupo de congruencia y $P = \{\pm T^k\}$ el estabilizador del ∞. Probar que la aplicación
$$\Gamma \backslash \mathrm{Sl}_2(\mathbb{Z}))/P \to \{ \text{ cúspides de } \Gamma\}$$
dada por
$$\Gamma\alpha P \mapsto \Gamma\alpha(\infty)$$
es una biyección.
3. Sea $\alpha \in \mathrm{Gl}_2^+(\mathbb{Q})$, y Γ un subgrupo de congruencia. Probar que $\Gamma_\alpha := \Gamma \cap \alpha^{-1}\Gamma\alpha$ también resulta de congruencia, no necesariamente del mismo nivel.
4. Deducir de la cota de Sturm que $\dim(\mathrm{M}_k(\Gamma)) < \infty$ para cualquier subgrupo de congruencia Γ.
5. Probar que $441E_4E_8 + 250E_6^2 = 691E_{12}$.
6. Sea Γ un subgrupo de congruencia.
   a) Probar que toda función modular para Γ satisface una ecuación polinomial de grado $[\mathrm{Sl}_2(\mathbb{Z}) : \Gamma]$ sobre el cuerpo $\mathbb{C}(j)$ de funciones modulares de peso 0 para $\mathrm{Sl}_2(\mathbb{Z})$.
   b) Probar que si Γ es normal y f es Γ-invariante, entonces lo es también $f[\alpha]_0$ para toda $\alpha \in \mathrm{Sl}_2(\mathbb{Z})$.
   c) Probar que el cuerpo de funciones modulares de peso 0 para Γ es una extensión de Galois de $\mathbb{C}(j)$ con grupo de Galois igual a $\mathrm{Sl}_2(\mathbb{Z})/\Gamma$.
7. Encontrar formas modulares de peso fijo con ν_∞ arbitrariamente grande.
8. Encontrar formas modulares para un subgrupo de congruencia fijo con ν_∞ arbitrariamente grande.
9. Sea $\mathrm{Gl}_2^+(\mathbb{Q})$ el subgrupo de matrices de determinante positivo de $\mathrm{Gl}_2(\mathbb{Q})$. Extendemos la definición 2.1 como $(f[\gamma]_k)(\tau) := (\det\gamma)^{k/2}(c\tau + d)^{-k} f(\tau)$. Probar que también vale (2.2) de la Observación 2.2.
10. Sea p un primo y $e \geq 1$
    a) Hallar el orden de $\mathrm{Sl}_2(\mathbb{Z}/p\mathbb{Z})$ y $\mathrm{Gl}_2(\mathbb{Z}/p\mathbb{Z})$.
    b) Hallar el núcleo del homomorfismo $\mathrm{Gl}_2(\mathbb{Z}/p^e\mathbb{Z}) \to \mathrm{Gl}_2(\mathbb{Z}/p\mathbb{Z})$.
    c) Hallar el orden de $\mathrm{Sl}_2(\mathbb{Z}/p^e\mathbb{Z})$ y $\mathrm{Gl}_2(\mathbb{Z}/p^e\mathbb{Z})$.
11. Sea $N = p_1^{e_1} \dots p_r^{e_r}$ la factorización en primos de N. Usando el Teorema Chino del Resto hallar el orden de $\mathrm{Sl}_2(\mathbb{Z}/N\mathbb{Z})$.
12. Probar que $[\mathrm{Sl}_2(\mathbb{Z}) : \Gamma(N)] = N^3 \prod_{p|N} \left(1 - \frac{1}{p^2}\right)$.
13. Probar que (2.3) implica $f[\gamma]_k$ holomorfa en ∞.
14. Completar los detalles de (2.5).
15. El objetivo de este ejercicio es demostrar la cota (2.4) a partir de las hipótesis del lema 2.8.
    a) Probar que $g(t) = t^r e^{-2\pi t y/N}$ es creciente en $[0, \frac{rN}{2\pi y}]$ y decreciente en $[\frac{rN}{2\pi y}, \infty]$.
    b) Probar que
$$|f(\tau)| \leq |a_0| + C \sum_{n=1}^{\infty} n^r e^{-2\pi n y/N} \leq C_1 + C_0 \left(\int_0^\infty g(t)\,\mathrm{d}\,t + \frac{1}{y^r} \right).$$

c) Concluir.

16. Probar que la condición 2 equivale a pedir que la forma $f(\tau)(\mathrm{d}\,\tau)^k$ sea Γ-invariante.
17. Adaptar el argumento de la proposición 1.13 para probar que $\mathrm{S}_k(\Gamma)$ se puede definir como aquellas $f \in \mathrm{M}_k(\Gamma)$ tales que $y^{k/2} f(x+iy)$ permanece acotada.
18. Sea
$$\lambda(\tau) := \left(\frac{\eta(\tau/2)\eta^2(2\tau)}{\eta^3(\tau)}\right)^8.$$
a) Usando la cota de Sturm para $\Gamma(2)$, probar que
$$j(\tau) = 256\frac{(1-\lambda(\tau)+\lambda(\tau)^2)}{\lambda(\tau)^2(1-\lambda(\tau))^2}.$$
b) Probar que λ es inyectiva restringida al interior de un dominio fundamental para $\Gamma(2)$.
c) Probar que la imagen de λ es $\mathbb{C} - \{0,1\}$.
d) Construir un revestimiento $p : B(0,1) \to \mathbb{C} - \{0,1\}$.
e) (Pequeño Teorema de Picard) Sea f una función entera que omite dos valores. Probar que f es constante.

2.1. El subgrupo de congruencia $\Gamma_0(N)$. Como veremos más adelante, el siguiente subgrupo de congruencia es de particular interés
$$\Gamma_0(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Sl}_2(\mathbb{Z}) :\ c \equiv 0 \mod(N) \right\}.$$
Al cociente $\Gamma_0(N)\backslash\overline{\mathcal{H}}$ se lo conoce como *curva modular clásica* y se la nota $X_0(N)$.

Ejemplo 2.11. Veamos que la función Θ^4 es una forma modular de peso 2 para $\Gamma_0(4)$.

La matriz $\left(\begin{smallmatrix} 0 & -1 \\ 4 & 0 \end{smallmatrix}\right)$ que manda τ a $-1/(4\tau)$, no tiene determinante 1, pero sí lo tiene
$$\begin{pmatrix} 0 & 1/4 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 4 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix}.$$
que manda τ a $\tau/(4\tau+1)$.

De ahí se tiene que
$$\Theta\left(\frac{\tau}{4\tau+1}\right) = \sqrt{4\tau+1}\Theta(\tau),$$
por lo que $\Theta^4[\left(\begin{smallmatrix} 1 & 0 \\ 4 & 1 \end{smallmatrix}\right)]_2 = \Theta^4$. También se tiene que $\Theta^4[\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)]_2 = \Theta^4$. Como $\Gamma_0(4)$ está generado por $\pm\left(\begin{smallmatrix} 1 & 0 \\ 4 & 1 \end{smallmatrix}\right), \pm\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, concluimos que Θ^4 es una función modular de peso 2 para $\Gamma_0(4)$.

Resta chequear que es holomorfa en las cúspides $\infty, 0$ y $-1/2$.

De la definición 0.5 de Θ se tiene que es holomorfa en ∞ y por lo tanto también lo es Θ^4.

Por el lema 2.8 se ve que también es holomorfa en 0 y $-1/2$.

Ejercicios.

1. Hallar $[\mathrm{Sl}_2(\mathbb{Z}) : \Gamma_0(N)]$.
2. Hallar $[\Gamma_0(N) : \Gamma(N)]$.
3. Encontrar un isomorfismo entre $\Gamma(N)$ y un subgrupo de $\Gamma_0(N^2)$ de índice $\varphi(N)$. Concluir que $\Gamma_0(4)$ es isomorfo a $\Gamma(2)$.
4. Probar que si $f \in \mathrm{M}_k(\mathrm{Sl}_2(\mathbb{Z}))$ entonces $g(\tau) := f(N\tau) \in \mathrm{M}_k(\Gamma_0(N))$.
5. Sea k par y f una función que verifica $f(\tau) = f(\tau+1)$ y $f(-1/4\tau) = (-4t^2)^{k/2} f(\tau)$. Probar que $f[\gamma]_k = f$ para toda $\gamma \in \Gamma_0(4)$.
6. Probar que la siguiente es una lista completa de representantes para $\Gamma_0(p^e)$ con p primo
$$\mathrm{Id};\quad T^{-k}S,\ k = 0,1,\ldots,p^e-1;\quad ST^{kp}S,\ k = 1,2,\ldots,p^{e-1}-1.$$
7. Probar que las cúspides para $\Gamma_0(p)$ son 0 e ∞, y que para $\Gamma_0(p^2)$ son $0, \infty$, y $-1/kp$ con $k = 1,\ldots,p-1$.
8. Probar que $\eta^8(4\tau)/\eta^4(2\tau) \in \mathrm{M}_2(\Gamma_0(4))$.

9. Probar que $E_2(\tau) - 3E_2(2\tau) + 2E_2(4\tau) \in \mathrm{M}_2(\Gamma_0(4))$.
10. Probar la identidad
$$q\prod_{n=1}^{\infty}(1-q^{4n})^4(1+q^{2n})^4 = \sum_{\substack{n>0\\ n \text{ impar}}} \sigma_1(n)q^n.$$
11. Probar que $E_2(\tau) - \frac{1}{N}\sum_{j=0}^{N-1} E_2(\tau + j/N) \in \mathrm{M}_2(\Gamma_0(N^2))$.
12. Sea N un entero positivo. Probar que $G_{2,N}(\tau) := G_2(\tau) - NG_2(N\tau) \in \mathrm{M}_2(\Gamma_0(N))$.
13. Probar que
$$G_{2,2} = -\frac{\pi^2}{3}\left(1 + 24\sum_{n=1}^{\infty}\left(\sum_{d|n\,,\,2\nmid d} d\right)q^n\right).$$
14. Probar que
$$G_{2,4} = -\pi^2\left(1 + 8\sum_{n=1}^{\infty}\left(\sum_{d|n\,,\,4\nmid d} d\right)q^n\right).$$
15. Sabiendo que $\dim(\mathrm{M}_2(\Gamma_0(4))) = 2$, probar que $G_{2,2}(\tau)$ y $G_{2,4}(\tau)$ son una base de $\mathrm{M}_2(\Gamma_0(4))$.
16. Probar (0.2) y (0.4).
17. Sea $k \geq 4$ y p un primo. Probar que $E_2(\tau) - pE_2(p\tau) \in \mathrm{M}_k(\Gamma_0(p))$ pero no es una forma cuspidal, a pesar de no tener término constante en su q-expansión.
18. Probar que la única forma cuspidal normalizada de $\mathrm{S}_2(\Gamma_0(32))$ es $\eta^2(4\tau)\eta^2(8\tau)$.

2.2. El subgrupo de congruencia $\Gamma_1(N)$. Otro subgrupo de congruencia de interés es
$$\Gamma_1(N) := \left\{\begin{pmatrix} a & b\\ c & d\end{pmatrix} \in \mathrm{Sl}_2(\mathbb{Z}) :\ a \equiv d \equiv 1,\ c \equiv 0 \mod(N)\right\}.$$
Al cociente $\Gamma_1(N)\backslash\overline{\mathcal{H}}$ se lo nota $X_1(N)$.

Observación 2.12. Sea $\alpha = \left(\begin{smallmatrix} N & 0\\ 0 & 1\end{smallmatrix}\right)$. Si $f \in \mathrm{M}_k(\Gamma(N))$, entonces
$$f[\alpha]_k \in \mathrm{M}_k(\alpha\Gamma(N)\alpha^{-1}) \subseteq \mathrm{M}_k(\Gamma_1(N^2)).$$
Es decir, a las formas modulares de peso k para $\Gamma(N)$ las podemos pensar dentro de $\mathrm{M}_k(\Gamma_1(N^2))$. Esto suele resultar cómodo dado que $\left(\begin{smallmatrix} 1 & 1\\ 0 & 1\end{smallmatrix}\right)$ pertenece a $\Gamma_1(N^2)$ pero no a $\Gamma(N)$ cuando $N > 1$, permitiendo desarrollar en series de q, en lugar de recurrir a q_N.

Definición 2.13. Se tiene que $\Gamma_1(N)$ es normal en $\Gamma_0(N)$ y su cociente es $(\mathbb{Z}/N\mathbb{Z})^*$. Llamamos *carácter de Dirichlet* a un homomorfismo de grupos $\chi : (\mathbb{Z}/N\mathbb{Z})^* \to \mathbb{C}^*$ extendido a $\mathbb{Z}$ de manera natural como $\chi(n) = 0$ para $\gcd(n, N) > 1$. Decimos que $f \in \mathrm{M}_k(\Gamma_1(N))$ es una *forma modular de peso k para $\Gamma_0(N)$ con carácter χ* si satisface
$$f[\gamma]_k = \chi(d)f \quad \text{para toda } \begin{pmatrix} a & b\\ c & d\end{pmatrix} \in \Gamma_0(N).$$
Notamos $\mathrm{M}_k(\Gamma_0(N), \chi)$ al espacio de tales funciones y $\mathrm{S}_k(\Gamma_0(N), \chi)$ al subespacio de formas cuspidales.

Descomponiendo en autoespacios la acción natural de
$$(\mathbb{Z}/N\mathbb{Z})^* = \Gamma_0(N)/\Gamma_1(N)$$
en $\mathrm{M}_k(\Gamma_1(N))$ se obtiene
$$\mathrm{M}_k(\Gamma_1(N)) = \bigoplus_{\chi} \mathrm{M}_k(\Gamma_0(N), \chi). \tag{2.7}$$

Ejercicios.

1. Probar que $[\Gamma_1(N) : \Gamma(N)] = N$
2. Probar que $[\Gamma_0(N) : \Gamma_1(N)] = \varphi(N)$.
3. Completar los detalles de la Observación 2.12.
4. Probar que si $\chi(-1) \neq (-1)^k$ entonces $\mathrm{M}_k(\Gamma_0(N), \chi) = 0$.
5. Sea χ_4 el carácter no trivial de $(\mathbb{Z}/4\mathbb{Z})^*$. Probar que
$$\mathrm{M}_k(\Gamma_1(4)) = \begin{cases} \mathrm{M}_k(\Gamma_0(4)) & \text{si } k \text{ es par} \\ \mathrm{M}_k(\Gamma_0(4), \chi_4) & \text{si } k \text{ es impar.} \end{cases}$$
6. Probar que $\mathrm{S}_5(\Gamma_1(4))$ está generado por $\eta^4(\tau)\eta^2(2\tau)\eta^4(4\tau)$.
7. Probar que $\mathrm{S}_8(\Gamma_1(4))$ está generado por $f(\tau) := (\eta(\tau)\eta(2\tau))^8$ y $g(\tau) = f(2\tau)$.
8. Sean N y k enteros positivos tales que $k(N+1) = 24$. Probar que $(\eta(\tau)\eta(N\tau))^k \in \mathrm{S}_k(\Gamma_0(N))$, salvo para $k = 1$ o $k = 3$ en cuyo caso $(\eta(\tau)\eta(N\tau))^k \in \mathrm{S}_k(\Gamma_0(N), \chi)$, con χ de orden 2.

2.3. Producto escalar de Petersson. Si $f, g \in \mathrm{M}_k(\Gamma)$, la función $f(\tau)\overline{g(\tau)}y^k$ resulta invariante por $\mathrm{Sl}_2(\mathbb{R}.)$ Cuando al menos una de ambas es cuspidal, definimos el *producto escalar de Petersson* como

$$\langle f, g\rangle := \frac{1}{[\mathrm{PSl}_2(\mathbb{Z}) : \mathrm{PSl}_2(\mathbb{Z}) \cap \Gamma]} \iint_{D'} f(\tau)\overline{g(\tau)}y^k \frac{\mathrm{d}\, x\, \mathrm{d}\, y}{y^2} \tag{2.8}$$

donde D' es un dominio fundamental para la acción de Γ en $\mathcal{H}$. Dicho producto interno está bien definido, la integral converge y no depende del dominio fundamental elegido. Si f y g también pertenecen a $\mathrm{M}_k(\Gamma')$ para otro subgrupo de congruencia Γ', el producto así definido no depende del espacio en el que se las considere.

Para $\alpha \in \mathrm{Gl}_2^+(\mathbb{Q})$, el subgrupo $\Gamma_\alpha := \Gamma \cap \alpha^{-1}\Gamma\alpha$ también resulta de congruencia (cf. ejercicio 3 de la sección 2) y, pensando a $f[\alpha]_k$ y $g[\alpha]_k$ en $\mathrm{M}_k(\Gamma_\alpha)$, se tiene que

$$\langle f[\alpha]_k, g[\alpha]_k\rangle = \langle f, g\rangle$$

y que $\langle f[\alpha]_k, g\rangle$ sólo depende de la coclase doble de α módulo Γ.

Ejercicios.

1. Sean $f \in \mathrm{M}_k(\Gamma_0(N), \chi)$ y $g \in \mathrm{M}_k(\Gamma_0(N), \mu)$ formas modulares para distintos caracteres χ, μ con al menos una de ellas cuspidal. Probar que $\langle f, g\rangle = 0$.

2.4. Operadores de Hecke. A continuación estudiaremos operadores análogos a los T_n de la sección 1.5, dando fórmulas que permitan calcularlos en q-expansiones.

Definición 2.14. Para $\alpha \in \mathrm{Gl}_2^+(\mathbb{Q})$, Γ_1 y Γ_2 dos subgrupos de congruencia, y $f \in \mathrm{M}_k(\Gamma_1)$, definimos el *operador de coclase doble* como

$$f[\Gamma_1\alpha\Gamma_2]_k := \sum_j f[\beta_j]_k,$$

donde los β_j son un conjunto de representantes de las órbitas de $\Gamma_1\backslash\Gamma_1\alpha\Gamma_2$, es decir

$$\Gamma_1\alpha\Gamma_2 = \cup_j \Gamma_1\beta_j.$$

El operador está bien definido, no depende de la elección de los representantes β_j, su imagen está contenida en $\mathrm{M}_k(\Gamma_2)$, manda formas cuspidales en formas cuspidales y además

1. si $\Gamma_2 \subseteq \Gamma_1$ y $\alpha = \mathrm{Id}$ se tiene la inclusión natural de $\mathrm{M}_k(\Gamma_1)$ en $\mathrm{M}_k(\Gamma_2)$,
2. si $\alpha^{-1}\Gamma_1\alpha = \Gamma_2$ se tiene $f[\Gamma_1\alpha\Gamma_2]_k = f[\alpha]_k$, el isomorfismo natural entre $\mathrm{M}_k(\Gamma_1)$ y $\mathrm{M}_k(\Gamma_2)$,
3. si $\Gamma_1 \subseteq \Gamma_2$ y $\alpha = \mathrm{Id}$ se tiene el *operador traza* que proyecta $\mathrm{M}_k(\Gamma_1)$ sobre $\mathrm{M}_k(\Gamma_2)$.

Observación 2.15. Todos los operadores de coclases dobles son una composición de operadores de estas formas. Dados Γ_1, Γ_2 y α, basta con tomar $\Gamma_3 = \alpha^{-1}\Gamma_1\alpha \cap \Gamma_2$, $\Gamma_3' = \alpha\Gamma_3\alpha^{-1}$ y los operadores correspondientes a $\Gamma_3' \subseteq \Gamma_1$, $\Gamma_3 = \alpha^{-1}\Gamma_3'\alpha$ y $\Gamma_3 \subseteq \Gamma_2$.

Como casos particulares tenemos los siguientes operadores

Definición 2.16. Sean $f \in \mathrm{M}_k(\Gamma_1(N))$ y $d \in (\mathbb{Z}/N\mathbb{Z})^*$. El *operador diamante* $\langle d\rangle$ se define como

$$\langle d\rangle : f \mapsto f[\Gamma_1(N)\alpha\Gamma_1(N)]_k$$

para cualquier $\alpha = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N)$.

Observación 2.17. Si $f \in \mathrm{M}_k(\Gamma_0(N), \chi)$, entonces $\langle d\rangle$ actúa como multiplicación por $\chi(d)$.

Definición 2.18. Sea N un entero positivo y p un número primo. El *operador de Hecke* T_p se define por la acción de $[\Gamma_1(N)\left(\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\right)\Gamma_1(N)]_k$ en $\mathrm{M}_k(\Gamma_1(N))$.

Observación 2.19. Se tiene que

$$T_p(f) = \begin{cases} \sum_{j=0}^{p-1} f[\left(\begin{smallmatrix} 1 & j \\ 0 & p \end{smallmatrix}\right)]_k & \text{si } p \nmid N \\ \sum_{j=0}^{p-1} f[\left(\begin{smallmatrix} 1 & j \\ 0 & p \end{smallmatrix}\right)]_k + f[\left(\begin{smallmatrix} m & n \\ N & p \end{smallmatrix}\right)\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)]_k & \text{si } p \mid N \end{cases} \tag{2.9}$$

donde m, n son enteros tales que $mp - Nn = 1$.

Observación 2.20. Cambiando $\Gamma_1(N)$ por $\Gamma_0(N)$ se definen los T_p de manera análoga y en la fórmula (2.9) se puede omitir $\left(\begin{smallmatrix} m & n \\ N & p \end{smallmatrix}\right)$ por estar en $\Gamma_0(N)$.

Definición 2.21. Los operadores de Hecke con índices compuestos se definen de manera inductiva por

$$T_{p^{r+1}} := T_p T_{p^r} - p^{k-1}\langle p\rangle T_{p^{r-1}} \tag{2.10}$$

para $r \geq 1$ y para m, n coprimos

$$T_{mn} := T_m T_n.$$

Observación 2.22. Sean n coprimo con N. La acción de T_n en una forma modular $f \in \mathrm{M}_k(\Gamma_0(N))$ no coincide con la acción de T_n en f considerada en $\mathrm{M}_k(\Gamma_0(nN))$.

Proposición 2.23. *Sea $f \in \mathrm{M}_k(\Gamma_0(N), \chi)$ una forma modular con expansión de Fourier $\sum_{n=0}^{\infty} a_n q^n$ y m un entero positivo. La acción del operador de Hecke T_m en los coeficientes de Fourier de f está dada por*

$$(T_m f)(\tau) = \sum_{n=0}^{\infty} b_n q^n,$$

donde

$$b_n = \sum_{1 \leq d \mid \gcd(m,n)} \chi(d) d^{k-1} a_{mn/d^2}. \tag{2.11}$$

Demostración. Supongamos $m = p$ un número primo. Para $0 \leq j < p$ tenemos

$$f[\left(\begin{smallmatrix} 1 & j \\ 0 & p \end{smallmatrix}\right)]_k(\tau) = p^{k-1}(0\tau + p)^{-k} f\left(\frac{\tau + j}{p}\right) = \frac{1}{p}\sum_{n=0}^{\infty} a_n q_p^n \zeta_p^{nj}$$

con $\zeta_p = e^{2\pi i/p}$ una raíz de la unidad de orden p. De donde

$$\sum_{j=0}^{p-1} f[\left(\begin{smallmatrix} 1 & j \\ 0 & p \end{smallmatrix}\right)]_k(\tau) = \sum_{n=0} a_{np} q^n.$$

Si $p \nmid N$ se tiene el término extra

$$f[\left(\begin{smallmatrix} m & n \\ N & p \end{smallmatrix}\right)\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)]_k(\tau) = (\langle d\rangle f)[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)]_k(\tau) = p^{k-1}\chi(d)\sum_{n=0}^{\infty} a_n q^{np},$$

de donde se prueba (2.11). El caso m compuesto sale por inducción de la definición 2.21. □

Observación 2.24. Sea N un entero positivo y $\chi : (\mathbb{Z}/N\mathbb{Z})^* \to \mathbb{C}^*$ un carácter de Dirichlet. Sea n un entero positivo coprimo con N. Si $f, g \in \mathrm{M}_k(\Gamma_0(N), \chi)$ (con al menos una cuspidal) entonces

$$\langle T_n f, g\rangle = \chi(n) \langle f, T_n g\rangle . \tag{2.12}$$

Tomando c_n cualquier raíz cuadrada de $\chi(n)$ se tiene

$$\langle c_n T_n f, g\rangle = \langle f, c_n T_n g\rangle . \tag{2.13}$$

Es decir, los operadores $c_n T_n$ de $\mathrm{S}_k(\Gamma_0(N), \chi)$ son hermitianos.

Proposición 2.25. *Sea N un entero positivo y $\chi : (\mathbb{Z}/N\mathbb{Z})^* \to \mathbb{C}^*$ un carácter de Dirichlet. El espacio $\mathrm{S}_k(\Gamma_0(N), \chi)$ admite una base de autovectores para todos los operadores de Hecke T_n con N y n coprimos.*

Demostración. Tales T_n resultan diagonalizables por la observación 2.24, y conmutan entre sí, de donde existe una base de autovectores comunes. □

Observación 2.26. Al ser $\mathrm{S}_k(\Gamma_0(N), \chi)$ de dimensión finita, sólo hace falta diagonalizar un número finito de operadores de Hecke.

Ejemplo 2.27. El espacio $\mathrm{S}_{28}(\mathrm{Sl}_2(\mathbb{Z}))$ tiene dimensión 2 y está generado por $f_1 = \Delta E_4^4$ y $f_2 = \Delta^2 E_4$.

Sus q-expansiones son

$$f_1 = q + 936q^2 + 331452q^3 + 53282368q^4 + O(q^5)$$

y

$$f_2 = q^2 + 192q^3 - 8280q^4 + O(q^5).$$

Se tiene que

$$T_2(f_1) = 936q + 187500096q^2 + O(q^3) = 936 f_1 + 18662400 f_2$$

y

$$T_2(f_1) = q + 8280q^2 + O(q^3) = f_1 - 9216 f_2.$$

Diagonalizando

$$\begin{pmatrix} 936 & 18662400 \\ 1 & -9216 \end{pmatrix}$$

obtenemos la siguiente base de autoformas

$$f_1 + (-5076 \pm 108\sqrt{18209}) f_2.$$

Ejercicios.

1. Probar que el cociente $Y_0(N) := \Gamma_0(N)\backslash\mathcal{H}$ clasifica clases de equivalencia de pares (Λ, S) con $\Lambda \subseteq \mathbb{C}$ un reticulado y $S \subseteq \mathbb{C}/L$ de orden N.
2. Probar que el cociente $Y_1(N) := \Gamma_1(N)\backslash\mathcal{H}$ clasifica clases de equivalencia de pares (Λ, P) con $\Lambda \subseteq \mathbb{C}$ un reticulado y $P \in \mathbb{C}/L$ de orden N.
3. Probar que los operadores T_{p^r} son polinomios en T_p y $\langle p\rangle$.
4. Sean $p \nmid N$ y m, n tales que $mp - nN = 1$. Probar que
$$\Gamma_1(N) \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right) \Gamma_1(N) = \Gamma_1(N) \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right) \Gamma_1(N) \left(\begin{smallmatrix} p & n \\ N & m \end{smallmatrix}\right).$$
5. Probar (2.12).
6. Sea ζ_n una raíz primitiva de la unidad de orden n. Probar que
$$\sum_{i=1}^{n} \zeta_n^k = \begin{cases} n & \text{si } n|k \\ 0 & \text{si } n \nmid k. \end{cases}$$

7. Sea $f(z) = \sum_{m=0}^{\infty} a_m z^m$ una función holomorfa en un entorno del cero y ζ_n como en el ejercicio anterior. Probar que

$$\frac{1}{n}\sum_{i=1}^{n} f(\zeta_n^k z) = \sum_{k=0}^{\infty} a_{nk} z^{nk}.$$

8. Probar que los operadores de Hecke preservan al espacio de formas cuspidales (comparar con ejercicio 17 de la sección 2).
9. Probar las afirmaciones de la primera oración del Ejemplo 2.27.
10. Probar que $\mathrm{S}_k(\mathrm{Sl}_2(\mathbb{Z}))$ tiene una base de autoformas cuyos coeficientes de Fourier son enteros algebraicos reales pertenecientes a una extensión finita de $\mathbb{Q}$.
11. Sea $f = \Delta E_4 \in S_{16}(\mathrm{Sl}_2(\mathbb{Z}))$. Probar que es un autovector para todos los T_n. Probar que no es autovector de T_2 en $M_k(\Gamma_0(2))$.
12. Sea $n \geq 1$ un entero y sea $f \in \mathrm{M}_k(\Gamma_0(N), \chi)$ un autovector para T_n. Supongamos que el término constante de la q-expansión de f es no nulo. Probar que el correspondiente autovalor de T_n es

$$\lambda_n = \sum_{1 \leq d | n} \chi(d) d^{k-1}.$$

13. Probar que existe una forma modular de peso 1 para $\Gamma_0(4)$ con carácter χ_4 cuya q-expansión comienza de la siguiente manera

$$\frac{1}{4} + q + q^2 + q^4 + 2q^5 + q^8 + \dots.$$

14. Probar que $\Theta^2 \in \mathrm{M}_1(\Gamma_0(4), \chi_4)$.
15. Probar (0.1) y (0.3).

3. Por último.

Mencionamos brevemente algunos temas relacionados a los vistos anteriormente, en los que no hemos tenido tiempo de profundizar. Mucho hay hecho sobre el problema de expresar un número como suma de cuadrados. Con técnicas similares se puede encarar el caso de una cantidad impar de cuadrados, considerando las llamadas *formas modulares de peso medio entero*. La definición es un poco más técnica, pero los resultados son similares. Un ejemplo de forma modular de peso medio entero es casualmente Θ y sus potencias impares. Para ver más sobre dicho problema recomendamos [5]. Un lindo texto introductorio sobre formas modulares de peso medio entero y curvas elípticas con multiplicación compleja es el [7], donde dichos temas son estudiados para encarar el problema de los *números congruentes* (i.e.: decidir si un número dado es o no el área de un triángulo rectángulo de lados racionales). Para una exposición accesible de la relación entre formas modulares, curvas elípticas y funciones L se puede consultar [8]. Un clásico bastante completo para estudiar formas modulares es [14], y uno más moderno es [4] donde, entre otras cosas, se abordan los teoremas de modularidad. Para saber más sobre los aspectos computacionales se puede consultar [15]. Recomendamos SAGE (`http://www.sagemath.org/`) como herramienta esencial. Puede usarse online aunque a la larga conviene instalarlo. En [2] hay muchísimos ejemplos de aplicaciones de formas modulares clásicas, formas de Siegel y de Hilbert. Para un recorrido interesante dentro de la Teoría de Números, pasando por la teoría de cuerpos de clases, funciones elípticas y formas modulares, el [3] no puede faltar. En [12] se construyen familias de grafos de Ramanujan, se estudia el problema de Rusiewisz sobre medidas invariantes de la esfera y el problema de Linnik sobre la distribución de representaciones de enteros como suma de tres cuadrados. Esperamos haber generado en el lector dudas suficientes como para seguir estudiando el tema.

Cerramos con algunos ejercicios surtidos para pensar en casa.

Ejercicios.

1. Probar que
$$\prod_{k=0}^{n-1}(1+q^k t)=\sum_{k=0}^{n} q^{\binom{k}{2}}\begin{bmatrix} n \\ k \end{bmatrix}_q t^k$$
donde $\begin{bmatrix} n \\ k \end{bmatrix}_q$ es el coeficiente q-binomial dado por la fórmula
$$\begin{bmatrix} n \\ k \end{bmatrix}_q=\frac{(1-q^n)(1-q^{n-1})\cdots(1-q^{n-k+1})}{(1-q)(1-q^2)\cdots(1-q^k)}.$$
2. Usando el ejercicio anterior, probar la siguiente identidad (*producto triple de Jacobi*)
$$\prod_{m=1}^{\infty}\left(1-x^{2m}\right)\left(1+x^{2m-1}y^2\right)\left(1+x^{2m-1}y^{-2}\right)=\sum_{n=-\infty}^{\infty}x^{n^2}y^{2n}.$$
3. Probar que
$$\Theta(\tau)=\frac{\eta^5(2\tau)}{\eta^2(\tau)\eta^2(4\tau)}.$$
4. Sean $\Theta_M(\tau):=\sum_{n\in\mathbb{Z}}(-1)^n q^{n^2}$ y $\Theta_F(\tau):=\sum_{m\in\mathbb{Z}+\frac{1}{2}} q^{m^2}$. Probar que
$$\Theta_M(\tau)=\frac{\eta^2(\tau)}{\eta(2\tau)}$$
y que
$$\Theta_F(\tau)=2\frac{\eta^2(4\tau)}{\eta(2\tau)}.$$
5. Probar que $\Theta^4=\Theta_M^4+\Theta_F^4$.
6. Sea $\Theta_Q(\tau):=\sum_{x,y\in\mathbb{Z}} q^{Q(x,y)}$ para $Q(x,y)=ax^2+bxy+cy^2$. Consideramos $Q_0(x,y)=x^2+xy+6y^2$, $Q_1(x,y)=2x^2+xy+3y^2$ y $Q_2(x,y)=2x^2-xy+3y^2$.
   a) Probar que cualquier forma $ax^2+bxy+cy^2$ con discriminante $D:=b^2-4ac$ igual a -23 es de la forma $Q_i\circ\gamma$ con $\gamma\in \mathrm{Sl}_2(\mathbb{Z})$ e $i=0,1$ o 2.
   b) Probar que $\Theta_{Q_1}=\Theta_{Q_2}$.
   c) Probar que
$$\frac{1}{2}\left(\Theta_{Q_0}+2\Theta_{Q_1}\right)=\frac{3}{2}+\sum_{n=1}^{\infty}\left(\sum_{d|n}\left(\frac{-23}{d}\right)\right)q^n$$
   donde $\left(\frac{-23}{d}\right)$ es el símbolo de Jacobi.
   d) Probar que
$$f:=\frac{1}{2}\left(\Theta_{Q_0}-\Theta_{Q_1}\right)$$
   es una autoforma normalizada de $\mathrm{S}_1(\Gamma_0(23),\left(\frac{-23}{\cdot}\right))$, por lo tanto coincide con $\eta(\tau)\eta(23\tau)$ (ver ejercicio 8 de la sección 2.2).
   e) Probar que la L-serie de la f anterior se factoriza como
$$L(f,s)=\prod_p\frac{1}{1-a_p p^{-s}+\left(\frac{-23}{p}\right)p^{-2s}}$$
   donde
$$a_p=\begin{cases}1 & \text{si } p=23,\\ 0 & \text{si } \left(\frac{p}{23}\right)=-1,\\ 2 & \text{si } \left(\frac{p}{23}\right)=1 \text{ y } p \text{ es de la forma } x^2+xy+6y^2,\\ -1 & \text{si } \left(\frac{p}{23}\right)=1 \text{ y } p \text{ es de la forma } 2x^2+xy+3y^2.\end{cases}$$

Referencias

[1] T. M. Apostol. *Modular functions and Dirichlet series in number theory*, Graduate Texts in Mathematics **41**. Springer-Verlag, New York, 1990.

[2] J. H. Bruinier, G. van der Geer, G. Harder, y D. Zagier. *The 1-2-3 of modular forms*, Universitext. Springer-Verlag, Berlin, 2008.

[3] D. Cox. *Primes of the form* $x^2 + ny^2$, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1989.

[4] F. Diamond, y J. Shurman. *A first course in modular forms*, Graduate Texts in Mathematics **228**. Springer-Verlag, New York, 2005.

[5] E. Grosswald. *Representations of integers as sums of squares*, Springer-Verlag, New York, 1985.

[6] F. Kirwan. *Complex algebraic curves*, London Mathematical Society Student Texts **23**. Cambridge University Press, Cambridge, 1992.

[7] N. Koblitz. *Introduction to elliptic curves and modular forms*, Graduate Texts in Mathematics **97**. Springer-Verlag, New York, 1984.

[8] Á. Lozano-Robledo. *Elliptic curves, modular forms, and their L-functions*, Student Mathematical Library **58**. American Mathematical Society, Providence, RI; Institute for Advanced Study (IAS), Princeton, NJ, 2011.

[9] R. Miranda. *Algebraic curves and Riemann surfaces*, Graduate Studies in Mathematics **5**. American Mathematical Society, Providence, RI, 1995.

[10] D. Mumford. *Algebraic geometry I*, Springer-Verlag, Berlin-New York, 1976.

[11] A. Pinkus y S. Zafrany. *Fourier series and integral transforms*, Cambridge University Press, Cambridge, 1977.

[12] P. Sarnak *Some applications of modular forms*, Cambridge Tracts in Mathematics **99**. Cambridge University Press, Cambridge, 1990.

[13] J.-P. Serre. *A course in arithmetic*, Springer-Verlag, New York-Heidelberg, 1973.

[14] G. Shimura. *Introduction to the arithmetic theory of automorphic functions*, Princeton University Press, Princeton, N.J., 1971.

[15] W. Stein. *Modular forms, a computational approach*, Graduate Studies in Mathematics **79**. American Mathematical Society, Providence, RI, 2007.

Universidad de Buenos Aires, FCEyN, Departamento de Matemáticas.
E-mail address: mmereb@gmail.com

LA FUNCIÓN ZETA DE DEDEKIND Y LA FÓRMULA PARA EL NÚMERO DE CLASES

NICOLÁS SIROLLI

RESUMEN. En estas notas damos una demostración de la fórmula del número de clases, que relaciona el residuo de la función Zeta de Dedekind de un cuerpo de números con algunos de sus invariantes aritméticos, y comparamos esta fórmula con la conjetura de Birch y Swinnerton-Dyer.

ÍNDICE

INTRODUCCIÓN

Una de las ideas más fructíferas de la teoría analítica de números consiste en asociarle una función generatriz al objeto que se quiera estudiar, y obtener información aritmética de este objeto a partir de información analítica de la función. El ejemplo paradigmático de esto es la función Zeta de Riemann

$$\zeta(s) = \prod_{p \text{ primo}} \left(1 - p^{-s}\right)^{-1},$$

cuyas propiedades analíticas dan información sobre la distribución de los números primos. Por ejemplo, el Teorema de los Números Primos equivale a que $\zeta(s) \neq 0$ si $\Re(s) = 1$.

En este curso estudiaremos la generalización de esta función a cuerpos de números K, llamada función Zeta de Dedekind y denotada ζ_K. Si bien esta función también da información sobre la distribución de los ideales primos de K, nos concentraremos en la fórmula para el número de clases (Teorema 2.1). Esta fórmula relaciona el residuo de ζ_K en $s = 1$ con los invariantes aritméticos más importantes del cuerpo. Entre ellos, el número de clases de K.

Date: 15 de julio de 2014.

Agradezco a Daniel Kohen y a Emilio Lauret por haberme ayudado a mejorar estas notas. También agradezco al Área de Matemática del PEDECIBA por financiar los gastos de mi traslado a La Falda.

La utilidad de esta fórmula reside en que, para ciertos cuerpos de números, dicho residuo se puede calcular de manera explícita en términos de sumas de Gauss, por lo que nos da una herramienta para calcular números de clases.

En lugar de profundizar sobre la fórmula en esa dirección, terminaremos el curso enunciando una fórmula similar: la conjetura de Birch y Swinnerton-Dyer, uno de los problemas del milenio. Esta conjetura relaciona ciertos invariantes algebraicos asociados a una curva elíptica con el primer coeficiente de la función generatriz correspondiente. El parecido entre las dos fórmulas no es casualidad: son ambas casos particulares de las conjeturas de Beilison y de Bloch-Kato (ver [Sch88]).

En la primera sección de estas notas daremos los resultados básicos de la teoría algebraica de números que se precisan para definir los invariantes involucrados en la fórmula para el número de clases. Algunas referencias sobre este tema son, por ejemplo, [Mar77, Capítulos 2 y 5] y [Neu99, Capítulo 1].

En la segunda sección definimos la función Zeta de Dedekind, y enunciamos y demostramos el Teorema 2.1, siguiendo de cerca a [Mar77, Capítulos 6 y 7]. Tanto en [Mar77] como en [BS66, Capítulo 5] se pueden encontrar cálculos explícitos del residuo en el caso de cuerpos cuadráticos y cuerpos ciclotómicos.

En la tercera sección enunciamos la conjetura de Birch y Swinnerton-Dyer tras introducir (muy informalmente) los invariantes involucrados en esta, y nos ocupamos de comparar estos términos con los que aparecen en la fórmula para el número de clases. Se puede consultar a [Sil09] sobre la teoría básica de curvas elípticas, y a [Dar09], [Wil06] para profundizar sobre la conjetura de Birch y Swinnerton-Dyer.

1. Resultados básicos de la teoría algebraica de números

Un *cuerpo de números* es una extensión finita $K/\mathbb{Q}$. Algunos ejemplos interesantes de cuerpos de números son:

- $K = \mathbb{Q}(\sqrt{m})$, con $m \in \mathbb{Z} \setminus \{0, 1\}$ libre de cuadrados. Estos son los llamados *cuerpos cuadráticos*.
- $K = \mathbb{Q}(\xi_n)$, con $\xi_n \in \mathbb{C}$ una raíz n-ésima primitiva de la unidad. Estos son los llamados *cuerpos ciclotómicos*.

Los cuerpos cuadráticos son, en términos del grado sobre $\mathbb{Q}$, los primeros cuerpos de números no triviales; a pesar de ello, ilustran la teoría que desarrollaremos sobradamente. Los cuerpos ciclotómicos son de interés, por ejemplo, porque toda extensión abeliana y finita de $\mathbb{Q}$ está contenida en uno de ellos.

Fijemos un cuerpo de números K, y denotemos por $d = [K : \mathbb{Q}]$ a la dimensión de K como $\mathbb{Q}$-espacio vectorial.

1.1. Aritmética de los cuerpos de números. El rol que juega $\mathbb{Z}$ como subanillo de $\mathbb{Q}$ es reemplazado en el cuerpo de números K por el *anillo de enteros*, que se define por

$$\mathcal{O}_K = \{\xi \in K : \exists\, f \in \mathbb{Z}[X] \text{ mónico tal que } f(\xi) = 0\}.$$

Este conjunto es en efecto un anillo, y sus elementos tienen norma y traza en $\mathbb{Z}$.

A los $\mathcal{O}_K$-módulos $\mathfrak{a} \subseteq K$ de tipo finito los llamaremos *ideales fraccionarios*. Aquellos que estén contenidos en $\mathcal{O}_K$ serán llamados *ideales*, a secas; son precisamente los ideales del anillo $\mathcal{O}_K$. El ideal $\{0\}$ quedará excluido de todo lo que sigue.

Proposición 1.1. *$\mathcal{O}_K$ es un $\mathbb{Z}$-módulo libre de rango d. Más aún, todo ideal fraccionario de K lo es.*

Gracias a este resultado, podemos introducir el *discriminante* de K. Se define como el determinante de la forma bilineal en $\mathcal{O}_K$ dada por $(\xi_1, \xi_2) \mapsto \mathrm{Tr}_{K/\mathbb{Q}}(\xi_1\xi_2)$, y se denota por $\mathrm{disc}(K)$.

Si bien el anillo de enteros no necesariamente es un dominio de factorización única (lo cual, en este caso, equivale a ser un dominio de ideales principales), sí se obtiene un resultado satisfactorio al considerar ideales en lugar de elementos. Más precisamente, se tiene el siguiente teorema.

Teorema 1.2. *Todo ideal fraccionario $\mathfrak{a} \subseteq K$ se escribe, de manera única, como*

$$\mathfrak{a} = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_n^{e_n},$$

donde los $\mathfrak{p}_i$ son ideales primos de $\mathcal{O}_K$, y los e_i son enteros no nulos.

Es decir, el grupo de ideales fraccionarios de K, que denotamos por $\mathrm{Frac}(K)$, es el grupo abeliano libre generado por los ideales primos de $\mathcal{O}_K$.

La medida de cuán lejos está el anillo de ser enteros de ser un dominio de ideales principales está dada por el *grupo de clases*, definido por

$$Cl(K) = \mathrm{Frac}(K)/P(K),$$

donde $P(K) = \{\mathfrak{a} \in \mathrm{Frac}(K) : \exists \xi \in K^\times \text{ tal que } \mathfrak{a} = (\xi)_{\mathcal{O}_K}\}$ es el subgrupo de los ideales fraccionarios principales. El grupo de clases es finito y su orden, que llamamos *número de clases*, se denota por h_K.

Dado un ideal fraccionario $\mathfrak{a}$, definimos su *norma* por $N(\mathfrak{a}) = [\mathcal{O}_K : \mathfrak{a}]$ (el índice de $\mathfrak{a}$ en $\mathcal{O}_K$). La función norma satisface las siguientes propiedades:

- $N(\mathfrak{ab}) = N(\mathfrak{a})N(\mathfrak{b}) \quad \forall\, \mathfrak{a}, \mathfrak{b} \in \mathrm{Frac}(K)$.
- $N((\xi)_{\mathcal{O}_K}) = |N_{K/\mathbb{Q}}(\xi)| \quad \forall\, \xi \in K^\times$.

1.2. Geometría de los cuerpos de números. A través de las inmersiones $K \hookrightarrow \mathbb{C}$, podemos utilizar herramientas de la geometría euclídea para estudiar a K, idea que se debe a Minkowski.

Distinguimos dos tipos de inmersiones.

- Las *reales*, aquellas $\sigma : K \hookrightarrow \mathbb{C}$ tales que $\sigma(K) \subseteq \mathbb{R}$.
- Las *complejas*, aquellas $\tau : K \hookrightarrow \mathbb{C}$ tales que $\tau(K) \not\subseteq \mathbb{R}$. Estas se pueden agrupar de a pares $(\tau, \bar{\tau})$.

Denotaremos entonces a las inmersiones de K por

$$\sigma_1, \ldots, \sigma_r, \tau_1, \bar{\tau_1}, \ldots, \tau_s, \bar{\tau_s},$$

con las σ_i reales y las τ_j complejas[1]. Aquí $r, s \in \mathbb{Z}_{\geq 0}$, y $r + 2s = d$.

A través de estas inmersiones definimos

$$\begin{aligned} \iota : K &\hookrightarrow \mathbb{R}^r \times \mathbb{C}^s \simeq \mathbb{R}^d, \\ \xi &\mapsto (\sigma_1(\xi), \ldots, \sigma_r(\xi), \tau_1(\xi), \ldots, \tau_s(\xi)). \end{aligned}$$

En $\mathbb{R}^r \times \mathbb{C}^s$ definimos una función *norma* por $N(x, z) = \prod_{i=1}^r x_i \cdot \prod_{j=1}^s z_j\overline{z_j}$. De esta manera, se tiene que $N(\iota(\xi)) = N_{K/\mathbb{Q}}(\xi)$ para todo $\xi \in K$.

Proposición 1.3. *$\Lambda_{\mathcal{O}_K} := \iota(\mathcal{O}_K)$ es un retículo completo en $\mathbb{R}^d$. Más aún,*

$$\sqrt{|\operatorname{disc}(K)|} = 2^s \cdot \mathrm{vol}(\mathbb{R}^d/\iota(\mathcal{O}_K)). \tag{1.1}$$

[1]Usaremos la letra s tanto para denotar a la cantidad de inmersiones complejas de K como para denotar a la variable compleja de nuestras funciones generatrices. El lector sabrá disculparnos.

Consideremos el grupo de unidades $\mathcal{O}_K^\times$. Denotemos por $T(\mathcal{O}_K^\times)$ al subgrupo de elementos de torsión de este grupo. Notemos que $T(\mathcal{O}_K^\times)$ coincide con el grupo de raíces de la unidad de K. Denotamos el orden de este grupo por ω_K.

Se puede obtener un resultado similar a la Proposición 1.3 mediante la utilización de logaritmos. Para esto, consideramos el morfismo de grupos

$$\begin{aligned}\log :&(\mathbb{R}^\times)^r \times (\mathbb{C}^\times)^s \to \mathbb{R}^{r+s}\\ &(x,z) \mapsto (\log(|x_1|),\dots,\log(|x_r|),2\log(|z_1|),\dots,2\log(|z_s|)),\end{aligned}$$

y denotamos $L = \log \circ \iota : K^\times \to \mathbb{R}^{r+s}$.

Como $N_{K/\mathbb{Q}}(\mathcal{O}_K^\times) \subseteq \mathbb{Z}^\times$, resulta que $L(\mathcal{O}_K^\times) \subseteq H$, donde H es el hiperplano dado por $H = \{v \in \mathbb{R}^{r+s} : \sum_k v_k = 0\}$. Además, como todo conjunto acotado de $\mathbb{R}^{r+s}$ tiene preimagen por L finita en $\mathcal{O}_K \setminus \{0\}$, se tiene que $\ker L = T(\mathcal{O}_K^\times)$.

Teorema 1.4 (Dirichlet)**.** *$\Lambda_{\mathcal{O}_K^\times} := L(\mathcal{O}_K^\times)$ es un retículo completo en H, y por lo tanto $\mathcal{O}_K^\times$ es producto directo de $T(\mathcal{O}_K^\times)$ y de un grupo abeliano libre de rango $r+s-1$.*

Este teorema nos permite definir otro de los invariantes importantes del cuerpo de números: el *regulador* de K, que se denota por $\operatorname{reg}(K)$ y está dado por

$$\operatorname{reg}(K) = \frac{\operatorname{vol}(H/\Lambda_{\mathcal{O}_K^\times})}{\sqrt{r+s}}. \tag{1.2}$$

Si $r+s=1$, por convención ponemos $\operatorname{reg}(K)=1$.

1.3. Ejercicios.

1. Sea $K = \mathbb{Q}(i)$.
   a) Probar que $\mathcal{O}_K = \mathbb{Z}[i]$, y calcular $\operatorname{disc}(K)$.
   b) Probar que $\mathbb{Z}[i]$ es un dominio euclídeo, utilizando como función "grado" a $\xi \mapsto N_{\mathbb{Q}(i)/\mathbb{Q}}(\xi)$.
   c) Probar que $T(\mathbb{Z}[i]^\times) = \{1,-1,i,-i\}$.
2. Sea p un primo distinto de 2. Probar que son equivalentes:
   a) p es reducible en $\mathbb{Z}[i]$.
   b) $p = a^2 + b^2$ con $a, b \in \mathbb{Z}$.
   c) -1 es un cuadrado en $\mathbb{Z}/p\mathbb{Z}$.
   d) $p \equiv 1 \mod 4$.
3. Sea $\mathfrak{p} = (1+i)$. Probar que $(2) = \mathfrak{p}^2$ es la factorización de (2) como producto de ideales primos de $\mathbb{Z}[i]$.
4. Sea p un primo distinto de 2.
   a) Probar que si $p \equiv 1 \mod 4$, entonces $(p) = \mathfrak{p}_1\mathfrak{p}_2$ con $\mathfrak{p}_1, \mathfrak{p}_2$ ideales primos (distintos) de $\mathbb{Z}[i]$.
   b) Probar que si $p \equiv 1 \mod 3$, entonces (p) es un ideal primo de $\mathbb{Z}[i]$.

2. La función Zeta de Dedekind y la fórmula para el número de clases

Sea K un cuerpo de números. Definimos la *función Zeta de Dedekind* de K por

$$\zeta_K(s) = \sum_{\mathfrak{a}} \frac{1}{N(\mathfrak{a})^s} \qquad (\Re(s) > 1), \tag{2.1}$$

donde $\mathfrak{a}$ recorre todos los ideales de $\mathcal{O}_K$. Para $K = \mathbb{Q}$, recuperamos la función Zeta de Riemann.

Al menos formalmente, el Teorema 1.2 nos permite escribir

$$\zeta_K(s) = \prod_{\mathfrak{p}} \left(1 - N(\mathfrak{p})^{-s}\right)^{-1}, \tag{2.2}$$

donde $\mathfrak{p}$ recorre todos los ideales primos de $\mathcal{O}_K$. Es decir, tal como le pasa a la función Zeta de Riemann, podemos desarrollar a ζ_K como un producto de Euler.

Veremos más adelante que (2.1) define una función holomorfa. Nuestro objetivo es probar el siguiente resultado.

Teorema 2.1. *La función ζ_K se puede extender de manera holomorfa a $\Re(s) > 1-1/d$, salvo por un polo simple en $s = 1$. El residuo en $s = 1$ está dado por*

$$\operatorname{Res}(\zeta_K, 1) = \frac{2^r(2\pi)^s}{\sqrt{|\operatorname{disc}(K)|}} \cdot h_K \cdot \frac{\operatorname{reg}(K)}{\omega_K}. \tag{2.3}$$

Probaremos un resultado algo más fuerte, que además muestra que los ideales están equidistribuidos en las clases de $Cl(K)$.

Teorema 2.2. *Sea κ la constante dada por*

$$\kappa = \frac{2^r(2\pi)^s}{\sqrt{|\operatorname{disc}(K)|}} \cdot \frac{\operatorname{reg}(K)}{\omega_K}.$$

Dada $\mathcal{C} \in Cl(K)$, consideremos la función $i_{\mathcal{C}}$ dada por

$$i_{\mathcal{C}}(t) = \#\left\{\mathfrak{a} \subseteq \mathcal{O}_K : \mathfrak{a} \in \mathcal{C}, N(\mathfrak{a}) \leq t\right\} \qquad (t \in \mathbb{R}_{\geq 0}).$$

Entonces, $i_{\mathcal{C}}(t) = \kappa t + O(t^{1-1/d})$.

Veamos cómo de este resultado se sigue el Teorema 2.1.

Para $n \in \mathbb{N}$, sea $j_n = \#\{\mathfrak{a} \subseteq \mathcal{O}_K : N(\mathfrak{a}) = n\}$. Reescribamos a ζ_K como la serie de Dirichlet

$$\zeta_K(s) = \sum_{n\geq 1} \frac{j_n}{n^s} = \sum_{n\geq 1} \frac{j_n - h_K\kappa}{n^s} + h_K\kappa\zeta(s). \tag{2.4}$$

Usaremos que el Teorema 2.1 es conocido para $K = \mathbb{Q}$. Esto es, la función ζ de Riemann $\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s}$ converge uniformemente en compactos de $\Re(s) > 1$ (ver Lema 2.3), y se puede extender de manera holomorfa a $\Re(s) > 0$, salvo por un polo simple en $s = 1$ en el que se tiene que $\operatorname{Res}(\zeta, 1) = 1$.

En cuanto al primer sumando del miembro derecho de (2.4), el Teorema 2.2 nos dice que

$$\sum_{n\leq t} j_n - \kappa h_K = \left(\sum_{\mathcal{C}\in Cl(K)} i_{\mathcal{C}}(t)\right) - \kappa h_K \lfloor t \rfloor = O(t^{1-1/d}).$$

Entonces, el Teorema 2.1 se sigue de (2.4) más el siguiente resultado sobre la convergencia de series de Dirichlet.

Lema 2.3. *Sea $(a_n)_{n\geq 1} \subseteq \mathbb{C}$ una sucesión tal que $\sum_{n\leq t} a_n = O(t^\alpha)$ para algún $\alpha \in \mathbb{R}$. Entonces, la serie de Dirichlet $\sum_{n\geq 1} \frac{a_n}{n^s}$ converge uniformemente en compactos de $\Re(s) > \alpha$.*

Demostración. Fijemos constantes $A, \varepsilon > 0$, y tomemos $s \in \mathbb{C}$ con $\alpha + \varepsilon \leq \Re(s) \leq A$. Denotemos $A_n = \sum_{k \leq n} a_k$. Dados $m, M \in \mathbb{N}$ con $m \leq M$ tenemos que

$$\sum_{n=m}^{M} \frac{a_n}{n^s} = \sum_{n=m}^{M} \frac{A_n}{n^s} - \sum_{n=m}^{M} \frac{A_{n-1}}{n^s} = \frac{A_M}{M^s} - \frac{A_{m-1}}{(m-1)^s} + \sum_{n=m}^{M-1} A_n \left(\frac{1}{n^s} - \frac{1}{(n+1)^s} \right).$$

Por hipótesis, existe $C > 0$ tal que $|A_n| \leq C n^\alpha$ para todo $n \in \mathbb{N}$. Por otra parte, se tiene que

$$\left| \frac{1}{n^s} - \frac{1}{(n+1)^s} \right| = \left| s \int_n^{n+1} \frac{dt}{t^{s+1}} \right| \leq \frac{A}{n^{\Re(s)+1}}.$$

Entonces

$$\lim_{m,M \to \infty} \left| \sum_{n=m}^{M} \frac{a_n}{n^s} \right| \leq \lim_{m,M \to \infty} C \left(\frac{1}{M^\varepsilon} + \frac{1}{(m-1)^\varepsilon} + A \sum_{n=m}^{M} \frac{1}{n^{1+\varepsilon}} \right) = 0,$$

de lo que se sigue el resultado. □

Observación 2.4. Como la serie de Dirichlet en (2.4) converge en $\Re(s) > 1$, lo hace absolutamente (!). Esto le da sentido a la expresión de ζ_K dada en (2.1), ya que los ideales $\mathfrak{a}$ sobre los que se suma no están ordenados a priori.

Comencemos con la demostración del Teorema 2.2. Tomemos una clase $\mathcal{C} \in Cl(K)$ y fijemos $\mathfrak{b} \in \mathcal{C}^{-1}$ un ideal (entero). Gracias a la biyección

$$\begin{aligned} \#\{\mathfrak{a} \subseteq \mathcal{O}_K : \mathfrak{a} \in \mathcal{C}, N(\mathfrak{a}) \leq t\} &\xrightarrow{\simeq} \{(\xi)_{\mathcal{O}_K} \subseteq \mathfrak{b} : |N_{K/\mathbb{Q}}(\xi)| \leq t N(\mathfrak{b})\}, \\ \mathfrak{a} &\mapsto \mathfrak{a}\mathfrak{b}, \end{aligned}$$

pasamos de tener que contar *ideales* a tener que contar *elementos*, ya que nos dice que

$$i_{\mathcal{C}}(t) = \#\left\{\xi \in \mathfrak{b} : |N_{K/\mathbb{Q}}(\xi)| \leq t N(\mathfrak{b})\right\} / \mathcal{O}_K^\times.$$

Equivalentemente, si usando el Teorema 1.4 escribimos $\mathcal{O}_K^\times = T(\mathcal{O}_K^\times) \cdot V$ con V un grupo abeliano libre de rango $r + s - 1$, nos dice que

$$\omega_K \cdot i_{\mathcal{C}}(t) = \#\left\{\xi \in \mathfrak{b} : |N_{K/\mathbb{Q}}(\xi)| \leq t N(\mathfrak{b})\right\} / V.$$

Podemos calcular el miembro derecho de esta igualdad a través de la geometría, si hallamos un dominio fundamental D para la acción de $\iota(V)$ en $(\mathbb{R}^\times)^r \times (\mathbb{C}^\times)^s$ y contamos los puntos de $\iota(\mathfrak{b})$ que estén en D. Siendo $\log|_{\iota(V)}$ un monomorfismo, se tiene que si D' es un dominio fundamental para la acción de $\Lambda_{\mathcal{O}_K^\times}$ en $\mathbb{R}^{r+s}$, entonces podemos tomar

$$D = \{(x, z) \in (\mathbb{R}^\times)^r \times (\mathbb{C}^\times)^s : \log((x, z)) \in D'\}.$$

Como $\Lambda_{\mathcal{O}_K^\times}$ es un retículo completo en H, podemos tomar $D' = P \oplus \mathbb{R} \cdot v_0$, con P el paralelogramo fundamental para $\Lambda_{\mathcal{O}_K^\times}$ y $v_0 \in \mathbb{R}^{r+s} \setminus H$. De hecho, tomamos

$$v_0 = (\underbrace{1, \ldots, 1}_{r \text{ veces}}, \underbrace{2, \ldots, 2}_{s \text{ veces}})$$

porque así D resulta *homogéneo* (es decir, satisface que $D = \lambda D$ para todo $\lambda \in \mathbb{R}^\times$).

Para $a > 0$, denotemos $D_a = \{(x, z) \in D : |N(x, z)| \leq a\}$. Entonces, por la homogeneidad de D tenemos que

$$\omega_K \cdot \iota_{\mathcal{C}}(t) = \#\iota(\mathfrak{b}) \cap D_{tN(\mathfrak{b})} = \#\iota(\mathfrak{b}) \cap \sqrt[d]{tN(\mathfrak{b})} D_1. \tag{2.5}$$

Para calcular el miembro derecho de (2.5), utilizaremos el siguiente resultado, de naturaleza puramente geométrica. Diremos que un conjunto $B \subseteq \mathbb{R}^d$ tiene borde *suficientemente lindo* si $\partial B \subseteq \cup_{i\in I} f_i([0,1]^{d-1})$, donde las $f_i : [0,1]^{d-1} \to \mathbb{R}^d$ son funciones Lipschitz y el conjunto I es finito.

Proposición 2.5. *Sea $\Lambda \subseteq \mathbb{R}^d$ un retículo completo, y sea $B \subseteq \mathbb{R}^d$ un conjunto acotado y medible. Si el borde de B es suficientemente lindo, entonces*

$$\#\Lambda \cap aB = \frac{|B|}{\operatorname{vol}(\mathbb{R}^d/\Lambda)} \cdot a^d + O(a^{d-1}) \qquad (a > 0).$$

Demostración. Tomemos un isomorfismo $\varphi : \mathbb{R}^d \to \mathbb{R}^d$ tal que $\varphi(\Lambda) = \mathbb{Z}^d$. Entonces $\varphi(B)$ también tiene borde suficientemente lindo. Además, por ser φ lineal, se tiene que $|B| = |\varphi(B)| \cdot \operatorname{vol}(\mathbb{R}^d/\Lambda)$. Esto prueba que podemos suponer $\Lambda = \mathbb{Z}^d$.

Consideremos las traslaciones del cubo $[0,1]^d$ con centros en los puntos de $\mathbb{Z}^d$. Los llamaremos *d-cubos*. Denotemos por N_a a la cantidad de d-cubos que intersecan a $\partial(aB)$. Entonces

$$\big|\#\mathbb{Z}^d \cap aB - \text{cantidad de } d\text{-cubos contenidos en } aB\big| \le N_a, \quad \text{y}$$
$$\big||aB| - \text{cantidad de } d\text{-cubos contenidos en } aB\big| \le N_a,$$

por lo que, siendo $|aB| = a^d|B|$, basta con probar que $N_a = O(a^{d-1})$.

Por otra parte,

$$\partial B \subseteq \bigcup_{i\in I} f_i([0,1]^{d-1}) \Longrightarrow \partial(aB) \subseteq \bigcup_{i\in I} a \cdot f_i([0,1]^{d-1}).$$

Esto nos permite suponer que $\partial(aB) = a \cdot f([0,1]^{d-1})$ con f una función Lipschitz.

Subdividamos, de la manera natural, al cubo $[0,1]^{d-1}$ en $\lfloor a \rfloor^{d-1}$ pequeños cubos de lado $\frac{1}{\lfloor a \rfloor}$. Sea C uno de estos cubos. C tiene diámetro igual a $\frac{\sqrt{d-1}}{\lfloor a \rfloor}$, por lo que si λ es la constante Lipschitz de f, entonces $f(C)$ tiene diámetro acotado por $\frac{\lambda\sqrt{d-1}}{\lfloor a \rfloor}$. Por lo tanto, suponiendo que $a \ge 1$, resulta que $a \cdot f(C)$ tiene diámetro acotado por $2\lambda\sqrt{d-1}$. Si denotamos

$$M = \left(2 + 2\left\lceil 2\lambda\sqrt{d-1}\right\rceil\right)^n,$$

esto implica que $a \cdot f(C)$ interseca a lo sumo M de los d-cubos. Entonces, como la cantidad de cubos pequeños es $\lfloor a \rfloor^{d-1}$, tenemos que $N_a = O(\lfloor a \rfloor^{d-1}) = O(a^{d-1})$, lo que termina la demostración. □

Aplicando este resultado a $\Lambda = \iota(\mathfrak{b})$ y $B = D_1$, de (2.5) se sigue que

$$\omega_K \cdot \iota_{\mathcal{C}}(t) = \frac{|D_1|}{\operatorname{vol}(\mathbb{R}^d/\iota(\mathfrak{b}))} \cdot tN(\mathfrak{b}) + O(t^{1-1/d}) = \frac{2^s|D_1|}{\sqrt{|\operatorname{disc}(K)|}} \cdot t + O(t^{1-1/d}),$$

donde para obtener la última igualdad usamos que dados dos retículos completos $\Lambda \subseteq \Lambda' \subseteq \mathbb{R}^d$ se tiene que $\operatorname{vol}(\mathbb{R}^d/\Lambda) = [\Lambda' : \Lambda] \operatorname{vol}(\mathbb{R}^d/\Lambda')$, junto con (1.1). Entonces, el Teorema 2.2 quedará demostrado una vez que probemos el siguiente resultado.

Lema 2.6. *D_1 tiene borde suficientemente lindo, y $|D_1| = 2^r\pi^s \operatorname{reg}(K)$.*

Demostración. Como D_1 es simétrico respecto al 0 en su primera coordenada, basta con probar que

$$D_1^+ = \{(x,z) \in D_1 : x_1, \dots, x_r \ge 0\}$$

tiene borde suficientemente lindo y $|D_1^+| = \pi^s \operatorname{reg}(K)$, ya que $|D_1| = 2^r|D_1^+|$.

Para probar ambas afirmaciones, parametrizaremos a D_1^+. Empecemos notando que

$$D_1 = \{(x,z) \in (\mathbb{R}^\times)^r \times (\mathbb{C}^\times)^s : \log(x,z) \in P \oplus (-\infty, 0] \cdot v_0\},$$

ya que $|N(x,z)| \leq 1$ si y solo si $\sum \log(x,z) \leq 0$, y las coordenadas de los puntos de P suman cero. Tomemos $v_1, \ldots, v_{r+s-1}$ base del retículo $\Lambda_{\mathcal{O}_K^\times}$. Escribamos $v_i = (v_i^{(1)}, \ldots, v_i^{(r+s)})$. Así, $(x,z) \in D_1^+$ si y solo si

$$\log(x_i) = \sum_{k=1}^{r+s-1} t_k v_k^{(i)} + u \qquad (1 \leq i \leq r),$$

$$2\log(|z_j|) = \sum_{k=1}^{r+s-1} t_k v_k^{(r+j)} + 2u \qquad (1 \leq j \leq s),$$

con $0 \leq t_k < 1$ y $-\infty < u \leq 0$. Pongamos $t_{r+s} = e^u$. Así, tenemos que $(x,z) \in D_1^+$ si y solo si

$$x_i = t_{r+s} \cdot \exp\left(\sum_{k=1}^{r+s-1} t_k v_k^{(i)}\right) \qquad (1 \leq i \leq r),$$

$$z_j = t_{r+s} \cdot \exp\left(\frac{1}{2}\sum_{k=1}^{r+s-1} t_k v_k^{(r+j)} + 2\pi i t_{r+s+j}\right), \qquad (1 \leq j \leq s),$$

con $t_{r+s} \in (0,1]$ y $t_k \in [0,1)$ para $1 \leq k \leq d, k \neq r+s$. La función $f : [0,1]^d \to \mathbb{R}^r \times \mathbb{C}^s$ dada por $t \mapsto (x,z)$ es Lipschitz, y se puede ver que $f([0,1]^d) = \overline{D_1^+}$ y $f(\partial([0,1]^d)) = \partial D_1^+$. Esto que muestra que D_1^+ es acotado, medible y tiene borde suficientemente lindo. Finalmente, utilizando esta parametrización se obtiene que $|D_1^+| = \pi^s \operatorname{reg}(K)$ (ver los ejercicios al final de esta sección). □

2.1. Ejercicios.

1. Sea $m \in \mathbb{N}$, y sea $\chi : (\mathbb{Z}/m\mathbb{Z})^\times \to \mathbb{C}^\times$ un morfismo. Este induce naturalmente una función en los enteros coprimos con m, que extendemos por 0 a todos los enteros; la denotamos también por χ. Definimos la *L-serie* de χ por

$$L(s,\chi) = \sum_{n \geq 1} \frac{\chi(n)}{n^s}.$$

   a) Probar que si el morfismo χ es no trivial, $L(s,\chi)$ define una función holomorfa en $\Re(s) > 0$. ¿Qué se puede decir cuando χ es trivial?
   b) Probar que

$$L(s,\chi) = \prod_{p \text{ primo}} \left(1 - \chi(p)p^{-s}\right)^{-1}.$$

2. Sea $K = \mathbb{Q}(i)$. Sea χ el carácter no trivial de $(\mathbb{Z}/4\mathbb{Z})^\times$. Probar que[2]

$$\zeta_K(s) = (1 - 2^{-s})^{-1} \cdot \zeta(s) \cdot L(s,\chi).$$

3. Probar la fórmula de Leibniz

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots$$

[2]Notar que $(\mathbb{Z}/4\mathbb{Z})^\times \simeq \operatorname{Gal}(K/\mathbb{Q})$. Se puede obtener una fórmula similar a esta para cuerpos ciclotómicos cualesquiera.

4. En este ejercicio probaremos que $|D_1^+| = \pi^s \operatorname{reg}(K)$, completando así la demostración del Lema 2.6.
   a) Sea $\tilde{f} : (0,1)^d \to \mathbb{R}^r \times \mathbb{C}^s$ la restricción a $(0,1)^d$ de la función f introducida en la demostración de dicho lema. Verificar que $\tilde{f}$ es la composición de las funciones

   $$(0,1)^d \xrightarrow{f_1} \mathbb{R}^d \xrightarrow{f_2} \mathbb{R}^d = \mathbb{R}^r \times \mathbb{R}^s \times \mathbb{R}^s \xrightarrow{f_3} \mathbb{R}^r \times \mathbb{R}^s \times \mathbb{R}^s \xrightarrow{f_4} \mathbb{R}^r \times \mathbb{C}^s,$$

   donde:
   - $f_1(t) = (t_1, \dots, t_{r+s-1}, \log(t_{r+s}), t_{r+s+1}, \dots, t_d)$.
   - f_2 es la multiplicación a derecha por la matriz por bloques

   $$M = \left(\begin{array}{c|c} \begin{matrix} v_1 \\ \vdots \\ v_{r+s-1} \\ v_0 \end{matrix} & \begin{matrix} \\ 0 \\ \\ \end{matrix} \\ \hline 0 & I_s \end{array}\right) \in \mathbb{R}^{d\times d}.$$

   - $f_3(\alpha, \beta, \gamma) = (e^{\alpha_1}, \dots, e^{\alpha_r}, \frac{1}{2}e^{\beta_1}, \dots, \frac{1}{2}e^{\beta_s}, 2\pi\gamma_1, \dots, 2\pi\gamma_s)$.
   - $f_4(x, \rho, \theta) = (x, \rho_1 e^{i\theta_1}, \dots, \rho_s e^{i\theta_s})$.

   b) Deducir que $\tilde{f}$ es abierta e inyectiva, y probar que $f(\partial([0,1]^d)) = \partial D_1^+$.
   c) Sea $g = f_3 \circ f_2 \circ f_1$. Probar que

   $$\det Dg(t) = \frac{\pi^s \det(M) \prod_{i=1}^r x_i(t) \cdot \prod_{j=1}^s \rho_j(t)}{t_{r+s}}.$$

   d) Deducir que

   $$\left|D_1^+\right| = \int_{[0,1]^d} |\det Dg(t)| \prod_{j=1}^s \rho_j(t)\, dt = \pi^s \frac{|\det(M)|}{d}.$$

   e) Sean A, B dos matrices cuadradas cuyas filas, salvo tal vez la última, son vectores cuyas coordenadas suman cero. Probar que si las últimas filas de A y de B son vectores cuyas coordenadas suman lo mismo, entonces $\det A = \det B$.
   f) Deducir que $\frac{|\det(M)|}{d} = \operatorname{reg}(K)$.

3. Una fórmula similar: la conjetura de Birch y Swinnerton-Dyer

Sea K un cuerpo de números. Una *curva elíptica* E/K es una curva sobre K, no singular y de género 1, junto con un punto K-racional $O \in E(K)$. Toda tal curva puede ser descrita por una *ecuación de Weierstrass*

$$E: \quad y^2 = x^3 + Ax + B, \tag{3.1}$$

con $A, B \in K$ satisfaciendo $-16(4A^3 + 27B^2) \neq 0$.

Comenzaremos definiendo la función generatriz correspondiente. La idea básica es reducir la ecuación (3.1) módulo $\mathfrak{p}$ para cada primo $\mathfrak{p}$ de K, y para aquellos primos $\mathfrak{p}$ para los cuales se obtenga una curva elíptica $\overline{E}$ sobre $k_\mathfrak{p} = \mathcal{O}_K/\mathfrak{p}$ (los llamados primos de *buena reducción*), contar la cantidad de puntos de $\overline{E}(k_\mathfrak{p})$. Hacer dicha reducción requiere de algún cuidado; sólo diremos que el conjunto de primos de mala reducción es finito y se puede determinar con precisión.

Se la llama *L-serie* asociada a E/K, y se define como un producto de Euler:

$$L(E/K, s) = \prod_{\mathfrak{p} \text{ primo}} L_\mathfrak{p}(N(\mathfrak{p})^{-s})^{-1} \qquad (s \in \mathbb{C}), \tag{3.2}$$

donde el factor local $L_\mathfrak{p}$ es el polinomio dado por

$$L_\mathfrak{p}(T) = \begin{cases} 1 - a_\mathfrak{p} T + N(\mathfrak{p})T^2, & \text{si } \mathfrak{p} \text{ es de buena reducción,} \\ 1, 1+T \text{ ó } 1-T, & \text{si no,} \end{cases}$$

donde $a_\mathfrak{p} = N(\mathfrak{p}) + 1 - \#\overline{E}(k_\mathfrak{p})$, y el factor correspondiente a cada uno de los primos de mala reducción queda determinado en términos de cómo sea dicha reducción. Estos números satisfacen que $|a_\mathfrak{p}| \leq 2\sqrt{N(\mathfrak{p})}$ (cota de Hasse), de lo que se deduce que $L(E/K, s)$ converge uniformemente sobre compactos de $\Re(s) > 3/2$.

Notar el parecido entre (3.2) y (2.2). Podemos pensar a ζ_K como una L-serie en la que todos los factores locales son iguales a $1 - T$.

Ahora del lado aritmético,

Proposición 3.1. *$E(\overline{K})$ es un grupo abeliano. Más aún, es un grupo algebraico definido sobre K, del cual $E(K)$ es un subgrupo.*

El objeto que nos interesa es el grupo $E(K)$. El primer resultado importante sobre la estructura de este grupo es el siguiente teorema, que entenderemos como análogo a la descripción de $\mathcal{O}_K^\times$ que nos da el Teorema 1.4.

Teorema 3.2 (Mordell-Weil)**.** *El grupo abeliano $E(K)$ es finitamente generado.*

A diferencia de lo que ocurre con el rango de $\mathcal{O}_K^\times$, el rango de $E(K)$ es difícil de calcular. Lo denotaremos por r_{MW}.

Así como teníamos en $\mathrm{reg}(K)$ una medida del "tamaño" de la parte libre de $\mathcal{O}_K^\times$, en este contexto tenemos el *regulador* de E/K, que se define como

$$\mathrm{reg}(E/K) = \det\left(\langle P_i, P_j\rangle\right)_{i,j}, \tag{3.3}$$

donde $P_1, \ldots, P_{r_{MW}}$ es una base para la parte libre de $E(K)$, y $\langle\ ,\ \rangle$ es una forma bilineal en $E(K)$, que se calcula en términos de la *altura de Néron-Tate* definida en $E(K)$. Notar la similitud entre (3.3) y (1.2).

Sin dudas, de los invariantes involucrados en la conjetura de Birch y Swinnerton-Dyer, el más complicado es el *grupo de Tate-Shafarevich*, que juega un rol análogo al de $Cl(K)$ en la fórmula para el número de clases.

Para definirlo, consideramos la acción de $G_K := \mathrm{Gal}(\overline{K}/K)$ en $E(\overline{K})$. Cada primo $\mathfrak{p}$ de K induce una valuación en $K^\times$, dada por $x \mapsto e_\mathfrak{p}$, si $(x)_{\mathcal{O}_K} = \prod_\mathfrak{q} \mathfrak{q}^{e_\mathfrak{q}}$ con $\mathfrak{q}$ primo y $e_\mathfrak{q} \in \mathbb{Z}$. Denotamos por $K_\mathfrak{p}$ a la completación de K con respecto a esta valuación. Podemos entonces considerar también la acción de $G_{K_\mathfrak{p}} := \mathrm{Gal}(\overline{K_\mathfrak{p}}/K_\mathfrak{p})$ en $E(\overline{K_\mathfrak{p}})$. Como $G_{K_\mathfrak{p}}$ es un subgrupo de G_K, podemos considerar el morfismo

$$H^1(G_K, E(\overline{K})) \longrightarrow \prod_\mathfrak{p} H^1(G_{K_\mathfrak{p}}, E(\overline{K_\mathfrak{p}})),$$

que en cada coordenada viene dado por la restricción de G_K a $G_{K_\mathfrak{p}}$. El grupo de Tate-Shafarevich es el núcleo de este morfismo, y se denota por $Ш(E/K)$. Se conjetura que este grupo abeliano es finito.

¿Qué relación hay entre el grupo de Tate-Shafarevich y el grupo de clases? Ambos miden la obstrucción a que elementos localmente triviales (una clase de cohomología, o un ideal fraccionario) sean globalmente triviales. Más concretamente, si consideramos la acción de G_K en $\mathcal{O}_{\overline{K}}^\times$, se puede probar que $Cl(K)$ es isomorfo al núcleo del morfismo

$$H^1(G_K, \mathcal{O}_{\overline{K}}^\times) \longrightarrow \prod_\mathfrak{p} H^1(G_{K_\mathfrak{p}}, \mathcal{O}_{\overline{K_\mathfrak{p}}}^\times).$$

Sobre el último de los invariantes que hace falta introducir, la *constante de Tamagawa*, no entraremos en detalles. Se define como un producto de factores locales elementales, determinados por la geometría de la curva en los primos de mala reducción y en el infinito. La denotaremos por c.

Ahora sí, podemos enunciar el análogo al Teorema 2.1.

Conjetura 3.3 (Birch y Swinnerton-Dyer)**.**

1. *La función $L(E/K, s)$ se puede continuar de manera holomorfa a todo $\mathbb{C}$.*
2. *Sea r el orden con el que se anula $L(E/K, s)$ en $s = 1$. Entonces, $r = r_{MW}$.*
3.

$$\frac{L^{(r)}(E/K, 1)}{r!} = \frac{c}{\sqrt{|\operatorname{disc}(K)|}} \cdot |Ш(E/K)| \cdot \frac{\operatorname{reg}(E/K)}{|T(E(K))|^2}. \tag{3.4}$$

A diferencia de lo que sucede con la fórmula para el número de clases, esta conjetura dista mucho de ser un teorema[3]. La parte 1 se sabe cierta para cuerpos K totalmente reales (i.e. con $s = 0$), y solo a partir de los trabajos de Wiles, Taylor y otros sobre la demostración de la conjetura de Shimura-Taniyama. El resto de la conjetura, cuando $K = \mathbb{Q}$, se sabe cierta si $r \leq 1$.

Las funciones generatrices $\zeta_K(s)$ y $L(E/K, s)$ se definen como productos cuyos factores contienen información local de K y de E/K. La fórmula para el número de clases y la conjetura de Birch y Swinnerton-Dyer relacionan esta información local con invariantes globales de estos objetos.

Concluimos estas notas comparando las igualdades (2.3) y (3.4) término a término.

- El miembro izquierdo de la igualdad es el "primer" coeficiente de la función generatriz correspondiente.
- El denominador $\sqrt{|\operatorname{disc}(K)|}$ aparece en ambas fórmulas.
- $Cl(K)$ se corresponde con $Ш(E/K)$. Mientras que la finitud del primero es conocida (y fácil de demostrar), la del segundo es conjetural.
- $\operatorname{reg}(K)$ se corresponde con $\operatorname{reg}(E/K)$.
- ω_K se corresponde con $|T(E(K))|$ (aunque uno aparece elevado al cuadrado, y el otro no).
- El factor $2^r(2\pi)^s$ se corresponde con la constante de Tamagawa.

Referencias

[BS66] A. I. Borevich and I. R. Shafarevich. *Number theory*. Pure and Applied Mathematics, Vol. 20. Academic Press, New York-London, 1966.

[Dar09] Henri Darmon. Rational points on curves. In *Arithmetic geometry*, volume 8 of *Clay Math. Proc.*, pages 7–53. Amer. Math. Soc., Providence, RI, 2009.

[Mar77] Daniel A. Marcus. *Number fields*. Springer-Verlag, New York-Heidelberg, 1977. Universitext.

[Neu99] Jürgen Neukirch. *Algebraic number theory*, volume 322 of *Fundamental Principles of Mathematical Sciences*. Springer-Verlag, Berlin, 1999.

[Sch88] Peter Schneider. Introduction to the Beĭlinson conjectures. In *Beĭlinson's conjectures on special values of L-functions*, volume 4 of *Perspect. Math.*, pages 1–35. Academic Press, Boston, MA, 1988.

[Sil09] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, Dordrecht, second edition, 2009.

[Wil06] Andrew Wiles. The Birch and Swinnerton-Dyer conjecture. In *The millennium prize problems*, pages 31–41. Clay Math. Inst., Cambridge, MA, 2006.

[3] De hecho, "esta fórmula relaciona el valor de una función en un punto en el que no está definida con el orden de un grupo cuya finitud no ha sido demostrada" (John Tate).

Instituto de Matemática y Estadística - Facultad de Ingeniería, Universidad de la República - Uruguay
E-mail address: nsirolli@fing.edu.uy