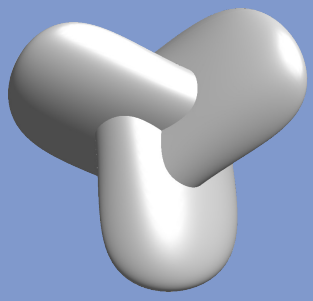


# Publicaciones Matemáticas del Uruguay

Proceedings de la

**III Escuela AGRA**

*editado por*

Harald Helfgott
Emilio Lauret
Ariel Pacetti
Gonzalo Tornaría

Volumen 18, Diciembre 2023

Publicaciones Matemáticas del Uruguay



**Publicada por:**

CMAT — Facultad de Ciencias
IMERL — Facultad de Ingeniería

Universidad de la República

http://pmu.uy/

**ISSN: 0797-1443**

**Créditos:**

Diseño de tapa: J. Rodriguez Hertz
Editor LaTeX: G. Tornaría

Publicaciones Matemáticas del Uruguay

Volumen 18 Diciembre 2023

# Prefacio

El presente volumen de las Publicaciones Matemáticas del Uruguay contiene las notas de cursos de la escuela AGRA III (Aritmética, Grupos y Análisis), realizada en la ciudad de Córdoba, Argentina, del 9 al 20 de julio de 2018.

La meta principal de la serie de escuelas AGRA es la formación de estudiantes graduados y jóvenes investigadores especializándose en teoría de números, teoría de grupos, geometría diofántica y áreas vecinas. Naturalmente, se ha tornado también en una oportunidad para que muchos investigadores más establecidos tengan un proyecto en común, y así ayuda a cohesionar una comunidad informal de personas activas en dichas áreas en Latinoamérica.

El programa del AGRA III consistió de los cursos siguientes.

1. Curvas sobre cuerpos finitos
   Miriam Abdón (Universidade Federal Fluminense), Cícero Carvalho (Universidade Federal de Uberlândia), Daniel Panario (Carleton University)
2. Equidistribución, teoría del potencial y aplicaciones aritméticas
   José Ignacio Burgos Gil (ICMAT, Madrid), Ricardo Menares (Pontificia Universidad Católica de Chile)
3. Representaciones de Galois
   Luis Dieulefait (Universitat de Barcelona), Ariel Pacetti (Universidad Nacional de Córdoba, Argentina), Fernando Rodriguez Villegas (ICTP Trieste)
4. Introducción a grupos aritméticos
   Emilio Lauret (Universidad Nacional del Sur, Argentina), Roberto Miatello (Universidad Nacional de Córdoba), Benjamin Linowitz (Oberlin College)
5. Primos, paridad y análisis
   Harald Helfgott (Universität Göttingen/CNRS), Adrián Ubis (Universidad Autónoma de Madrid)
6. Variedades abelianas, una introducción
   Marc Hindry (Université Paris 7), Marusia Rebolledo (Université Clermont Auvergne), David Roberts (University of Minnesota)

Cada unidad fue acompañada por sesiones guiadas de problemas, a menudo incluyendo mini-clases sobre técnicas útiles. También hubo reuniones nocturnas donde se siguió trabajando sobre las hojas de problemas con empanadas a disposición libre.

Se agradece el patrocinio de la fundación Humboldt (a través de la cátedra Humboldt de H. Helfgott), ICTP (The Abdus Salam International Center for Theoretical Physics, Trieste), CIMPA (Centre International de Mathématiques Pures et Appliquées, Nice), la Academia Nacional de Ciencias, la Fundación Compositio Mathematica y el Ministerio de Ciencia y Tecnología (Gobierno de Córdoba), así como el rol logístico de la FAMAF (Facultad de Matemática, Astronomía, Física y Computación, Universidad Nacional de Córdoba), y el CIEM-CONICET (Centro de Investigación y Estudios de Matemática, Consejo Nacional de Investigaciones Científicas y Técnicas).

El comité científico consistió de Michael Harris (Université Paris 7//Columbia University), Harald Helfgott, Roberto Miatello, Fernando Rodriguez Villegas y Nuria Vila Oliva (Universitat de Barcelona). El comité nacional estuvo compuesto de María Chara (Universidad Nacional del Litoral), Emilio Lauret (Universidad Nacional del Sur), Ariel Pacetti (Universidade de Aveiro), Ricardo Podestá (Universidad Nacional de Córdoba), Diego Sulca (Universidad Nacional de Córdoba) y Ángel Villanueva (Universidad Nacional de Cuyo).

La escuela se realizó en el local histórico de la Academia Nacional de Ciencias, en Córdoba. Si vimos o no cóndores en la Quebrada del Condorito queda como pregunta indecisa y probablemente indecidible.

*Harald Helfgott, Emilio Lauret, Ariel Pacetti y Gonzalo Tornaría*

Córdoba, Julio 2018 – Montevideo, Diciembre 2023.

Curso

# CURVAS SOBRE CUERPOS FINITOS

MIRIAM ABDÓN, CÍCERO CARVALHO, Y DANIEL PANARIO

Publicaciones Matemáticas del Uruguay
Volumen 18, Diciembre 2023, Páginas 1–57
S 0797-1443

# CURVAS SOBRE CUERPOS FINITOS

MIRIAM ABDÓN, CÍCERO CARVALHO, Y DANIEL PANARIO

Resumen. El objetivo principal de estas notas es presentar resultados de la teoría de curvas algebraicas definidas sobre cuerpos finitos, con énfasis en el estudio de curvas maximales. Iniciamos con una exposición de resultados de la teoría de cuerpos finitos que serán necesarios en el estudio de curvas. En seguida pasamos al estudio de cuerpos de funciones algebraicas en una variable, que corresponde al estudio de la geometría intrínseca de las curvas algebraicas, y presentamos también una aplicación de resultados de cuerpos de funciones a la teoría de códigos. Finalmente pasamos a algunos de los principales resultados de la teoría de curvas algebraicas, especialmente a los que se refieren al número de puntos racionales de la curva.

## Índice



Versión final: 18 de mayo de 2019.
Cicero Carvalho: trabajo parcialmente financiado por Fapemig (proc. CEX-APQ-01645-16) y CNPq.
Daniel Panario: trabajo parcialmente financiado por NSERC de Canada.
Estas notas corresponden al curso dictado por los autores en la escuela AGRA III, Aritmética, Grupos y Análisis, del 9 al 20 de Julio de 2018 en Córdoba, Argentina.

## 1. Introducción

El origen de los cuerpos finitos se remonta a los siglos XVII y XVIII. Los primeros en estudiarlos fueron: Fermat (1601-1665), Euler(1707-1783), Lagrange (1736-1813) y Legendre (1752-1833). Todos ellos trabajaron sobre determinados cuerpos finitos: $\mathbb{F}_p$, donde $p$ es un número primo. Más adelante se verá que existe otro tipo de cuerpos finitos.

La teoría de cuerpos finitos tal y como se conoce hoy en día fue construida a finales del siglo XVIII y principios del XIX. Los principales investigadores en el área fueron: Carl Friedrich Gauss (1777-1855) y Évariste Galois (1811-1832). El artículo de Galois *Sur la théorie des nombres* marcó el inicio de los *cuerpos finitos*.

El siguiente gran paso en la construcción de cuerpos finitos fue dado por Richard Dedekind en 1857. Él caracterizó a los cuerpos finitos de orden $p^n$ como anillos de clases residuales

$$\mathbb{F}_p[x]/(f)$$

donde $f$ es un polinomio irreducible de grado $n$ sobre $\mathbb{F}_p$. También introdujo la fórmula de inversión de Möbius en cuerpos finitos para estudiar el número de polinomios irreducibles de cierto grado.

Finalmente, Eliakim H. Moore en 1893, demostró que los cuerpos finitos deben tener $p^n$ elementos si $p$ es un número primo.

A finales del siglo XIX, toda la estructura de los cuerpos finitos era conocida. El libro de Dickson (1901) ya tenía todos los elementos importantes de tal estructura.

### 1.1. Resultados fundamentales.

1. En cualquier cuerpo finito, el número de elementos es potencia de un número primo, este último es la característica del cuerpo.
2. Si $p$ es un primo y $m$ un número positivo, entonces existe un cuerpo finito de orden $p^m$, el cual es único salvo isomorfismos.
3. El grupo multiplicativo de elementos no nulos de $\mathbb{F}_q$, $\mathbb{F}_q^\times$, es cíclico. Cualquier elemento generador es un elemento primitivo de $\mathbb{F}_q$.
4. Si $q = p^m$ entonces cada subcuerpo de $\mathbb{F}_q$ tiene orden $p^d$, donde $d$ es un divisor positivo de $m$. Recíprocamente, si $d|m$ entonces existe exactamente un subcuerpo de $\mathbb{F}_q$ de orden $p^d$.
5. Cada elemento $a \in \mathbb{F}_q$ cumple $a^q = a$.
6. Un cuerpo finito $\mathbb{F}_q$ es isomorfo al cuerpo de descomposición de $x^q - x$ sobre $\mathbb{F}_p$, donde $p$ es la característica de $\mathbb{F}_q$.

El siglo XX fue la época en la que se desarrollaron aplicaciones de cuerpos finitos, debido mayormente a la aparición de las computadoras.

Las áreas de aplicación más importantes son: criptografía y teoría de códigos. Sin embargo, hoy en día el uso de los cuerpos finitos se ha expandido.

El principal libro para ahondarse en la teoría de cuerpos finitos es de Lidl y Niederreiter [40]; para una colección actualizada de temas de investigación en cuerpos finitos ver el manual de cuerpos finitos de Mullen y Panario [41].

### 1.2. Anillos y cuerpos.

**Definición 1.1.** Un *anillo* $(R, +, \cdot)$ es un conjunto $R$ junto con dos operaciones "$+$" y "$\cdot$", tal que:

1. $(R, +)$ es un grupo abeliano;
2. $\cdot$ es asociativo, es decir, para todo $a, b, c \in R, a \cdot (b \cdot c) = (a \cdot b) \cdot c$;

3. las leyes distributivas se cumplen: para todo $a, b, c \in R$, $a \cdot (b+c) = a \cdot b + a \cdot c$ y $(b + c) \cdot a = b \cdot a + c \cdot a$.

**Definición 1.2.** Sea $R$ un anillo.

1. Un anillo se denomina **anillo con identidad**, si el anillo tiene identidad multiplicativa.
2. Un anillo es **conmutativo** si bajo "$\cdot$" es conmutativo.
3. Un anillo se denomina **anillo de división** si los elementos distintos de cero forman un grupo bajo "$\cdot$".
4. Un anillo se denomina **cuerpo** si es un anillo de división conmutativo con identidad.

Dicho de otra manera, un cuerpo $(F, +, \cdot)$ es un conjunto $F$ junto con operaciones $+$ y $\cdot$ tal que:

1. $(F, +)$ es un grupo abeliano con identidad 0;
2. $(F \setminus \{0\}, \cdot)$ es un grupo abeliano con identidad 1;
3. las leyes distributivas se cumplen, i.e., para todo $a, b, c \in F$ se cumple

$$a \cdot (b + c) = a \cdot b + a \cdot c,$$
$$(b + c) \cdot a = b \cdot a + c \cdot a.$$

Si $|F|$ es finito, entonces se dice que $F$ es un *cuerpo finito*. El número de elementos en $F$ es el *orden* del cuerpo finito.

La definición anterior implica que, excepto el 0, todos los elementos de $F$ tienen inverso.

Es bien conocido que $\mathbb{Z}/(p)$ es un cuerpo si y sólo si $p$ es un número primo. Por ejemplo, $\mathbb{Z}/(6)$ no es un cuerpo finito puesto que $2 \cdot 3 \equiv 0 \pmod 6$. Dicho de otra forma, 2 no tiene inverso multiplicativo en $\mathbb{Z}/(6)$.

**Definición 1.3.** Sea $p$ un número primo, $\mathbb{F}_p$ el conjunto $\{0, 1, \ldots, p-1\}$ de enteros y $\phi : \mathbb{Z}/(p) \to \mathbb{F}_p$, la aplicación: $\phi([a]) = a$ para $a = 0, 1, \ldots, p-1$. Entonces, $\mathbb{F}_p$ posee la estructura de cuerpo inducida por $\mathbb{Z}/(p)$, por lo tanto es un cuerpo finito de orden $p$.

Se denotará a un cuerpo finito con $q$ elementos por $\mathbb{F}_q$. Más adelante se verá que $q$ debe ser una potencia de un primo y que salvo isomorfismos hay solamente un cuerpo finito con $q$ elementos.

En $\mathbb{Z}$, para un entero $a \neq 0$, $an = 0$ (donde $n \in \mathbb{N}$) implica que $n = 0$.

Ahora consideremos $\mathbb{Z}/(p)$ y nuevamente tomemos $a \neq 0$. Entonces se puede demostrar que $ap = 0$ y $p$ es el entero positivo más pequeño con esta propiedad.

**Definición 1.4.** Si $R$ es un anillo arbitrario y existe un entero positivo $n$ tal que $nr = 0$ para todo $r \in R$, entonces el entero positivo más pequeño $n$ es la *característica* del anillo y se dice que $R$ tiene *característica positiva*. De no ser así, se dice que $R$ es de *característica cero*.

**Teorema 1.5.** *Un anillo $R \neq \{0\}$ con característica positiva que tiene una identidad y ningún divisor de cero trivial, debe tener característica prima.*

**Corolario 1.6.** *Un cuerpo finito tiene característica prima.*

### 1.3. Propiedades básicas.

**Teorema 1.7.** *Si $\mathbb{F}_q$ es un cuerpo de característica prima $p$ y $n \geq 1$, entonces*

$$(a+b)^{p^n} = a^{p^n} + b^{p^n} \text{ y } (a-b)^{p^n} = a^{p^n} - b^{p^n}.$$

*Demostración.* (Esbozo.) Usamos inducción sobre $n$:

1. Base (cuando $n = 1$): $(a+b)^p = a^p + b^p$. Se desarrolla el binomio y se verifica que cada coeficiente $0 < i < p$ es cero, puesto que
$$\left( \begin{array}{c} p \\ i \end{array} \right) = \frac{p(p-1)\dots(p-i+1)}{i \dots 2\,1} \equiv 0 \mod p.$$
2. De $(a+b)^{p^n} = a^{p^n} + b^{p^n}$ se deduce que $(a-b)^{p^n} = a^{p^n} - b^{p^n}$ puesto que
$$a^{p^n} = ((a-b)+b)^{p^n} = (a-b)^{p^n} + b^{p^n}.$$

□

**Teorema 1.8.** *El grupo multiplicativo de elementos distintos de cero en $\mathbb{F}_q$, denotado por $\mathbb{F}_q^\times$, es cíclico.*

*Demostración.* (Esbozo.) Si $q = 2$ es fácil ver que el resultado es cierto. Suponga que $q \geq 3$. El orden de $\mathbb{F}_q^\times$ es $q-1$. Ahora considere la factorización en primos de $h = q-1$

$$h = p_1^{r_1} p_2^{r_2} \cdots p_m^{r_m},$$

donde $p_1, p_2, \dots, p_m$ son números primos distintos y $r_1, r_2, \dots, r_m$ son enteros positivos.

Para cada $i, 1 \leq i \leq m$, considere el polinomio $x^{h/p_i} - 1$. Puesto que el grado de este es $h/p_i$, tiene como máximo $h/p_i$ raíces. Ahora bien, $h/p_i < h = q-1$, por lo tanto hay elementos en $\mathbb{F}_q^\times$ que no son raíces de $x^{h/p_i} - 1$.

Sea $a_i$ un elemento de $\mathbb{F}_q^\times$ que no es una raíz de $x^{h/p_i} - 1$. Definamos el elemento

$$b_i = a_i^{h/(p_i)^{r_i}}.$$

Dejamos como ejercicio probar que el orden de $b_i$ es $p_i^{r_i}$ y que si $b = b_1, b_2, \dots b_m$, entonces el orden de $b$ es $q-1$ y por lo tanto es un generador de $\mathbb{F}_q^\times$. □

**Ejemplo 1.9.** Consideremos $\mathbb{F}_7$ y su grupo multiplicativo $\mathbb{F}_7^\times = \{1,2,3,4,5,6\}$. Es fácil verificar que 2 no es un elemento primitivo, pero en el caso de 3 se tiene que:

$$3^1 = 3,\ 3^2 = 2,\ 3^3 = 6,\ 3^4 = 4,\ 3^5 = 5,\ 3^6 = 1,$$

de donde se concluye que 3 es un elemento primitivo.

Existen algoritmos para encontrar elementos primitivos, pero ninguno de ellos se ejecuta en un tiempo polinomial en el tamaño de la entrada. Encontrar un tal algoritmo es un problema abierto difícil.

El siguiente teorema caracteriza los elementos que pertenecen a un cuerpo finito.

**Teorema 1.10.** *Si $\mathbb{F}_q$ es un cuerpo finito de $q$ elementos, entonces cada $a \in \mathbb{F}_q$ satisface $a^q = a$.*

*Demostración.* Si $a = 0$, entonces es claro que $a^q = a$. Si $a \neq 0$, entonces $a^{q-1} = 1$, puesto que $q-1$ es el orden del grupo multiplicativo in $\mathbb{F}_q$. □

**1.4. Polinomios sobre cuerpos finitos.** Un polinomio sobre un anillo $R$ es una expresión de la forma

$$p(x) = \sum_{i=0}^{n} a_i x^i,$$

donde $n$ es un entero no negativo y $a_i \in R$, para todo $i = 0, 1, \ldots n$. Un polinomio es *mónico* si el término líder tiene coeficiente 1.

**Definición 1.11.** El anillo formado por los polinomios sobre $R$ con operaciones suma y producto de polinomios, se denomina **anillo de polinomios sobre** $R$ y se denota como $R[x]$.

Un polinomio $f(x) = a_n x^n + \cdots + a_0$ tiene grado $n$ si $a_n \neq 0$. Por convención, el polinomio $f(x) = 0$ tiene grado $-\infty$.

**Teorema 1.12.** *Si $f, g \in R[x]$, entonces*

$$\begin{aligned} grado(f+g) &\leq max(grado(f), grado(g)) \\ grado(fg) &\leq grado(f) + grado(g). \end{aligned}$$

Sea $F$ un cuerpo. Un polinomio $g \in F[x]$ *divide a* un polinomio $f \in F[x]$, si existe un polinomio $h \in F[x]$ tal que $f = gh$. Se dice entonces que $g$ es un *divisor* de $f$.

**Teorema 1.13** (Algoritmo de la división)**.** *Si $g \in F[x], g \neq 0$ y $F$ es un cuerpo, entonces para cualquier $f \in F[x]$ existen polinomios únicos $q, r \in F[x]$ tales que $f = qg + r$ y $grado(r) < grado(g)$.*

Algunas clases importantes de polinomios sobre cuerpos finitos incluyen:

- Un polinomio $f \in \mathbb{F}_q[x]$ es *irreducible* sobre $\mathbb{F}_q$ si $f$ tiene grado positivo y $f = gh$ con $g, h \in \mathbb{F}_q[x]$ implica que $g$ o $h$ es una constante. De otra forma $f$ es reducible.
- Sea $f \in \mathbb{F}_q[x]$ un polinomio distinto del polinomio idénticamente nulo. Si $f(0) \neq 0$, entonces al entero positivo más pequeño $e$ para el cual $f(x)$ divide a $x^e - 1$, se le denomina el *orden* de $f$ y se denota como $\text{ord}(f)$. Si $f(x) = x^h g(x)$ con $g(0) \neq 0$, entonces $\text{ord}(f) = \text{ord}(g)$. Un polinomio mónico $f \in \mathbb{F}_q[x]$ de grado $m$ es *primitivo* sobre $\mathbb{F}_q$ si $f(0) \neq 0$ y $\text{ord}(f) = q^m - 1$.
- Un polinomio $f \in \mathbb{F}_q[x]$ es una *permutación polinomial* sobre $\mathbb{F}_q$ si la función polinomial asociada $f : c \mapsto f(c)$ de $\mathbb{F}_q$ en $\mathbb{F}_q$ es una permutación de $\mathbb{F}_q$.

En las presentes notas no nos adentraremos en aplicaciones de estos polinomios, pero existen innumerables aplicaciones de ellos en criptografía, sucesiones sobre cuerpos finitos, combinatoria y geometría finita, entre otras áreas de investigación.

Cada aplicación de $\mathbb{F}_q$ en sí mismo puede expresarse como un polinomio. Es claro que si $\phi : \mathbb{F}_q \to \mathbb{F}_q$ es una función arbitraria de $\mathbb{F}_q$ en $\mathbb{F}_q$, entonces existe un polinomio único $g \in \mathbb{F}_q$ con $\text{grado}(g) < q$ representando a $\phi$, es decir, $g(c) = \phi(c)$ para toda $c \in \mathbb{F}_q$. Es posible hallar al polinomio $g$ usando algún método de interpolación (como el de Lagrange) para la función $\phi$.

Los polinomios irreducibles son los elementos "primos" de los polinomios. Al igual que los números primos para los números enteros, los polinomios irreducibles tienen un papel preponderante para los cuerpos finitos. En el caso de los enteros, se tiene que

$$\mathbb{Z}/(p) \text{ es un cuerpo si y sólo si } p \text{ es un número primo.}$$

Los siguientes teoremas garantizan que lo mismo se cumple para los polinomios sobre cuerpos finitos.

**Teorema 1.14.** *Para $f \in \mathbb{F}_q[x]$, el anillo de clases residuales $\mathbb{F}_q[x]/(f)$ es un cuerpo si y sólo si $f$ es irreducible.*

Por lo tanto, un problema de suma importancia es hallar polinomios irreducibles en cuerpos finitos. En aplicaciones, como criptografía por ejemplo, la elección del polinomio irreducible para la construcción de una extensión de un cuerpo finito juega un papel muy importante en la eficiencia de los métodos considerados.

El teorema anterior garantiza que se obtendrá un cuerpo si y solamente si, el polinomio que define la estructura es irreducible. Observamos que si $f$ es un polinomio mónico irreducible sobre $\mathbb{F}_p$ y grado $(f) = n$, entonces el número de elementos de $\mathbb{F}_p/(f)$ es $p^n$. Así que $\mathbb{F}_p/(f)$ es un cuerpo finito con $p^n$ elementos.

**Ejemplos:** Consideremos primero $x^2 + 1 \in \mathbb{F}_2[x]$ como el polinomio que define al "cuerpo" $\mathbb{F}_4$. Usándolo se genera la tabla para el producto que se muestra a continuación:

| $\cdot$ | $0$ | $1$ | $x$ | $x+1$ |
|---|---|---|---|---|
| $0$ | $0$ | $0$ | $0$ | $0$ |
| $1$ | $0$ | $1$ | $x$ | $x+1$ |
| $x$ | $0$ | $x$ | $1$ | $x+1$ |
| $x+1$ | $0$ | $x+1$ | $x+1$ | $0$ |

.

Como se podrá observar no se obtuvo un cuerpo finito, esto se debe a que $(x + 1)(x + 1) = 0$ en $\mathbb{F}_2[x]/(x^2 + 1)$ y $x + 1 \neq 0$, es decir, se tienen divisores de cero diferentes de cero y por tanto no puede ser un cuerpo (finito). Adicionalmente, $x + 1$ no tiene inverso ya que no existe un elemento que multiplicado por $x + 1$ dé 1 como resultado y los elementos en un cuerpo que son distintos de cero, deben tener un inverso.

Ahora dado $x^2 + x + 1 \in \mathbb{F}_2[x]$, se tiene que

| $\cdot$ | $0$ | $1$ | $x$ | $x+1$ |
|---|---|---|---|---|
| $0$ | $0$ | $0$ | $0$ | $0$ |
| $1$ | $0$ | $1$ | $x$ | $x+1$ |
| $x$ | $0$ | $x$ | $x+1$ | $1$ |
| $x+1$ | $0$ | $x+1$ | $1$ | $x$ |

.

En este caso sí se obtiene un cuerpo. La diferencia está en el polinomio que se usó. En el primer ejemplo, el polinomio usado $x^2 + 1$ era *reducible* sobre $\mathbb{F}_2$, mientras que en el segundo ejemplo, $x^2 + x + 1$ es *irreducible* sobre $\mathbb{F}_2$. (Ejercicio: verificar que $x^2 + x + 1$ es irreducible sobre $\mathbb{F}_2$.)

Terminamos esta sección con un teorema fundamental para polinomios sobre cuerpos finitos.

**Teorema 1.15** (Factorización única en $\mathbb{F}_q[x]$.)**.** *Cualquier polinomio $f \in \mathbb{F}_q[x]$ de grado positivo se puede expresar como*

$$f = a f_1^{e_1} f_2^{e_2} \dots f_k^{e_k}$$

*donde $a \in \mathbb{F}_q$, $f_1, f_2, \dots, f_k$ son polinomios mónicos irreducibles distintos y $e_1, e_2, \dots, e_k$ son enteros positivos. Además esta factorización es única independientemente del orden en el que aparezcan los factores.*

La demostración de este teorema no es constructiva. En otras palabras, no ofrece ningún algoritmo para factorizar polinomios. Sin embargo, hoy en día existen métodos muy eficientes para factorizar polinomios sobre cuerpos finitos; ver, por ejemplo [23].

**1.5. Estructura de los cuerpos finitos.** Sea $p$ un número primo. Es sabido que:

1. $\mathbb{Z}/(p)$ es un cuerpo finito;
2. $\mathbb{F}_p[x]/(f)$ es un cuerpo finito si $f$ es irreducible sobre $\mathbb{F}_p$;
3. si grado $(f) = n$ y $f$ es irreducible entonces $\mathbb{F}_p[x]/(f)$ tiene $p^n$ elementos.

¿Son las anteriores las únicas posibles opciones para construir cuerpos finitos? Lo que se desea, es justamente caracterizar a todos los cuerpos finitos posibles.

**Definición 1.16.** Sea $F$ un cuerpo y $K \subseteq F$. Si $K$ es en sí mismo un cuerpo bajo las operaciones de $F$, a $K$ se le denomina *subcuerpo* de $F$ y a $F$ se le llama una *extensión* de $K$. Si $K \neq F$, se dice que $K$ es un *subcuerpo propio* de $F$.

Observe que $\mathbb{F}_p$ no tiene subcuerpos propios. Es claro que si $K \subset \mathbb{F}_p$ y $K$ es un cuerpo, entonces 0 y 1 son elementos de $K$. Puesto que $K$ es un cuerpo, debe ser cerrado bajo la suma, así que cada elemento en $\mathbb{F}_p$ está en $K$. En consecuencia $K = \mathbb{F}_p$.

**Definición 1.17.** Un cuerpo que no tiene subcuerpos propios se llama un **cuerpo primo**.

Los cuerpos primos se obtienen considerando la intersección de todas las colecciones distintas de cero, de subcuerpos de un cuerpo dado. El siguiente teorema caracteriza a los cuerpos primos.

**Teorema 1.18.** *El subcuerpo primo de un cuerpo $F$ es isomorfo a $\mathbb{F}_p$ o bien a $\mathbb{Q}$, dependiendo de si la característica de $F$ es prima o cero.*

**Definición 1.19.** Si $K$ es un subcuerpo de $F$ y $M$ cualquier subconjunto de $F$. Entonces el cuerpo $K(M)$ está definido como la intersección de todos los subcuerpos de $F$ que contienen tanto a $K$ como a $M$ y se le denomina extensión del cuerpo $K$, obtenida adjuntando los elementos de $M$.

Para un conjunto finito $M = \{\theta_1, \theta_2, \ldots, \theta_n\}$, se escribe $K(M) = K(\theta_1, \theta_2, \ldots, \theta_n)$. Si $M$ tiene sólo un elemento $\theta \in F$, entonces $L = K(\theta)$ se llama una *extensión simple de $K$* y $\theta$ será el *elemento de definición* de $L$ sobre $K$.

Si $L$ es una extensión del cuerpo $K$, entonces es posible ver a $L$ como un espacio vectorial sobre $K$, puesto que los elementos de $L$ forman un grupo abeliano bajo la suma ($L$ es un cuerpo) y la multiplicación escalar de un elemento $\alpha \in L$ por un elemento $r \in K$ da como resultado $r\alpha \in L$, el cual cumple que:

$$\begin{aligned} r(\alpha + \beta) &= r\alpha + r\beta \\ (r + s)\alpha &= r\alpha + s\alpha \\ (rs)\alpha &= r(s\alpha) \\ 1 \cdot \alpha &= \alpha \end{aligned}$$

para todas $r, s \in K$ y $\alpha, \beta \in L$. La dimensión de este espacio vectorial es el grado de la extensión, si se tiene un espacio de dimensión finita.

**Definición 1.20.** Sea $L$ una extensión de un cuerpo $K$. Si $L$, considerado como un espacio vectorial sobre $K$, tiene dimensión finita, entonces a $L$ se le denomina una *extensión finita de $K$*. A la dimensión del espacio vectorial se le llama el *grado* de $L$ sobre $K$ y se denota por $[L : K]$

**Teorema 1.21.** *Si $L$ es una extensión finita de $K$ y $M$ es una extensión finita de $L$, entonces $M$ es una extensión finita de $K$ y $[M : K] = [M : L][L : K]$.*

**Definición 1.22.** Si $f \in K[x]$ de grado positivo y $F$ es una extensión de $K$, entonces se dice que $f$ descompone en $F$, si $f$ puede escribirse como el producto de factores lineales en $F[x]$. Es decir, $f$ descompone en $F$ si existen $\alpha_1, \alpha_2, \ldots, \alpha_n \in F$ tales que

$$f(x) = a(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n),$$

donde $a \in K$ es el coeficiente líder de $f$. Al cuerpo $F$ se le llama *cuerpo de descomposición* de $f$ sobre $K$ si $f$ se descompone en $F$ y si $F = K(\alpha_1, \alpha_2, \ldots, \alpha_n)$.

El siguiente teorema caracteriza a los cuerpos de descomposición.

**Teorema 1.23** (Existencia y unicidad de los cuerpos de descomposición.)**.** *Si $K$ es un cuerpo y $f$ es cualquier polinomio de grado positivo en $K[x]$, entonces existe un cuerpo de descomposición de $f$ sobre $K$. Cualesquiera dos cuerpos de descomposición de $f$ sobre $K$ son isomorfos bajo un isomorfismo que mantiene a los elementos de $K$ fijos y lleva las raíces de $f$ entre sí.*

**Teorema 1.24.** *Si $F$ es un cuerpo finito. Entonces $F$ tiene $p^n$ elementos donde $p$ es la característica de $F$ y $n$ es el grado de extensión de $F$ sobre su cuerpo primo.*

*Demostración.* Puesto que $F$ es finito, la característica de $F$ es un número primo, y esto implica que el subcuerpo primo $K$ de $F$ es isomorfo a $\mathbb{F}_p$. Por lo tanto contiene $p$ elementos.

Es posible ver a los elementos en $F$ como elementos de un espacio vectorial de $F$ sobre $K$. Por lo tanto, existe una base para $F$ sobre $K$ formada por $\beta_1, \beta_2, \ldots, \beta_n$. Entonces cualquier elemento en $F$ se puede escribir como

$$a_1\beta_1 + a_2\beta_2 + \cdots + a_n\beta_n$$

con $a_1, a_2, \ldots a_n \in K$. Puesto que $a_i \in K$ para todo $i = 1, \ldots, n$, se tienen $p$ posibles valores, el número total de elementos en $F$ es $p^n$. □

Ahora estamos listos para dar uno de los resultados más importantes de cuerpos finitos.

**Teorema 1.25** (Existencia y unicidad de los cuerpos finitos.)**.** *Para cada primo $p$ y cada entero positivo $n$, existe un cuerpo finito con $p^n$ elementos. Cualquier cuerpo finito con $p^n$ elementos es isomorfo al cuerpo de descomposición de $x^{p^n} - x$ sobre $\mathbb{F}_p$.*

Por ejemplo, este teorema garantiza que existe un cuerpo finito con 8 elementos, puesto que $8 = 2^3$ y 2 es primo. Sin embargo, este cuerpo con 8 elementos no es $\mathbb{Z}/(8)$, puesto que $\mathbb{Z}/(8)$ no es un cuerpo; por ejemplo, 4 no tiene inverso. Para encontrar $\mathbb{F}_{2^3}$ hay que hallar un polinomio irreducible de grado 3 sobre $\mathbb{F}_2$. (Ejercicio: hallar un polinomio irreducible de grado 3 sobre $\mathbb{F}_2$ y construir $\mathbb{F}_{2^3}$.)

**Teorema 1.26.** *Si $\mathbb{F}_q$ es un cuerpo finito con $q = p^n$ elementos. Entonces cada subcuerpo de $\mathbb{F}_q$ tiene orden $p^m$, donde $m$ es un divisor positivo de $n$. Recíprocamente, si $m$ es un divisor positivo de $n$, entonces existe exactamente un subcuerpo de $\mathbb{F}_q$ con $p^m$ elementos.*

**Ejemplos:**

1. $\mathbb{F}_{2^{10}}$ tiene subcuerpos $\mathbb{F}_{2^2}$ y $\mathbb{F}_{2^5}$, cada uno de los cuales tiene a $\mathbb{F}_2$ como subcuerpo.

2. $\mathbb{F}_{3^{18}}$ tiene subcuerpos $\mathbb{F}_{3^6}$ y $\mathbb{F}_{3^9}$; $\mathbb{F}_{3^6}$ tiene subcuerpos $\mathbb{F}_{3^2}$ y $\mathbb{F}_{3^3}$, mientras que $\mathbb{F}_{3^9}$ tiene como subcuerpo a $\mathbb{F}_{3^3}$; finalmente, cada uno de ellos, tiene como subcuerpo a $\mathbb{F}_3$.
3. $\mathbb{F}_8$ no es un subcuerpo de $\mathbb{F}_{16}$, aunque $8|16$. Como 3 no divide a 4, por lo tanto $\mathbb{F}_8$ no es un subcuerpo de $\mathbb{F}_{16}$.

Los ejemplos anteriores están ilustrados en la Figura 1.

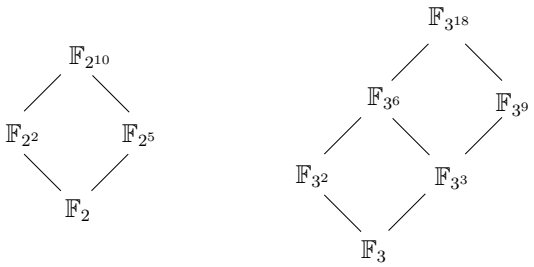



FIGURA 1. Subcuerpos de $\mathbb{F}_{2^{10}}$ e $\mathbb{F}_{3^{18}}$.

Nuestro próximo paso es definir los *conjugados* de un elemento de un cuerpo finito.

**Definición 1.27.** Sean $\mathbb{F}_{q^m}$ una extensión de $\mathbb{F}_q$ y $\alpha$ un elemento en $\mathbb{F}_{q^m}$. Los elementos $\alpha, \alpha^q, \alpha^{q^2}, \ldots, \alpha^{q^{m-1}}$ son los *conjugados* de $\alpha$ con respecto a $\mathbb{F}_q$.

Sea $\alpha \in \mathbb{F}_{q^n}$ con polinomio minimal sobre $\mathbb{F}_q$ de grado $d$. Consideremos el conjunto $\alpha, \alpha^q, \alpha^{q^2}, \ldots, \alpha^{q^{n-1}}$ de conjugados de $\alpha$. Los elementos de este conjunto son distintos si $n = d$; sino, cada conjugado distinto aparece repetido $n/d$ veces.

**Teorema 1.28.** *Los automorfismos distintos de $\mathbb{F}_{q^n}$ sobre $\mathbb{F}_q$ son las funciones $\sigma_0, \sigma_1, \ldots, \sigma_{n-1}$, donde $\sigma_j \colon \mathbb{F}_{q^n} \to \mathbb{F}_{q^n}$ definida como $\sigma_j(\alpha) = \alpha^{q^j}$ para cada $\alpha \in \mathbb{F}_{q^n}$.*

El conjunto de automorfismos de $\mathbb{F}_q$ forma un grupo con la operación de composición funcional llamado *grupo de Galois de $\mathbb{F}_{q^n}$ sobre $\mathbb{F}_q$*. Es un grupo cíclico con generador $\sigma_1 \colon \mathbb{F}_{q^n} \to \mathbb{F}_{q^n}$ que lleva $\alpha \in \mathbb{F}_{q^n}$ a $\alpha^q$ llamado *automorfismo de Frobenius*. Los conjugados de $\alpha$ son, entonces, los elementos a los cuales $\alpha$ es enviado aplicando iterativamente el automorfismo de Frobenius.

La suma y el producto de los conjugados de $\alpha$ producen dos funciones especiales muy usadas en aplicaciones.

**Definición 1.29.** Para cada $\alpha \in \mathbb{F}_{q^m}$, la *traza* de $\alpha$ sobre $\mathbb{F}_q$ es definida por

$$\mathrm{Tr}_{\mathbb{F}_{q^m}/\mathbb{F}_q}(\alpha) = \alpha + \alpha^q + \alpha^{q^2} + \cdots + \alpha^{q^{m-1}}.$$

La próxima proposición, que dejamos como ejercicio, contiene algunas propiedades de la traza.

**Proposición 1.30.** *Sean $\mathbb{F}_{q^m}$ una extensión de $\mathbb{F}_q$, $\alpha, \beta \in \mathbb{F}_{q^m}$ y $a, b \in \mathbb{F}_q$. Entonces,*

1. $\mathrm{Tr}_{\mathbb{F}_{q^m}/\mathbb{F}_q}(\alpha) \in \mathbb{F}_q$;
2. $\mathrm{Tr}_{\mathbb{F}_{q^m}/\mathbb{F}_q}(a\alpha + b\beta) = a\,\mathrm{Tr}_{\mathbb{F}_{q^m}/\mathbb{F}_q}(\alpha) + b\,\mathrm{Tr}_{\mathbb{F}_{q^m}/\mathbb{F}_q}(\beta)$;

3. $\mathrm{Tr}_{\mathbb{F}_{q^m}/\mathbb{F}_q}$ *es una transformación linear de* $\mathbb{F}_{q^m}$ *en* $\mathbb{F}_q$*, donde* $\mathbb{F}_{q^m}$ *y* $\mathbb{F}_q$ *son vistos como espacios vectoriales sobre* $\mathbb{F}_q$*;*
4. $\mathrm{Tr}_{\mathbb{F}_{q^m}/\mathbb{F}_q}(a) = ma$*;*
5. $\mathrm{Tr}_{\mathbb{F}_{q^m}/\mathbb{F}_q}(\alpha^q) = \mathrm{Tr}_{\mathbb{F}_{q^m}/\mathbb{F}_q}(\alpha)$.

Observamos que

$$\begin{aligned} f(x) &= (x-\alpha)(x-\alpha^q)\dots(x-\alpha^{q^{n-1}}) \\ &= x^n - (\alpha+\alpha^q+\dots+\alpha^{q^{n-1}})x^{n-1} + \dots + (-1)^n \alpha\alpha^q\dots\alpha^{q^{n-1}}, \end{aligned}$$

entonces $\mathrm{Tr}_{\mathbb{F}_{q^n}/\mathbb{F}_q}(\alpha) = -a_{n-1}$.

Podemos definir la traza de $\mathbb{F}_{q^n}$ sobre un subcuerpo $\mathbb{F}_{q^m}$:

$$\mathrm{Tr}_{\mathbb{F}_{q^n}/\mathbb{F}_{q^m}}(x) = x + x^{q^m} + \dots + \left(x^{q^m}\right)^{\frac{n-1}{m}}.$$

Cuando tenemos una cadena de extensiones de cuerpos, podemos calcular la composición de trazas.

**Teorema 1.31.** *Sea* $K$ *un cuerpo finito,* $F$ *una extensión de* $K$ *y* $E$ *una extensión de* $F$*. Entonces, para* $\alpha \in E$*,*

$$\mathrm{Tr}_{E/K}(\alpha) = \mathrm{Tr}_{F/K}\left(\mathrm{Tr}_{E/F}(\alpha)\right).$$

*Demostración.* Sea $K = \mathbb{F}_q$, $[F:K] = n$, $[E:F] = m$ y entonces $[E:K] = mn$. Para $\alpha \in E$ tenemos

$$\begin{aligned} \mathrm{Tr}_{F/K}\left(\mathrm{Tr}_{E/F}(\alpha)\right) &= \sum_{i=0}^{n-1} \left(\mathrm{Tr}_{E/K}(\alpha)\right)^{q^i} = \sum_{i=0}^{n-1} \left( \sum_{j=0}^{m-1} \alpha^{(q^n)^j} \right)^{q^i} \\ &= \sum_{i=0}^{n-1}\sum_{j=0}^{m-1} \alpha^{q^{nj+i}} = \sum_{k=0}^{mn-1} \alpha^{q^k} = \mathrm{Tr}_{E/K}(\alpha). \end{aligned}$$

□

La función traza $\mathrm{Tr}_{F/K}$ para una extensión $F$ de $K$ es una transformación linear de $F$ en $K$ que describe todas las posibles transformaciones de $F$ en $K$ (*funcionales lineales* de $F$).

**Teorema 1.32.** *Sea* $F$ *un cuerpo, extensión finita del cuerpo finito* $K$*, donde ambos son considerados como espacios vectoriales sobre* $K$*. Las transformaciones lineales de* $F$ *en* $K$ *son exactamente las funciones* $L_\beta$*,* $\beta \in F$*, donde* $L_\beta(\alpha) = \mathrm{Tr}_{F/K}(\beta\alpha)$ *para todo* $\alpha \in F$*. Además, tenemos* $L_\beta \neq L_\gamma$ *cuando* $\beta$ *y* $\gamma$ *son elementos distintos de* $F$*.*

Otra función interesante de un cuerpo finito a un subcuerpo es la *norma*.

**Definición 1.33.** Para $\alpha \in \mathbb{F}_{q^n}$, la *norma* $\mathrm{N}_{\mathbb{F}_{q^n}/\mathbb{F}_q}(\alpha)$ sobre $\mathbb{F}_q$ es definida como

$$\mathrm{N}_{\mathbb{F}_{q^n}/\mathbb{F}_q}(\alpha) = \alpha\alpha^q\dots\alpha^{q^{n-1}} = \alpha^{\frac{q^n-1}{q-1}}.$$

Si $f(x) = (x-\alpha)\dots(x-\alpha^{q^{n-1}}) = \sum_{i=0}^n \alpha_i x^i$, entonces $\mathrm{N}_{\mathbb{F}_{q^n}/\mathbb{F}_q}(\alpha) = (-1)^n a_0$.

**Teorema 1.34.** *La función norma de* $\mathbb{F}_{q^n}$ *sobre* $\mathbb{F}_q$ *verifica*

(a) $\mathrm{N}_{\mathbb{F}_{q^n}/\mathbb{F}_q}(\alpha\beta) = \mathrm{N}_{\mathbb{F}_{q^n}/\mathbb{F}_q}(\alpha)\mathrm{N}_{\mathbb{F}_{q^n}/\mathbb{F}_q}(\beta)$ *para todo* $\alpha, \beta \in \mathbb{F}_{q^n}$*;*
(b) $\mathrm{N}_{\mathbb{F}_{q^n}/\mathbb{F}_q}$ *lleva* $\mathbb{F}_{q^n}$ *en* $\mathbb{F}_q$ *y* $\mathbb{F}_q^*$ *en* $\mathbb{F}_q^*$*;*

(c) $\mathrm{N}_{\mathbb{F}_{q^n}/\mathbb{F}_q}(\alpha) = \alpha^n$, *para todo* $\alpha \in \mathbb{F}_q$;
(d) $\mathrm{N}_{\mathbb{F}_{q^n}/\mathbb{F}_q}(\alpha^q) = \mathrm{N}_{\mathbb{F}_{q^n}/\mathbb{F}_q}(\alpha)$, *para todo* $\alpha \in \mathbb{F}_{q^n}$.
(e) Transitividad de la norma: *Si* $K$ *es una extensión de un cuerpo finito* $K$ *y* $E$ *es una extensión de un cuerpo finito* $F$, *entonces* $\mathrm{N}_{E/K}(\alpha) = \mathrm{N}_{F/K}\left(\mathrm{N}_{E/F}(\alpha)\right)$, *para todo* $\alpha \in E$.

Ejercicio: probar el Teorema 1.34.

## 2. Cuerpos de funciones, semigrupos de Weierstrass y códigos de Goppa

**2.1. Cuerpos de funciones de una variable.** Sean $K$ y $F$ cuerpos tales que $K \subset F$. Como se ha visto anteriormente, se dice que $F$ es una *extensión* de $K$ y escribimos $F \,|\, K$. También se ha visto que podemos pensar a $F$ como un $K$-espacio vectorial. Si $F$ es de dimensión finita, con $\dim_K F = d$, entonces decimos que la extensión $F \,|\, K$ es *finita, de grado* $d$. Si $y \in F$ es tal que para un polinomio distinto de cero $p(X) \in K[X]$ tenemos $p(y) = 0$, entonces decimos que $y$ es un elemento *algebraico* sobre $K$, de lo contrario, decimos que $y$ es *trascendente* sobre $K$. Observamos que si $F \,|\, K$ es una extensión finita, entonces cada elemento $y \in F$ es algebraico sobre $K$: de hecho, digamos que $\dim_K F = n$, tenemos que $\{1, y, \ldots, y^n\}$ es un conjunto linealmente dependiente, por lo tanto, existen $a_0, \ldots, a_n \in K$, no todos iguales a cero, tal que $a_0 + a_1 y + \cdots + a_n y^n = 0$, es decir, $y$ es una raíz de $p(X) = \sum_{i=0}^{n} a_i X^i \in K[X]$. Si $y \in F$ es trascendente sobre $K$, entonces es fácil comprobar que la intersección de todos los subcuerpos de $F$ que contienen $K$ e $y$ es el subcuerpo $K(y) := \{p(y)/q(y) \in F \,|\, p(X), q(X) \in K[X], q(X) \neq 0\}$, que es isomorfo al cuerpo de fracciones de polinomios $K(X) = \{p(X)/q(X) \,|\, p(X), q(X) \in K[X], q(X) \neq 0\}$. Esto refleja el hecho de que si $y$ es trascendente sobre $K$, entonces se comporta como "una variable" sobre $K$, ya que para cualquier $a_0, \ldots, a_n \in K$ tenemos $\sum_{i=0}^{n} a_i y^i = 0$ si y sólo si $a_i = 0$ para todo $i = 0, \ldots, n$. Decimos que $F \,|K$ es una *extensión algebraica* si cada elemento de $F$ es una raíz de algún polinomio distinto de cero en $K[X]$, de lo contrario, decimos que $F \,|\, K$ es una extensión *trascendente*.

En lo que sigue vamos a trabajar con un objeto básico: *un cuerpo de funciones algebraicas* $F \,|\, K$ *de una variable*. Esta es una extensión $F \,|\, K$ con la propiedad de que existe un elemento $x \in F$, trascendente sobre $K$ y tal que la extensión $F \,|\, K(x)$ es finita.

$$\begin{array}{cl} F & \\ \big| \big) & \text{extensión finita} \\ K(x) & \\ \big| \big) & \text{extensión trascendente} \\ K & \end{array}$$

Siempre asumiremos que $K$ *es algebraicamente cerrado en* $F$, lo que significa que si $f \in F$ es un elemento algebraico sobre $K$, entonces $f \in K$ (en otras palabras, excepto los elementos de $K \subset F$, que obviamente son algebraicos sobre $K$, no hay otros elementos de $F$ que sean algebraicos sobre $K$). Nuestra referencia, en esta sección, es el primer capítulo del libro de H. Stichtenoth ([47]). No tenemos tiempo para demostrar la mayor parte de lo que necesitaremos, por lo que solo indicaremos los resultados y el lector podrá ver las pruebas en ese libro.

El ejemplo más básico de un cuerpo de funciones se obtiene al tomar $F = K(X)$: claramente $X$ es trascendente sobre $K$ y la extensión $F \mid K(X)$ es finita de grado uno (que es solo una forma elaborada de decir que $F = K(X)$).

**Definición 2.1.** Un *anillo de valoración* del cuerpo de funciones $F \mid K$ es un anillo $\mathcal{O}$ tal que:

(1) $K \subsetneq \mathcal{O} \subsetneq F$;
(2) para cualquier $f \in F$ tenemos $f \in \mathcal{O}$ o $f^{-1} \in \mathcal{O}$.

**Lema 2.2.** *Sea $\mathcal{O}$ un anillo de valoración de $F \mid K$. Entonces $\mathcal{O}$ es un anillo local, es decir, $\mathcal{O}$ tiene un único ideal maximal, que es el conjunto $P := \mathcal{O} \setminus \mathcal{O}^*$ (donde $\mathcal{O}^*$ denota el conjunto de elementos invertibles de $\mathcal{O}$).*

*Demostración.* Veamos que $P$ es un ideal, y comenzamos por observar que $0 \in P$. Sea $z \in P$ y $f \in \mathcal{O}$, si $zf =: u \in \mathcal{O}^*$ tenemos $(zu^{-1})f = 1$ y $f \in \mathcal{O}^*$, luego $z = uf^{-1} \in \mathcal{O}^*$ lo cual es absurdo, por lo tanto, $zf \in P$. Dado $f, g \in P \setminus \{0\}$ tenemos que $f/g \in \mathcal{O}$ o $g/f \in \mathcal{O}$, digamos $f/g \in \mathcal{O}$. Luego $1 + f/g \in \mathcal{O}$ y $f + g = g(1 + f/g) \in P$ por lo que acabamos de probar que $P$ es un ideal de $\mathcal{O}$. Obviamente es un ideal maximal porque si $J \subset \mathcal{O}$ es un ideal tal que $P \subsetneq J \subsetneq \mathcal{O}$, entonces $J$ debe contener un elemento de $\mathcal{O}^*$, de modo que $J = \mathcal{O}$. □

**Definición 2.3.** Un subconjunto $P \subset F$ que es un ideal maximal de algún anillo de valoración de $F \mid K$ se llama *lugar* de $F \mid K$.

El teorema a continuación enumera las propiedades importantes de los lugares.

**Teorema 2.4.** *Sea $\mathcal{O}$ un anillo de valoración de $F \mid K$ y sea $P \subset \mathcal{O}$ su ideal maximal. Entonces:*

(1) *$P$ es un ideal principal;*
(2) *sea $t \in P$ tal que $P = t\mathcal{O}$, luego cualquier elemento distinto de cero $z \in F$ se escribe de manera única como $z = t^n u$, con $n \in \mathbb{Z}$ y $u \in \mathcal{O}^*$;*
(3) *$z \in \mathcal{O}$ si y sólo si $z = t^n u$, con $n \in \mathbb{Z}$, $n \geq 0$ y $u \in \mathcal{O}^*$;*
(4) *el número entero $n$ es el mismo para cualquier generador de $P$.*

*Demostración.* (1) Véase [47, Teorema 1.1.6].
(2) Véase [47, Teorema 1.1.6]. Aunque no probaremos la existencia, la parte de unicidad del elemento es fácil de verificar. De hecho, suponga que $z = t^{n_1}u_1 = t^{n_2}u_2$ para $n_1, n_2 \in \mathbb{Z}$ y $u_1, u_2 \in \mathcal{O}^*$. Supongamos que $n_1 \geq n_2$, lo que implica que $1 = t^{n_1 - n_2}u_1u_2^{-1}$ y como $1 \notin P$ debemos tener $n_1 = n_2$, y luego $u_1 = u_2$.
(3) Si $z = t^n u$ con $u \in \mathcal{O}^*$ y $n < 0$, entonces no podemos tener $z \in \mathcal{O}$ porque en este caso $1 = zt^{-n}u^{-1} \in P$, lo cual es absurdo. La recíproca es obvia.
(4) Supongamos ahora que $P = t\mathcal{O} = z\mathcal{O}$, sabemos que existen $n \in \mathbb{Z}$ y $u \in \mathcal{O}$ únicos tales que $z = t^n u$, y como $z \in P \subset \mathcal{O}$ debemos tener $n \geq 0$. No podemos tener $n = 0$ ya que esto implicaría $z \in \mathcal{O}^*$ (y luego $1 \in \mathcal{O}$) por lo que $n > 0$. Como $P = z\mathcal{O} = t^n\mathcal{O}$ debemos tener $n = 1$ (porque $t \in \mathcal{O}$ y debido a la unicidad probada en el ítem (2)). Así $z = tu$ y dado $f \in F$, $f \neq 0$ tenemos $f = t^m v$ con $m \in \mathbb{Z}$ y $v \in \mathcal{O}$, entonces $f = z^m(u^m v)$ y así el entero $m$ es el mismo, independientemente si usamos $t$ o $z$. □

En la definición siguiente, $\infty$ denota un elemento tal que $\infty + \infty = n + \infty = \infty + n = \infty$ y $\infty > n$ para cualquier $n \in \mathbb{Z}$.

**Definición 2.5.** Una *valoración (discreta)* de $F \mid K$ es una función $v : F \to \mathbb{Z} \cup \{\infty\}$ con las siguientes propiedades:

(1) $v(f) = \infty \Leftrightarrow f = 0$;
(2) $v(fh) = v(f) + v(h)$ para cualquier $f, h \in F$;
(3) $v(f + h) \geq \text{mín}\{v(f), v(h)\}$ para cualquier $f, h \in F$;
(4) existe $t \in F$ tal que $v(t) = 1$;
(5) $v(a) = 0$ para cualquier $a \in K$, $a \neq 0$.

Observe que si $a \in K^*$ y $f \in F$ tenemos $v(af) = v(a)+v(f) = v(f)$, en particular $v(-f) = v(f)$. Una propiedad importante de las valoraciones discretas es la llamada *desigualdad triangular estricta*, que enunciamos a continuación.

**Lema 2.6.** *Sea $v$ una valoración discreta de $F \mid K$ y sean $f, h \in F$. Si $v(f) \neq v(h)$, entonces $v(f + h) = \text{mín}\{v(f), v(h)\}$.*

*Demostración.* Supongamos que $v(f) < v(h)$ y supongamos también que $v(f+h) \neq \text{mín}\{v(f), v(h)\} = v(f)$. De (3) obtenemos $v(f+h) > v(f)$ y $v(f) = v((f+h)-h) \geq \text{mín}\{v(f + h), v(-h)\} > v(f)$, una contradicción. □

**Definición 2.7.** Sea $P$ un lugar de $F \mid K$ y definamos una función $v_P : F \to \mathbb{Z} \cup \{\infty\}$ de la siguiente manera: $v(0) = \infty$; si $f \neq 0$ sea $t \in P$ tal que $P = t\mathcal{O}$ y escribamos $f = t^n u$, con $u \in \mathcal{O}^*$; luego tomamos $v(f) := n$. Esta función se llama la *valoración asociada a $P$*. Del Teorema 2.4 esta función está bien definida, es decir, no depende de la elección del generador para $P$ y para cualquier generador $t$ tenemos $v_P(t) = 1$. Un generador para $P$ también se llama *parámetro local en $P$*.

Dejamos al lector, como ejercicio, la prueba de que $v_P$ es de hecho una valoración discreta de $F \mid K$. El resultado a continuación puede ayudar a probar que $v_P$ tiene la propiedad (3) en la definición de valoraciones.

**Lema 2.8.** *Sea $P$ un lugar de $F \mid K$, $\mathcal{O}$ su anillo de valoración y sea $v_P$ la valoración asociada. Entonces:*

(1) $\mathcal{O} = \{f \in F \mid v_P(f) \geq 0\}$;
(2) $\mathcal{O}^* = \{f \in F \mid v_P(f) = 0\}$;
(3) $P = \{f \in F \mid v_P(f) > 0\}$.

*Demostración.* Sea $t \in P$ tal que $P = t\mathcal{O}$.
(1) Claramente, $t^n u \in \mathcal{O}$ siempre que $n \geq 0$ y $u \in \mathcal{O}^*$. Por otro lado, del Teorema 2.4 (3) tenemos $\mathcal{O} \subset \{f \in F \mid v_P(f) \geq 0\}$.
(2) Sea $f = t^n u \in \mathcal{O}^*$. Como $f^{-1} = t^{-n}u^{-1} \in \mathcal{O}$, debemos tener $n \geq 0$ y $-n \geq 0$. Por lo tanto $n = 0$.
(3) Esto es claro ya que $P = \mathcal{O} \setminus \mathcal{O}^*$. □

Por lo tanto, un lugar determina una valoración de $F \mid K$, y del ítem (1) del Lema anterior vemos que hay un único anillo de valoración que contiene un lugar dado. Es fácil probar el siguiente resultado, que es una especie de implicación inversa a la del lema: sea $v : F \to \mathbb{Z} \cup \{\infty\}$ una valoración de $F \mid K$ y sea $\mathcal{O} := \{f \in F \mid K \mid v(f) \geq 0\}$; entonces $\mathcal{O}$ es un anillo de valoración de $F \mid K$, que tiene como ideal maximal el conjunto $P := \{f \in F \mid v(f) > 0\}$. Uno incluso puede probar que hay una biyección entre los conjuntos de funciones de valoraciones discretas de $F \mid K$ y los lugares de $F \mid K$. Además, hay una biyección entre el conjunto de anillos de valoración de $F \mid K$ y el conjunto de lugares de $F \mid K$ (es decir, el conjunto de subconjuntos de $f$ que son ideales maximales para anillos de valoración de $F \mid K$).

Dado que $P$ es un ideal maximal de $\mathcal{O}$, obtenemos que $\mathcal{O}/P$ es un cuerpo. Además, como $K \subset \mathcal{O}$ podemos escribir $K \subset \mathcal{O}/P$, es decir, tenemos una extensión $(\mathcal{O}/P) \,|\, K$. Uno puede mostrar que esta es una extensión finita, y su grado se llama *grado de* $P$ (notación: $\deg P$). Sea $z \in \mathcal{O}$, es habitual escribir $z(P)$, en lugar de $\overline{z}$, para denotar la clase de $z$ en el cuerpo $\mathcal{O}/P$; esta notación se usará libremente en las secciones 3 y 4 a continuación.

Cuando $\deg P = 1$ decimos que $P$ es un *lugar racional de* $F \,|\, K$.

**Ejemplo 2.9.** Veamos dos ejemplos de cuerpos de funciones y examinemos sus anillos de valoración, lugares y valoraciones discretas.
1) Sea $K$ un cuerpo (que, en una primera lectura, se puede pensar que sea el cuerpo de números complejos $\mathbb{C}$) y considere el cuerpo de funciones $F \,|\, K$ donde $F = K(X) = \{f(X)/g(X) \,|\, f(X), g(X) \in K[X], g(X) \neq 0\}$. Esto cuerpo se llama *cuerpo de funciones racionales*). Sea $p(X) \in K[X]$ un polinomio irreducible (entonces, si $K = \mathbb{C}$ debemos tener $p(X) = X - \alpha$, para determinado $\alpha \in \mathbb{C}$). Recordamos que $K[X]$ es un dominio euclidiano, en particular es un dominio de factorización única. Luego, dado $f(X)/g(X) \in K(X)$, podemos encontrar un único $n \in \mathbb{Z}$ tal que $f(X)/g(X) = p(X)^n \tilde{f}(X)/\tilde{g}(X)$, donde $\tilde{f}(X), \tilde{g}(X) \in K[X]$ y ni $\tilde{f}(X)$ ni $\tilde{g}(X)$ son múltiplos de $p(X)$. Definiendo $v_P(f(X)/g(X)) = n$ y $v_P(0) = \infty$ obtenemos una función $v_P : F \to \mathbb{Z} \cup \{\infty\}$. No es difícil probar que $v_P$ es una valoración de $K(x)$. El anillo de valoración correspondiente es

$$\mathcal{O}_{p(X)} = \{f(X)/g(X) \in K(X) \,|\, g(X) \text{ no es un múltiplo de } p(X)\},$$

y el lugar correspondiente es el conjunto

$$P_{p(X)} = \{f(X)/g(X) \in K(X) \,|\, f(X) \text{ es un múltiplo de } p(X) \text{ pero } g(X) \text{ no es un múltiplo de } p(X)\} = p(X)\mathcal{O}_{p(X)};$$

entonces, por supuesto,

$$\mathcal{O}^*_{p(X)} = \{f(X)/g(X) \in K(X) \,|\, f(X) \text{ y } g(X) \text{ no son múltiplos de } p(X)\}.$$

Vamos a demostrar ahora que $\deg P_{p(X)} = \deg p(X)$. Sea $f(X)/g(X) \in \mathcal{O}_{p(X)}$. En $K[X]$ tenemos $\operatorname{mcd}(p(X), g(X)) = 1$ (ya que $p(X) \nmid g(X)$) por lo tanto, existen $t(X), s(X) \in K[X]$ tales que $t(X)p(X) + s(X)g(X) = 1$, luego $s(X)g(X) - 1 \in (p(X))$ y en $\mathcal{O}_{p(X)}/P_{p(X)}$ tenemos $\overline{s(X)/1} = \overline{1/g(X)}$; esto muestra que $\overline{f(X)/g(X)} = \overline{f(X)s(X)/1}$ en $\mathcal{O}_{p(X)}/P_{p(X)}$. Ahora sea $r(X)$ el resto en la división (euclidiana) de $f(X)s(X)$ por $p(X)$, luego obtenemos $\overline{f(X)s(X)/1} = \overline{r(X)/1}$ con $r(X) = 0$ o $\deg r(X) < \deg p(X)$. De aquí es fácil concluir que $\mathcal{O}_{p(X)}/P_{p(X)}$ se genera, como un $K$-espacio vectorial, por $\overline{1/1}$, $\overline{X/1}$, …, $\overline{X^{\deg(p(X))-1}/1}$.

Otra valoración de $K(x)$, generalmente denotada por $v_\infty$ es la función definida por $v_\infty(0) = \infty$ y $v_\infty(f(X)/g(X)) = \deg g(X) - \deg f(X)$ para todo $f(X)/g(X) \in K(X) \setminus \{0\}$. El anillo de valoración correspondiente es

$$\mathcal{O}_\infty = \{f(X)/g(X) \in K(X) \,|\, \deg f(X) \leq \deg g(X)\},$$

el lugar correspondiente es

$$P_\infty = \{f(X)/g(X) \in K(X) \,|\, \deg f(X) < \deg g(X)\} = (1/X)\mathcal{O}_\infty$$

y luego $\mathcal{O}^*_\infty = \{f(X)/g(X) \in K(X) \,|\, \deg f(X) = \deg g(X)\}$. Vamos a demostrar ahora que $\deg P_\infty = 1$. De hecho, en $\mathcal{O}_\infty/P_\infty$ tenemos $\overline{f(X)/g(X)} \neq \overline{0}$ si y solo

si $\deg f(X) = \deg g(X)$. Suponga que $\overline{f(X)/g(X)} \neq \overline{0}$ y sean $a$ y $b$, respectivamente, los coeficientes líderes de $f(X)$ y $g(X)$. Tenemos $\overline{f(X)/g(X)} - \overline{ab^{-1}/1} = \overline{(f(X) - ab^{-1}g(X))/g(X)} = \overline{0}$, y por lo tanto $\mathcal{O}_\infty/P_\infty = K$.

Se puede mostrar que estas son todas las valoraciones de $K(X)\,|\,K$ (ver [47, Proposition 1.2.1]).

Supongamos ahora que $K = \mathbb{C}$ y sea $p(X) = X - \alpha$, para algún $\alpha \in \mathbb{C}$ (observe que todos los polinomios irreducibles de $\mathbb{C}[X]$ son de esta forma). Entonces $\deg P_{x-\alpha} = 1$ y puede escribir $\mathcal{O}_{X-\alpha}/P_{X-\alpha} = \mathbb{C}$. De lo que hicimos arriba, sabemos que los elementos de $\mathcal{O}_{X-\alpha}/P_{X-\alpha}$ son de la forma $\overline{z(X)/1}$, donde $z(X) \in \mathbb{C}[X]$. Está claro que $\overline{X/1} = \alpha$ así que $\overline{z(X)/1} = z(\alpha)$. Esta es la motivación para la notación $z(P)$ presentada justo antes de este ejemplo.

2) Sea $f(X, Y) = Y^2 - X(X-1)(X-2)(X-3)(X-4)(X-5)(X-6) \in \mathbb{C}(X)[Y]$; no es difícil ver que $f(X, Y)$ (como un polinomio en la variable $Y$) es irreducible, luego $\mathbb{C}(X)[Y]/(f(X, Y))$ es un cuerpo. Escribimos

$$x := \overline{X},\ y := \overline{Y} \text{ y } \mathbb{C}(x, y) := \mathbb{C}(X)[Y]/(f(X, Y)).$$

Como $y^2 = \prod_{i=0}^{6}(x - i)$ vemos que cualquier elemento de $f$ puede escribirse como $f(x) + g(x)y$. Observe que $\mathbb{C}(x, y)\,|\,\mathbb{C}$ es un cuerpo de funciones.

$$\begin{array}{ll} \mathbb{C}(x,y) & \\ \big| \big) & \text{extensión finita de grado 2: } \mathbb{C}(x,y) = \mathbb{C}(x) \oplus \mathbb{C}(x)y \\ \mathbb{C}(x) & \\ \big| \big) & \text{extensión trascendente} \\ \mathbb{C} & \end{array}$$

Damos ahora alguna información sobre las valoraciones de $\mathbb{C}(x, y)$. Sea $v_\infty$ la valoración de $\mathbb{C}(x)$ que describimos antes, entonces la función $w_\infty : \mathbb{C}(x, y) \to \mathbb{Z} \cup \{\infty\}$ definida en un elemento no nulo por

$$w_\infty(f(x) + g(x)y) = \text{mín}\{2v_\infty(f(x)), 2v_\infty(g(x)) - 7\}$$

(y $w_\infty(0) = \infty$) es una valoración de $\mathbb{C}(x, y)\,|\,\mathbb{C}$. Para $i \in \{0, 1, 2, 3, 4, 5, 6\}$ tenemos que

$$w_i(f(x) + g(x)y) := \text{mín}\{2v_{x-i}(f(x)), 2v_{x-i}(g(x)) + 1\}$$

(y $w_i(0) = \infty$) también es una valoración (aquí $v_{x-i}$ es la valoración de $\mathbb{C}(x)$ asociada con el polinomio irreducible $x - i$, como se vio anteriormente), cuyo anillo de valoración (observe que $\mathbb{C}(x, y)$ es el cuerpo de fracciones de $\mathbb{C}[x, y] := \mathbb{C}[X, Y]/(f(X, Y))$) es ahora

$$\mathcal{O}_i = \left\{ \frac{p(x) + q(x)y}{m(x) + n(x)y} \in \mathbb{C}(x, y)\,|\,p(x), q(x), m(x), n(x) \in \mathbb{C}[x] \text{ y } m(i) \neq 0 \right\}.$$

Además, para $a \in \mathbb{C} \setminus \{0, 1, 2, 3, 4, 5, 6\}$, sea $b \in \mathbb{C}$ tal que $b^2 = \prod_{i=0}^{6}(a - i)$; luego $b \neq 0$ y existen dos valoraciones distintas de $\mathbb{C}(x, y)$, digamos $w_{a,b}$ y $w_{a,-b}$, con

anillos de valoración iguales a

$$\mathcal{O}_{a,b} = \left\{ \frac{p(x)+q(x)y}{m(x)+n(x)y} \in \mathbb{C}(x,y) \,|\, p(x), q(x), m(x), n(x) \in \mathbb{C}[x] \right. \\ \left. \text{y } m(a)+n(a)b \neq 0 \right\}$$

y

$$\mathcal{O}_{a,-b} = \left\{ \frac{p(x)+q(x)y}{m(x)+n(x)y} \in \mathbb{C}(x,y) \,|\, p(x), q(x), m(x), n(x) \in \mathbb{C}[x] \right. \\ \left. \text{y } m(a)-n(a)b \neq 0 \right\}.$$

Estos son todos los anillos de valoración de $\mathbb{C}(x,y)$. Además, todos los lugares son racionales. Esto es cierto porque, como se observó justo antes de estos ejemplos, para un lugar $P$ en un cuerpo de funciones $F \,|\, K$ la extensión $(\mathcal{O}_P/P) \,|\, K$ es finita, por lo tanto es una extensión algebraica, pero en nuestro caso $K = \mathbb{C}$, un cuerpo algebraicamente cerrado, entonces debemos tener $\mathcal{O}_P/P = \mathbb{C}$ para todos los lugares $P$ de $\mathbb{C}(x,y) \,|\, \mathbb{C}$.

**Definición 2.10.** Sea $\mathcal{P}$ el conjunto de lugares de $F \,|\, K$. Una suma formal de lugares $\sum_{P\in\mathcal{P}} n_P P$, donde $n_P \in \mathbb{Z}$ para todo $P \in \mathcal{P}$ y $n_P \neq 0$ solo para un número finito de lugares $P$, se llama *divisor* de $F \,|\, K$. El conjunto $\mathcal{D}$ de divisores es un grupo abeliano con la suma $\sum_{P\in\mathcal{P}} n_P P + \sum_{P\in\mathcal{P}} s_P P := \sum_{P\in\mathcal{P}} (n_P + s_P)P$. Definimos un orden parcial en el conjunto de divisores $\mathcal{D}$ como $\sum_{P\in\mathcal{P}} n_P P \leq \sum_{P\in\mathcal{P}} s_P P$ si y sólo si $n_P \leq s_P$ para todo $P \in \mathcal{P}$. El *grado* de un divisor es el número entero $\sum_{P\in\mathcal{P}} n_P \deg P$. El *soporte* de un divisor $\sum_{P\in\mathcal{P}} n_P P$ es el conjunto (finito) de lugares $P$ tal que $n_P \neq 0$.

Sea $f \in F$ con $f \neq 0$. Podemos mostrar que $v_P(f) \neq 0$ sólo para un número finito de lugares $P$ (ver [47, Corollary 1.3.4]). Si $P$ es tal que $v_P(f) > 0$, entonces se dice que $P$ es un *cero* de $f$; si $v_P(f) < 0$, entonces se dice que $P$ es un *polo* de $f$. Definimos el *divisor de* $f$ como $\operatorname{div}(f) := \sum_{P\in\mathcal{P}} v_P(f)P$. Observe que si $f, h \in F \setminus \{0\}$, a partir de las propiedades de las valoraciones, obtenemos $\operatorname{div}(fh) = \operatorname{div}(f) + \operatorname{div}(h)$; en particular, ya que $\operatorname{div}(1) = 0$, obtenemos $\operatorname{div}(f^{-1}) = -\operatorname{div}(f)$. Los divisores del tipo $\operatorname{div}(f)$, con $f \in F$, $f \neq 0$, se llaman *divisores principales* y el siguiente lema enumera algunas propiedades relacionadas con estos divisores (ver [47, Corollary 1.1.20 y Theorem 1.4.11]).

**Lema 2.11.** *Sea $f \in F \setminus K$. Entonces $f$ tiene al menos un cero y un polo. Además,* $\deg(\operatorname{div}(f)) = 0$.

**Ejemplo 2.12.** Sea $\mathbb{R}(X)$ el cuerpo de funciones racionales sobre los números reales $\mathbb{R}$, y sea $f = (X^2+1)(X^2+2)^3(X-1)^2/((X-3)(X-5)^4)$. Luego, del Ejemplo 2.9 (1) (utilizamos la notación de ese ejemplo) y las propiedades de las valoraciones vemos que $v_{X^2+1}(f) = 1$, $v_{X^2+2}(f) = 3$, $v_{X-1}(f) = 2$, $v_{X-3}(f) = -1$, $v_{X-5}(f) = -4$ y $v_\infty(f) = -5$, además $v_P(f) = 0$ para cualquier $P \notin \{P_{X^2+1}, P_{X^2+2}, P_{X-1}, P_{X-3}, P_{X-5}, P_\infty(f)\}$. Así $\operatorname{div}(f) = P_{X^2+1} + 3P_{X^2+2} + 2P_{X-1} - P_{X-3} - 4P_{X-5} - 5P_\infty(f)$, los ceros de $f$ son $\{P_{X^2+1}, P_{X^2+2}, P_{X-1}\}$, los polos de $f$ son $\{P_{X-3}, P_{X-5}, P_\infty(f)\}$ y, recordando que $\deg P_{X^2+1} = \deg P_{X^2+2} = 2$, es fácil comprobar que $\deg(\operatorname{div}(f)) = 0$.

**Definición 2.13.** Sea $D$ un divisor de $F \mid K$. Denotamos como $L(D)$ el conjunto $L(D) = \{f \in F \setminus \{0\} \mid \operatorname{div}(f) + D \geq 0\} \cup \{0\}$ (de manera equivalente, si $D = \sum_{P \in \mathcal{P}} n_P P$ entonces $L(D)$ es el conjunto de funciones $f$ tal que $v_P(f) + n_P \geq 0$ para todo $P \in \mathcal{P}$). De las propiedades (3) y (5) de las valoraciones vemos que $L(D)$ es un $K$-espacio vectorial, generalmente llamado *espacio de Riemann-Roch asociado a* $D$; denotaremos su dimensión por $\ell(D)$.

No es trivial, pero uno puede mostrar que $\ell(D)$ es finito para cualquier divisor $D$ de $F \mid K$.

**Definición 2.14.** Decimos que los divisores $D$ y $E$ son *linealmente equivalentes* si existe una función $f \in F$ tal que $E = D + \operatorname{div}(f)$.

La equivalencia lineal es una relación de equivalencia en el conjunto de divisores de $F \mid K$. A continuación, enumeramos algunos resultados relacionados con el espacio $L(D)$.

**Lema 2.15.** *Sean $D$ un divisor de $F \mid K$ y $f \in F$, $f \neq 0$. Sea $E := \operatorname{div}(f) + D$, entonces $L(D) \cong L(E)$.*

*Demostración.* Mostraremos que $\psi : L(D) \to L(E)$ definido por $\psi(h) = h/f$ es un isomorfismo de $K$-espacios vectoriales. Sea $h \in L(D)$, luego $\operatorname{div}(h) + D \geq 0$, por lo tanto $\operatorname{div}(h/f) + \operatorname{div}(f) + D \geq 0$, es decir $\operatorname{div}(h/f) + E \geq 0$; esto muestra que $\psi$ está bien definida. Está claro que $\psi$ es lineal e inyectiva, veamos que es suryectiva. Si $z \in L(E)$, entonces $\operatorname{div}(z) + E \geq 0$ por lo tanto $\operatorname{div}(z) + \operatorname{div}(f) + D \geq 0$, por lo tanto $zf \in L(D)$ y por supuesto $\psi(zf) = z$. □

**Lema 2.16.** *Sea $D$ un divisor de $F \mid K$ tal que $\deg D < 0$, entonces $\ell(D) = 0$ (i.e., $L(D) = \{0\}$).*

*Demostración.* Suponga que existe $f \in L(D)$, $f \neq 0$, luego $\operatorname{div}(f) + D \geq 0$. Observe que $\deg(\operatorname{div}(f) + D) = \deg(\operatorname{div}(f)) + \deg D = \deg D$ y tomando el grado en ambos lados de esta desigualdad obtenemos $\deg D \geq 0$, una contradicción con la hipótesis. □

**Lema 2.17.** *Si $D \geq 0$ tenemos $K \subset L(D)$ y si $D = 0$, entonces $L(D) = K$ (y tenemos $\ell(D) = 1$).*

*Demostración.* Suponga que $D \geq 0$ y sea $a \in K \setminus \{0\}$. Como $v_P(a) = 0$ para todos los lugares $P$ de $F \mid K$ obtenemos $\operatorname{div}(a) = 0$, por lo tanto $\operatorname{div}(a) + D \geq 0$ y tenemos $K \subset L(D)$. Ahora suponga $D = 0$; si $f \in L(D) \cap (F \setminus K)$ entonces del Lema 2.11 existe un lugar $P$ tal que $v_P(f) < 0$ por lo que no podemos tener $\operatorname{div}(f) + 0 \geq 0$ y luego $L(D) = K$. □

**Lema 2.18.** *Si $D \leq E$, entonces $L(D) \subset L(E)$ y $\dim(L(E)/L(D)) \leq \deg(E - D)$.*

*Demostración.* Si $h \in L(D)$ entonces $\operatorname{div}(h) + D \geq 0$, por lo tanto $\operatorname{div}(h) + E = \operatorname{div}(h) + D + (E - D) \geq E - D \geq 0$ y tenemos $h \in L(E)$. Supongamos ahora que $E = D + P$ y sea $n_P$ el coeficiente de $P$ en el divisor $D$; si $h \in L(D + P)$ tenemos $v_P(h) + n_p + 1 \geq 0$. Sea $t$ un parámetro local en $P$ y sea $\psi : L(D + P) \to \mathcal{O}_P/P$ la transformación $K$-lineal definida por $\psi(h) = \overline{ht^{n_P+1}}$, es fácil verificar que el núcleo de $\psi$ es exactamente $L(D)$ y por lo tanto $\dim(L(D + P)/L(D)) \leq \dim \mathcal{O}_P/P = \deg P$. Ahora el lema sigue por inducción y del hecho que si $U \subset W \subset V$ son espacios vectoriales $K$ de dimensión finita, entonces $\dim V/U = \dim V/W + \dim W/U$. □

El principal teorema sobre los espacios $L(D)$ es el llamado *Teorema de Riemann-Roch* (ver [47, Section I.5]) que indicamos a continuación.

**Teorema 2.19.** *Sea $F \mid K$ un cuerpo de funciones y sea $D$ un divisor de $F \mid K$. Existe un entero no negativo $g$ y un divisor $C$ tal que*

$$\ell(D) = \deg\, D + 1 - g + \ell(C - D).$$

*Además, $C$ puede ser reemplazado por cualquier divisor en su clase de equivalencia lineal.*

Este teorema presenta el invariante más importante de $F \mid K$, el entero $g$, que se llama el *género* de $F \mid K$. La clase de equivalencia de $C$ se llama *clase canónica* de divisores de $F \mid K$ y cualquier divisor en ella se llama *divisor canónico*.

**Corolario 2.20.** *Sea $C$ un divisor de $F \mid K$. Entonces $C$ es un divisor canónico si y sólo si* $\deg\, C = 2g - 2$ *y* $\ell(C) = g$.

*Demostración.* Supongamos que $C$ sea un divisor canónico; aplicando el teorema de Riemann-Roch a $D = 0$ y recordando del Lema 2.17 que $\ell(0) = 1$ obtenemos $1 = 0 + 1 - g + \ell(C)$, luego $\ell(C) = g$. Ahora aplicamos el teorema de Riemann-Roch a $D = C$ y usamos $\ell(C) = g$ y $\ell(C - C) = \ell(0) = 1$, así $g = \deg\, C + 1 - g + 1$ para que $\deg\, C = 2g - 2$. Para probar la recíproca, sea $C'$ un divisor canónico y suponga que $\deg\, C = 2g - 2$ y $\ell(C) = g$. Del teorema de Riemann-Roch aplicado a $D = C$ obtenemos $g = 2g - 2 + 1 - g + \ell(C' - C)$ y $\ell(C' - C) = 1$. Esto significa que existe $z \in F \setminus \{0\}$ tal que $\mathrm{div}(z) + (C' - C) \geq 0$, es decir, $\mathrm{div}(z) + C' \geq C$. Observe que ambos lados tienen el mismo grado, por lo tanto, debemos tener $\mathrm{div}(z) + C' = C$, por lo que $C$ es linealmente equivalente a $C'$ y, por lo tanto, es un divisor canónico. □

**Corolario 2.21.** *Sea $D$ un divisor de $F \mid K$ tal que* $\deg D > 2g - 2$. *Entonces* $\ell(D) = \deg D + 1 - g$.

*Demostración.* Sea $C$ un divisor canónico, luego $\deg C = 2g - 2$ y obtenemos $deg(C - D) = \deg C - \deg D < 0$. Así, del Lema 2.16 obtenemos $\ell(C - D) = 0$ y del teorema de Riemann-Roch tenemos $\ell(D) = \deg D + 1 - g$. □

**Ejemplo 2.22.** El corolario anterior proporciona un método para calcular el género de un cuerpo de funciones, aunque generalmente no sea práctico. La idea es calcular $\deg(D) + 1 - \ell(D)$ para divisores cuyos grados crecen hasta el infinito, obteniendo así una secuencia de enteros que debe ser constante después de que el grado "sea lo suficientemente grande".
1) Usaremos este método para calcular el género de $K(x)$. Sea

$$P_\infty = \{f(X)/g(X) \in K(X) \mid \deg f(X) < \deg g(X)\}$$

(consulte el Ejemplo 2.9 (1) para recordar las definiciones y la notación que usaremos aquí) y sea $n$ un entero positivo. Entonces

$$L(nP_\infty) = \{z \in K(X) \mid v_\infty(z) + n \geq 0 \text{ y } v_{p(X)}(z) \geq 0 \text{ para todo irreducible } p(X) \in K[X]\}.$$

Vamos a escribir $z \in K(X)$ como $z = \prod_{p(X)} p(X)^{n_{p(X)}}$, donde $p(X)$ recorre el conjunto de polinomios irreducibles en $K[X]$ y $n_{p(X)}$ es un número entero (por supuesto, $n_{p(X)} = 0$ excepto por un número finito de $p(X)$). Al recordar la definición de $v_\infty$ vemos que $L(nP_\infty) = \{f(X) \in K[X] \mid \deg\, f(X) \leq n\} \cup \{0\}$, luego $\ell(nP_\infty) = n+1 = \deg\, nP_\infty + 1$. Esto muestra que $K(x)$ tiene género cero. Uno puede demostrar

la recíproca: si $F \mid K$ tiene género cero y un lugar racional, entonces $F = K(X)$ (vea [47, Proposition 1.6.3]).

2) También aplicaremos el método anterior para calcular el género del cuerpo de funciones $\mathbb{C}(x, y)$ descrito en el Ejemplo 2.9 (2) (consulte este ejemplo para las definiciones y la notación). Es un poco más elaborado, pero vamos a delinear el razonamiento. Para $n$ un entero positivo, queremos determinar $L(nP_\infty)$ para calcular $\ell(nP_\infty)$. Tenemos

$$L(nP_\infty) = \{z \in \mathbb{C}(x, y) \mid w_\infty(z) + n \geq 0 \text{ y } w(z) \geq 0 \text{ siempre que } w \neq w_\infty\}.$$

Para $L' := \{z \in \mathbb{C}(x, y) \mid w(z) \geq 0 \text{ siempre que } w \neq w_\infty\}$, vamos a mostrar que

$$L' = \{p(x) + q(x)y \in \mathbb{C}(x, y) \mid p(x), q(x) \in \mathbb{C}[x]\}.$$

Sabemos que los elementos de $\mathbb{C}(x, y)$ pueden escribirse como $p(x) + q(x)y$ con $p(x), q(x) \in \mathbb{C}(x)$ y sea $\sigma : \mathbb{C}(x, y) \to \mathbb{C}(x, y)$ el automorfismo de $\mathbb{C}(x, y)$ definido por $\sigma(p(x) + q(x)y) = p(x) - q(x)y$ (en otras palabras, $\sigma$ es el automorfismo de $\mathbb{C}(x, y)$ definido por $\sigma(x) = x$ y $\sigma(y) = -y$). Sea $w$ una valoración de $\mathbb{C}(x, y)$. No es difícil comprobar que la composición $w \circ \sigma$ es también una valoración de $\mathbb{C}(x, y)$; además, si $P \neq P_\infty$ tenemos $w_P \circ \sigma \neq w_\infty$. De hecho, $(w_P \circ \sigma)(x) = w_P(\sigma(x)) = w_P(x) \geq 0$ ya que $x \in \mathcal{O}_P$, mientras que $w_\infty(x) = \text{mín}\{2 \cdot (-1), \infty\} = -2$. De esto podemos concluir que $\sigma(L') \subset L'$ porque si $z \in L'$ entonces $w(\sigma(z)) = (w \circ \sigma)(z) \geq 0$ para todos $w \neq w_\infty$. Por lo tanto, si $p(x) + q(x)y \in L'$ entones $p(x) - q(x)y \in L'$ y luego $p(x), q(x)y \in L'$. Ahora $L' \cap \mathbb{C}(x) = \{a(x) \in \mathbb{C}(x) \mid w(a(x)) \geq 0 \text{ siempre que } w \neq w_\infty\} = \{a(x) \in K(x) \mid a(x) \in \mathcal{O}_P \text{ para todos los lugares } P \neq P_\infty\}$ y de la descripción de los anillos de valoración en el Ejemplo 2.9 (2) obtenemos $L' \cap \mathbb{C}(x) = \mathbb{C}[x]$. Por lo tanto, $p(x) \in \mathbb{C}[x]$ y $(q(x)y)^2 = q(x)^2 \prod_{i=0}^{6}(x - i) \in \mathbb{C}[x]$; escribiendo $q(x)$ como el cociente de dos polinomios vemos que debemos tener $q(x) \in \mathbb{C}[x]$. Así $L' = \{p(x) + q(x)y \in \mathbb{C}(x, y) \mid p(x), q(x) \in \mathbb{C}[x]\}$ y como $L(nP_\infty) = L' \cap \{z \in F \mid w_\infty(z) + n \geq 0\}$ obtenemos

$$\begin{aligned} L(nP_\infty) &= \{p(x) + q(x)y \mid p(x), q(x) \in \mathbb{C}[x] \text{ y } w_\infty(p(x) + q(x)y) \geq -n\} \\ &= \{p(x) + q(x)y \mid p(x), q(x) \in \mathbb{C}[x],\ 2v_\infty(p(x)) \geq -n \\ &\qquad\qquad \text{y } 2v_\infty(q(x)) - 7 \geq -n\} \\ &= \{p(x) + q(x)y \mid p(x), q(x) \in \mathbb{C}[x],\ \deg(p(x)) \leq n/2 \\ &\qquad\qquad \text{y } \deg(q(x)) \leq (n-7)/2\}. \end{aligned}$$

De esto concluimos que

$$L(nP_\infty) = \begin{cases} [n/2] + 1 & \text{si } 1 \leq n \leq 6, \\ n + 1 - 3 & \text{si } n \geq 7, \end{cases}$$

y luego el género de $\mathbb{C}(x, y)$ es 3.

Necesitaremos, en una demostración en la próxima sección, el resultado a continuación (ver [47], sección 4.2, y especialmente el Teorema 4.2.6, para obtener resultados más generales). Recordemos que el anillo *de la serie formal de Laurent* sobre un cuerpo $K$, en la variable $t$, es el anillo $K((t)) := \{\sum_{i=n}^{\infty} a_i t^i \mid n \in \mathbb{Z}, a_i \in K \text{ para todo } i \geq n\}$ (este conjunto es un anillo con la suma y el producto habituales de la serie). Recuerde también que un *cuerpo perfecto* $K$ es un cuerpo de característica cero o, si la característica es $p > 0$, entonces cada elemento tiene una raíz $p$-ésima en $K$. Por lo tanto, los cuerpos finitos son ejemplos de cuerpos perfectos (este es el ejemplo que nos interesará, especialmente en las dos últimas secciones).

**Teorema 2.23.** *Sea $F \mid K$ un cuerpo de funciones donde $K$ es un cuerpo perfecto. Sea $t$ un parámetro local en un lugar racional $Q$. Luego, existe un monomorfismo de anillos $\Phi : F \to K((t))$ que asocia a cada elemento $z \in F \setminus \{0\}$ una serie $\Phi(z) = \sum_{i=n}^{\infty} a_i t^i$, donde $a_n \neq 0$ y $n = v_P(z)$. La serie $\Phi(z)$ se denomina* expansión local de $z$ en $Q$.

**2.2. Semigrupos de Weierstrass de varios puntos.** En lo que sigue, denotaremos el conjunto de enteros no negativos como $\mathbb{N}_0$. Sea $F \mid K$ un cuerpo de funciones de una variable y sea $f \in F$, $f \neq 0$. Ya vimos que $f$ tiene un número finito de ceros y polos.

**Definición 2.24.** Llamamos al divisor $\mathrm{div}_0(f) := \sum_{P \in \mathcal{P} \text{ y } v_p(f)>0} v_P(f)P$ *divisor de ceros* de $f$, mientras que el divisor $\mathrm{div}_\infty(f) := \sum_{P \in \mathcal{P} \text{ y } v_p(f)<0} (-v_P(f))P$ se llamará *divisor de polos* de $f$.

Observe que $\mathrm{div}_0(f) \geq 0$, $\mathrm{div}_\infty(f) \geq 0$, $\mathrm{div}(f) = \mathrm{div}_0(f) - \mathrm{div}_\infty(f)$ y que $\deg(\mathrm{div}_0(f)) = \deg(\mathrm{div}_\infty(f))$.

**Definición 2.25.** El conjunto $H(Q) = \{n \in \mathbb{N}_0 \mid \exists f \in F \text{ tal que } \mathrm{div}_\infty(f) = nQ\}$ se denomina *semigrupo de Weierstrass asociado a $Q$*, o el *semigrupo de Weierstrass en $Q$*.

Recordamos que un *semigrupo* es una estructura algebraica que consiste en un conjunto no vacío $H$ sobre el cual está definida una operación binaria asociativa y un *monoide* es un semigrupo que tiene un elemento neutro relativo a la operación (es decir, existe $0 \in H$ tal que $h + 0 = 0 + h = h$ para todos los $h \in H$). Observe que el conjunto $H(Q)$ es un subsemigrupo de $\mathbb{N}_0$, ya que si $n_1, n_2 \in H(Q)$ entonces existen $f_1, f_2 \in F$ tales que $\mathrm{div}_\infty(f_1) = n_1 Q$ y $\mathrm{div}_\infty(f_2) = n_2 Q$ y tomando $f := f_1 f_2$ obtenemos $\mathrm{div}_\infty(f) = (n_1 + n_2)Q$, es decir $n_1 + n_2 \in H(Q)$. En realidad, como $0 \in H(Q)$ (porque $\mathrm{div}_\infty(1) = 0Q$) tenemos que $H(Q)$ es un submonoide de $\mathbb{N}_0$, pero llamar a $H(Q)$ un semigrupo ya es un procedimiento estándar.

Un resultado muy útil al estudiar $H(Q)$ es el siguiente.

**Lema 2.26.** *Existe $f \in F \setminus \{0\}$ tal que $\mathrm{div}_\infty(f) = nQ$ si y sólo si $L((n-1)Q) \subsetneqq L(nQ)$.*

*Demostración.* Tenemos: $\mathrm{div}_\infty(f) = nQ$ para algún $f \in F \Leftrightarrow \mathrm{div}(f) + nQ \geq 0$ pero $\mathrm{div}(f) + (n-1)Q \not\geq 0 \Leftrightarrow f \in L(nQ) \setminus L((n-1)Q)$. □

Sea $g$ el género de $F \mid K$. Si $n \geq 2g$, entonces del teorema de Riemann-Roch obtenemos $\ell(nQ) = n + 1 - g$, $\ell((n-1)Q) = (n-1) + 1 - g$, por lo tanto $\ell(nQ) - \ell((n-1)Q) = 1$, así que $L((n-1)Q) \subsetneqq L(nQ)$. Por lo tanto $\{n \in \mathbb{N}_0 \mid n \geq 2g\} \subset H(Q)$ y en particular, si $g = 0$, entonces $H(Q) = \mathbb{N}_0$. Supongamos que $g = 1$. Luego de Corolario 2.21 obtenemos $\ell(Q) = 1$ y como $\ell(0) = 1$ debemos tener $H(Q) = \{n \in \mathbb{N}_0 \mid n = 0 \text{ o } n \geq 2\}$. Entonces, asumimos que $g \geq 2$ y analizamos el caso donde $n \in \{1, \ldots, 2g-1\}$. Observe que $\ell(0Q) = 1$ y $\ell((2g-1)Q) = g$ y del Lema 2.18 obtenemos $0 \leq \ell(nQ) - \ell((n-1)Q) \leq 1$, por lo tanto, cuando $n$ va de 0 a $2g-1$ existen $2g-1$ "saltos", mientras que la dimensión $\ell(nQ)$ salta $g-1$ veces, de 1 a $g$. Por lo tanto, podemos concluir que hay exactamente $g-1$ elementos $n \in \{1, \ldots, 2g-1\}$ satisfaciendo $L((n-1)Q) \subsetneqq L(nQ)$ y de lo que se hizo arriba obtenemos

$$H(Q) = \{n \in \mathbb{N}_0 \mid \exists f \in K \text{ tal que } \mathrm{div}_\infty(f) = nQ\} = \{0, \gamma_1, \ldots, \gamma_{g-1}\} \cup \{n \in \mathbb{N}_0 \mid n \geq 2g\}$$

donde $0 < \gamma_1 < \ldots < \gamma_{g-1} \leq 2g-1$, y luego $\#(\mathbb{N}_0 \setminus H(Q)) = g$, para cualquier $g \geq 0$.

**Definición 2.27.** El conjunto $\mathbb{N}_0 \setminus H(Q)$ se llama *conjunto de lagunas* o *secuencia de lagunas* en $Q$ y sus $g$ elementos se llaman *lagunas de Weierstrass en* $Q$; los elementos de $H(Q)$ a menudo se denominan *no-lagunas en* $Q$.

El semigrupo $H(Q)$ es un objeto clásico y muy estudiado, de enorme importancia en el estudio de los cuerpos de funciones y de las curvas algebraicas. En la Sección 2.4 daremos algunas aplicaciones de estos semigrupos a la teoría de codificación, pero su importancia va mucho más allá. Suponiendo que $K$ es un cuerpo algebraicamente cerrado, se puede probar que para casi todos los lugares de $F \,|\, K$ (es decir, todos los lugares, excepto un número finito) la secuencia de lagunas es la misma (y si $K = \mathbb{C}$, entonces esta secuencia de lagunas es $\{1, \ldots, g\}$) y los lugares que tienen una secuencia de lagunas diferente se llaman *lugares de Weierstrass*. El lector interesado puede encontrar más información sobre los semigrupos y lugares de Weierstrass en [26, Sección 4.4], [29, Section 7.6] y [49]. Por ahora, demostramos un resultado simple que usaremos más tarde.

**Lema 2.28.** *Sea $Q$ un lugar racional de un cuerpo de funciones $F \,|\, K$, sea $\alpha$ una laguna en $Q$, y sea $C$ un divisor canónico de $F \,|\, K$. Entonces existe un elemento $h \in F$ tal que* $\operatorname{div}(h) + C = (\alpha - 1)Q + E$, *con $E \geq 0$ y $Q \notin \operatorname{supp} E$.*

*Demostración.* Si $\alpha$ es una laguna en $Q$, entonces $\ell(\alpha Q) = \ell((\alpha-1)Q)$. Del teorema de Riemann-Roch obtenemos $\alpha+1-g+\ell(C-\alpha Q) = (\alpha-1)+1-g+\ell(C-(\alpha-1)Q)$ y, por lo tanto, $\ell(C-(\alpha-1)Q) = \ell(C-\alpha Q)+1$. Entonces existe $h \in F$ tal que $\operatorname{div}(h) + C - (\alpha-1)Q \geq 0$ pero $\operatorname{div}(h) + C - \alpha Q \not\geq 0$. Así, tomando $E := \operatorname{div}(h)+C-(\alpha-1)Q$ tenemos $E \geq 0$, $\operatorname{div}(h)+C = (\alpha-1)Q+E$ y $Q \notin \operatorname{supp} E$. $\square$

Presentamos ahora una generalización directa del concepto de semigrupo de Weierstrass en un punto, a un conjunto de varios puntos. Por lo que sabemos, apareció por primera vez en [7, p. 366], pero su estudio sistemático comenzó con Kim ([36]) y Homma ([30]). Sea $m$ un entero positivo. Observe que $\mathbb{N}_0^m$ es un semigrupo con la adición coordenada a coordenada de $m$-tuplas.

**Definición 2.29.** Sea $m$ un entero positivo y sea $Q_1, \ldots, Q_m$ lugares de $F \,|\, K$ de grado uno. El subsemigrupo de $\mathbb{N}_0^m$ dado por

$$H(Q_1, \ldots, Q_m) := \{(\alpha_1, \ldots, \alpha_m) \in \mathbb{N}_0^m : \\ \exists f \in F \text{ con } \operatorname{div}_\infty(f) = \alpha_1 Q_1 + \ldots + \alpha_m Q_m\}$$

se llama *semigrupo de Weierstrass en* $Q_1, \ldots, Q_m$. A partir de las propiedades de las valoraciones, es fácil comprobar que se trata de un subsemigrupo de $\mathbb{N}_0^m$.

El conjunto complementario $G(Q_1, \ldots, Q_m) := \mathbb{N}_0^m \setminus H(Q_1, \ldots, Q_m)$ se llama el *conjunto de lagunas en* $Q_1, \ldots, Q_m$, y los elementos de $H(Q_1, \ldots, Q_m)$ se llaman *no-lagunas*.

De nuevo, observamos que $H(Q_1, \ldots, Q_m)$ es en realidad un monoide conmutativo, ya que, por ejemplo, $\operatorname{div}_\infty(1) = 0 = \sum_{i=1}^m 0Q_i$ y así $(0, \ldots, 0) \in H(Q_1, \ldots, Q_m)$.

En lo que sigue, usaremos la siguiente notación. Denotamos por $\mathbf{0}$ la $m$-tupla de $\mathbb{N}_0^m$ teniendo todas las entradas igual a cero; cuando escribimos $\boldsymbol{\alpha} \in \mathbb{N}_0^m$ debe entenderse que las entradas de esta $m$-tupla son $\boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_m)$ (de manera similar para $\boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}_0^m$). Además, para $i \in \{1, \ldots, m\}$ denotamos por $\boldsymbol{e}_i$ la $m$-tupla

que tiene todas las entradas iguales a cero, excepto la $i$-ésima entrada, que es igual a 1 y por $v_i$ nos referimos a la valoración asociada a $Q_i$. Si $\boldsymbol{\alpha} \in \mathbb{N}_0^m$ anotamos como $L(\boldsymbol{\alpha})$ al espacio vectorial de Riemann-Roch $L(\alpha_1 Q_1 + \cdots + \alpha_m Q_m)$ y por $\ell(\boldsymbol{\alpha})$ nos referimos a $\dim L(\boldsymbol{\alpha})$. Sumamos $m$-tuplas de $\mathbb{N}_0^m$ y las multiplicamos por enteros de la manera habitual. Denotamos la $m$-tupla de los puntos $(Q_1, \ldots, Q_m)$ por $\mathbf{Q}_m$ y generalmente escribiremos $H(\mathbf{Q}_m)$ en lugar de $H(Q_1, \ldots, Q_m)$ y $G(\mathbf{Q}_m)$ en lugar de $G(Q_1, \ldots, Q_m)$.

Siempre asumiremos que $\#K \geq m$, si $K$ es un cuerpo finito.

**Lema 2.30.** *Sea $\boldsymbol{\alpha} \in \mathbb{N}_0^m \backslash \{\mathbf{0}\}$. Entonces las siguientes afirmaciones son equivalentes*:

(1) $\boldsymbol{\alpha} \in H(\mathbf{Q}_m)$;
(2) $\ell(\boldsymbol{\alpha}) = \ell(\boldsymbol{\alpha} - \boldsymbol{e}_i) + 1$, *para todo $i \in \{1, \ldots, m\}$ tal que $\alpha_i > 0$.*

*Demostración.* Tenemos que (1) implica (2) simplemente porque si $\boldsymbol{\alpha} \in H(\mathbf{Q}_m)$ entonces existe $f \in F$ tal que $\mathrm{div}_\infty(f) = \alpha_1 Q_1 + \cdots + \alpha_m Q_m$, luego $f \in L(\boldsymbol{\alpha}) \setminus L(\boldsymbol{\alpha} - \boldsymbol{e}_i)$ para todo $i \in \{1, \ldots, m\}$ tal que $\alpha_i > 0$. Para ver que (2) implica (1), suponga que $\alpha_i > 0$ para algún $i \in \{1, \ldots, m\}$ y sean $f_1, \ldots, f_m \in F$ tales que $v_i(f_i) = -\alpha_i$ y $v_j(f_i) \geq -\alpha_j$ para todo $j \in \{1, \ldots, m\}$. Vamos a mostrar que existen elementos $\alpha_1, \ldots, \alpha_m \in K$ tales que el divisor de polos de $\sum_{i=1}^m \alpha_i f_i$ es precisamente $\sum_{i=1}^m \alpha_i Q_i$. Para cada $i = 1, \ldots, m$, sea $t_i$ un parámetro local en $Q_i$. Sea

$$f_i = a_{i,j} t_j^{v_j(f_i)} + \cdots \in K((t_j))$$

la expansión local de $f_i$ en $Q_j$ (cf. Teorema 2.23). Luego, $\mathrm{div}_\infty(\sum_{i=1}^m \alpha_i f_i) \neq \sum_{i=1}^m \alpha_i Q_i$ si y sólo si existe $j \in \{1, \ldots, m\}$ tal que $\alpha_j > 0$ (es decir, $v_j(f_j) = -\alpha_j < 0$) y $v_j(\sum_{i=1}^m \alpha_i f_i) > -\alpha_j$), es decir, $\sum_{i=1}^m \alpha_i a_{i,j} = 0$; por lo tanto, para tener $\mathrm{div}_\infty(\sum_{i=1}^m \alpha_i f_i) = \sum_{i=1}^m \alpha_i Q_i$ es suficiente elegir una $m$-tupla $(\alpha_1, \ldots, \alpha_m) \in K^m$ fuera de la unión de (como máximo) $m$ subespacios lineales de dimensión $m-1$. Cada uno de estos subespacios lineales tiene $(\#K)^{m-1}$ elementos y el origen es un elemento común a todos estos subespacios. Por lo tanto, debemos evitar como máximo un total de $m((\#K)^{m-1} - 1) + 1$ elementos y esto es posible porque $\#K \geq m$, luego $\#K^m > m(\#K)^{m-1} - m + 1$; y la prueba está completa. □

**Lema 2.31.** *Sea $\boldsymbol{\alpha} \in \mathbb{N}_0^m$ tal que $\alpha_i > 0$ para algún $i \in \{1, \ldots, m\}$. Las siguientes afirmaciones son equivalentes.*

(1) $\ell(\boldsymbol{\alpha}) = \ell(\boldsymbol{\alpha} - \boldsymbol{e}_i) + 1$;
(2) $\{\boldsymbol{\beta} \in H(\mathbf{Q}_m) \,|\, \beta_i = \alpha_i \text{ y } \beta_j \leq \alpha_j \text{ para todo } j = 1, \ldots, m\} \neq \emptyset$.

*Demostración.* Si $\ell(\boldsymbol{\alpha}) = \ell(\boldsymbol{\alpha} - \boldsymbol{e}_i) + 1$ entonces existe $f \in F$ tal que $\mathrm{div}_\infty(f) \leq \sum_{j=1}^m \alpha_j Q_j$ y $v_i(f) = -\alpha_i$. Así, escribiendo $\mathrm{div}_\infty(f) = \sum_{j=1}^m \beta_j P_j$ tenemos que $\boldsymbol{\beta} \in H(\mathbf{Q}_m)$, con $\beta_i = \alpha_i$ y $\beta_j \leq \alpha_j$ para todo $j = 1, \ldots, m$.

La implicación (2) $\Rightarrow$ (1) es fácil. □

En lo que sigue vamos a denotar el conjunto

$$\{\boldsymbol{\beta} \in H(\mathbf{Q}_m) \,|\, \beta_i = \alpha_i \text{ y } \beta_j \leq \alpha_j \text{ para todo } j = 1, \ldots, m\} \neq \emptyset$$

por $\nabla_i(\boldsymbol{\alpha})$, $i \in \{1, \ldots, m\}$.

Nos gustaría llamar la atención del lector sobre dos observaciones importantes: 1) El número de lagunas es siempre finito. De hecho, dado $\boldsymbol{\alpha} \in \mathbb{N}_0^m$, si $\sum_{j=1}^m \alpha_j \geq 2g$ entonces, del teorema de Riemann-Roch tenemos que $\ell(\boldsymbol{\alpha}) = \sum_{j=1}^n \alpha_j + 1 - g$ y para todo $i \in \{1, \ldots, m\}$ tales que $\alpha_i > 0$ obtenemos $\ell(\boldsymbol{\alpha} - \boldsymbol{e}_i) = (\sum_{j=1}^n \alpha_j - 1) + 1 - g = \ell(\boldsymbol{\alpha}) - 1$ y por lo tanto $\boldsymbol{\alpha} \in H(\mathbf{Q}_m)$.

2) Sea $i \in \{1, \ldots, m\}$, existe una biyección entre el semigrupo de Weierstrass (habitual) $H(Q_i)$ y el conjunto $\{\boldsymbol{\beta} \in H(\mathbf{Q}_m) \mid \boldsymbol{\beta} = a\boldsymbol{e}_i \text{ para algún } a \in \mathbb{N}_0\}$, definida por $a \mapsto a\boldsymbol{e}_i$ para todos los $a \in H(Q_i)$.

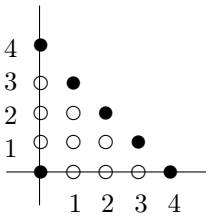
Como consecuencia de los lemas anteriores, tenemos el siguiente resultado.

**Corolario 2.32.** *Sea $\boldsymbol{\alpha} \in \mathbb{N}_0^m$ y supongamos que $\#K \geq m$. Entonces las siguientes afirmaciones son equivalentes:*

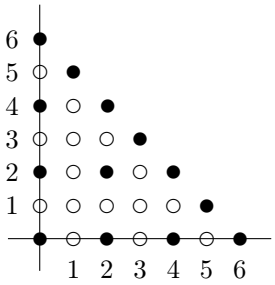
(1) $\boldsymbol{\alpha} \in G(\mathbf{Q}_m)$;
(2) *existe* $i \in \{1, \ldots, m\}$ *tal que* $\alpha_i > 0$ *y* $\ell(\boldsymbol{\alpha}) = \ell(\boldsymbol{\alpha} - \boldsymbol{e}_i)$;
(3) *existe* $i \in \{1, \ldots, m\}$ *tal que* $\alpha_i > 0$ *y* $\nabla_i(\boldsymbol{\alpha}) = \emptyset$.

En [36], Kim comenzó el estudio de $H(\mathbf{Q}_m)$ tratando el caso $m = 2$ (aunque él trabajó bajo las hipótesis de que $K$ es algebraicamente cerrado y char$K = 0$, sus resultados son válidos para cuerpos perfectos de cualquier característica). Presentamos a continuación dos ejemplos de semigrupos de Weierstrass, basados en ejemplos de su trabajo.

**Ejemplo 2.33.** Sea $F = \mathbb{C}[x, y]$, donde $y^2 = \prod_{j=0}^{6}(x - j)$.
(i) Sean $Q_1$ y $Q_2$ los lugares tales que $\ell(Q_1 + Q_2) = 1$ (y luego no existe una función $f \in F \setminus K$ tal que $Q_1 + Q_2 \geq \text{div}_\infty(f)$). Tome, por ejemplo, los lugares asociados a los puntos $(7!, \sqrt{7!})$ y $(8!, \sqrt{8!})$. Se puede mostrar que en este caso $H(Q_1, Q_2)$ es el subsemigrupo de $\mathbb{N}_0^2$ cuyas lagunas se indican abajo como círculos vacíos:




Por otro lado, si consideramos que $Q_1$ y $Q_2$ son puntos de Weierstrass (por ejemplo, $(0, 0)$ y $(1, 0)$), entonces $H(Q_1, Q_2)$ es el subsemigrupo de $\mathbb{N}_0^2$ que tiene las siguientes lagunas:




Como podemos observar a partir de estos ejemplos, el número de lagunas puede variar con la elección de $Q_1, \ldots, Q_m$. Uno puede mostrar que

$$\#G(\mathbf{Q}_m) \geq \binom{m+g}{g}$$

(ver [36] para el caso $m = 2$, [33] para el caso $m \geq 2$ y [9] para una prueba de este hecho en una situación más general).

Como señalan Homma y Kim (véase [31]), para las aplicaciones de la teoría de la codificación, el siguiente concepto es importante.

**Definición 2.34.** Una *laguna de Weierstrass pura* de $H(\mathbf{Q}_m)$ es una laguna $\boldsymbol{\alpha}$ tal que, para todo $i \in \{1, \ldots, m\}$, tenemos $\alpha_i > 0$ y $\ell(\boldsymbol{\alpha}) = \ell(\boldsymbol{\alpha} - \boldsymbol{e}_i)$ (o, de manera equivalente, $\nabla_i(\boldsymbol{\alpha}) = \emptyset$). El conjunto de lagunas puras se denotará por $G_0(\mathbf{Q}_m)$.

De la equivalencia de (2) y (3) en el Corolario 2.32 obtenemos que $(1, 1)$, $(1, 2)$ y $(2, 1)$ son lagunas puras del semigrupo de Weierstrass del Ejemplo 2.33 (i), mientras que en el Ejemplo 2.33 (ii) las lagunas puras son los puntos $(a, b)$ con $a, b \in \{1, 3, 5\}$.

**Lema 2.35.** *Si $\boldsymbol{\alpha} \in G_0(\mathbf{Q}_m)$, entonces $\alpha_i$ es una laguna de Weierstrass en $Q_i$, para todo $i \in \{1, \ldots, m\}$.*

*Demostración.* Sea $i \in \{1, \ldots, m\}$. De $\nabla_i(\boldsymbol{\alpha}) = \emptyset$ tenemos $\alpha_i \boldsymbol{e}_i \notin H(\mathbf{Q}_m)$, luego $\alpha_i \notin H(Q_i)$. □

**Lema 2.36.** *Las siguientes afirmaciones son equivalentes:*

(i) $\boldsymbol{\alpha} \in G_0(\mathbf{Q}_m)$;
(ii) $\ell(\boldsymbol{\alpha}) = \ell(\boldsymbol{\alpha} - \mathbf{1})$.

*Demostración.* Para demostrar que (i) implica (ii) supongamos que existe $f \in L(\boldsymbol{\alpha}) \setminus L(\boldsymbol{\alpha} - \mathbf{1})$, entonces debemos tener $v_i(f) = -\alpha_i$ para $i \in \{1, \ldots, m\}$, luego $\nabla_i(\boldsymbol{\alpha}) \neq \emptyset$ y $\boldsymbol{\alpha} \notin G_0(\mathbf{Q}_m)$. Para demostrar la recíproca, observamos que para cualquier $i \in \{1, \ldots, m\}$ tenemos $\ell(\boldsymbol{\alpha}) \geq \ell(\boldsymbol{\alpha} - \boldsymbol{e}_i) \geq \ell(\boldsymbol{\alpha} - \mathbf{1})$, luego las hipótesis implican $\ell(\boldsymbol{\alpha}) = \ell(\boldsymbol{\alpha} - \boldsymbol{e}_i)$ para todo $i \in \{1, \ldots, m\}$. □

**2.3. Códigos de Goppa.** Un *código* es un conjunto formado por combinaciones de un conjunto de símbolos que se llama el *alfabeto* del código. Una combinación del alfabeto que pertenece al código se llama *palabra* del código. Los códigos se utilizan para transmitir información, generalmente a través de un medio que puede manipular las palabras del código, transformándolas en otras palabras o en combinaciones que no están en el código. Los llamados *códigos correctores de errores* tienen la propiedad de que, si el medio no cambia "demasiado" una palabra, la palabra original puede recuperarse de la palabra recibida. Esta propiedad se deriva del hecho de que cada palabra, en dichos códigos, transporta no solo información, sino también cierta "redundancia de información" que se utiliza para recuperar palabras que han sido cambiadas.

Uno de los tipos más comunes de códigos correctores de errores es el *código lineal*.

**Definición 2.37.** Un *código lineal de longitud $n$* sobre un cuerpo finito $K$ es simplemente un subespacio vectorial de $K^n$. Sea $\mathcal{C} \subset K^n$ un código lineal y sean $v, w \in \mathcal{C}$. La *distancia de Hamming* entre $v = (v_1, \ldots, v_n)$ y $w = (w_1, \ldots, w_n)$ se define como $d(v, w) := \#\{i \mid v_i \neq w_i\}$. La *distancia mínima de un código* es el número $\min\{d(v, w) \mid v, w \in \mathcal{C}, v \neq w\}$, o equivalentemente, $\min\{d(v, 0) \mid v \in \mathcal{C}, v \neq 0\}$.

Siempre trabajaremos con códigos lineales y por eso omitiremos la palabra lineal. Un *código* $[n, k, d]$ es un código de longitud $n$, dimensión $k$ y distancia mínima $d$. Para aplicaciones, generalmente queremos códigos con una distancia mínima $d$ grande, ya que no es difícil mostrar que si un medio de transmisión cambia como máximo $(d - 1)/2$ entradas en una palabra $w$, convirtiéndola en una $n$-tupla $w'$,

entonces $w$ será la única palabra del código cuya distancia para $w'$ es como máximo $(d-1)/2$ y decodificaremos la $n$-tupla $w'$ recibida como $w$. También es deseable tener una gran dimensión $k$, ya que esto significa que el código puede transmitir grandes cantidades de información. Sin embargo, para un $n$ de longitud fija, no podemos tener grandes $d$ y $k$, debido a la cota de Singleton.

**Lema 2.38** (Cota de Singleton)**.** *Sea $C$ un código $[n,k,d]$. Entonces*

$$k+d \leq n+1.$$

*Demostración.* Sea $w$ el subespacio de $K^n$ definido por $W = \{(a_1,\ldots,a_n) \in K^n \mid a_i = 0 \text{ para todo } i \geq d\}$. Como $d(w,0) \leq d-1$ para todo $w \in W$, tenemos $W \cap C = \{0\}$. De $\dim W = d-1$ obtenemos $k+(d-1) \leq n$, que prueba el lema. □

**Definición 2.39.** Si $\mathcal{C} \subset K^n$ es un código, entonces

$$\mathcal{C}^{\perp} := \{w \in K^n \mid \langle w, v\rangle = 0 \text{ para todo } v \in \mathcal{C}\}$$

se llama *código dual* de $\mathcal{C}$ (aquí $\langle\ ,\ \rangle$ denota el producto interno habitual en $K^n$).

Presentamos ahora la construcción, debido a V.D. Goppa, que usa material de las secciones anteriores, para obtener los llamados *códigos (geométricos) de Goppa.*

Sea $F \mid K$ un cuerpo de funciones de una variable. Sean $G$ y $D$ divisores de $F \mid K$ con soporte disjunto, con $D$ de la forma $D = P_1 + \cdots + P_n$, donde $P_1,\ldots,P_n$ son lugares racionales de $F \mid K$. Sea $\varphi : L(G) \to K^n$ la función definida por $\varphi(f) = (f(P_1),\ldots,f(P_n))$ (recuerde que $f(P)$ es la clase de $f \in \mathcal{O}_P$ en $\mathcal{O}_p/P$, cf. pág. 14); luego, $\varphi$ es una transformación lineal de $K$-espacios vectoriales y $\varphi(L(G))$ es un código de longitud $n$; denotaremos este código como $\mathcal{C}(D,G)$. Una ventaja de esta construcción es que de inmediato podemos calcular o encontrar una estimación de los principales parámetros del código.

**Teorema 2.40.** *$\mathcal{C}(D,G)$ es un código $[n,k,d]$, donde $k = \dim L(G) - \dim L(G-D)$ y $d \geq n - \deg G$.*

*Demostración.* Observe que $\varphi$ es una transformación lineal suryectiva sobre $\mathcal{C}(D,G)$ y

$$\begin{aligned}
\operatorname{Ker}(\varphi) &= \{f \in L(G) \mid f(P_i) = 0 \text{ para todo } i = 1,\ldots,n\}\\
&= \{f \in L(G) \mid v_{P_i}(f) > 0 \text{ para todo } i = 1,\ldots,n\}\\
&= \{f \in L(G) \mid \operatorname{div}_0(f) \geq D\}\\
&= \{f \in F \mid \operatorname{div}_0(f) \geq D \text{ y } \operatorname{div}_0(f) - \operatorname{div}_\infty(f) + G \geq 0\}\\
&= \{f \in F \mid \operatorname{div}(f) + G - D \geq 0\} = L(D-G).
\end{aligned}$$

La definición de distancia mínima tiene sentido sólo si el código tiene un elemento distinto de cero, lo que asumimos ahora. Sea $x \in L(G)$ tal que $\varphi(x) \neq 0$ (en particular $x \neq 0$) y tal que $d = d(\varphi(x),0)$. Luego, existe un subconjunto $\Lambda \subset \{1,\ldots,n\}$ de cardinalidad $n-d$ tal que si $i \in \Lambda$ entonces $v_{P_i}(x) > 0$ (y $v_{P_j}(x) \neq 0$ si $j \notin \Lambda$). Por lo tanto $x \in L(G - \sum_{i\in\Lambda} P_i)$, luego $\deg(G - \sum_{i\in\Lambda} P_i) \geq 0$ y tenemos $d \geq n - \deg G$. □

En lo que sigue, denotamos por $g$ el género de $F \mid K$.

**Corolario 2.41.** *Supongamos que $\deg G < \deg D = n$. Entonces $\varphi : L(G) \to \mathcal{C}(D,G)$ es inyectiva y:*

(1) $\mathcal{C}(D,G)$ *es un código* $[n,k,d]$ *con* $d \geq n - \deg G$ *y* $k = \dim G \geq \deg G + 1 - g$ *(luego* $k + d \geq n + 1 - g$*);*
(2) *suponiendo además que* $2g - 2 < \deg G < n$*, tenemos* $k = \deg G + 1 - g$.

*Demostración.* Si $\deg(G-D) < 0$ entonces $\dim L(G-D) = 0$ y del teorema anterior tenemos $k = \dim L(G)$, por lo tanto, $\varphi$ es inyectiva. Del teorema de Riemann-Roch obtenemos $\dim L(G) = \deg G + 1 - g + \dim L(C - G) \geq \deg G + 1 - g$, donde $C$ es un divisor canónico; si $\deg G < \deg C = 2g - 2$ entonces $\dim L(C - G) = 0$, luego $k = \deg G + 1 - g$. □

En estas notas llamamos al número $n - \deg G$ la *cota de Goppa para la distancia mínima de* $\mathcal{C}(D,G)$.

Se puede probar [47, Proposition 2.2.10] que el código que es el dual de $\mathcal{C}(D,G)$ es el código $\mathcal{C}(D, D - G + C)$, donde $C$ es un divisor canónico tal que $D - G + C$ es un divisor cuyo soporte es disjunto del soporte de $D$ (en el curso de la determinación de $\mathcal{C}(D,G)^\perp$ uno ve que de hecho existe tal divisor canónico). No lo probaremos, pero demostramos al menos que $\mathcal{C}(D, D - G + C)$ tiene la dimensión correcta. Pero primero, presentamos algunos hechos sobre los parámetros de $\mathcal{C}(D,G)^\perp$.

**Corolario 2.42.** $\mathcal{C}(D,G)^\perp$ *es un código* $[n,k',d']$*, con* $k' = \ell(C-(G-D))-\ell(C-G)$ *y* $d' \geq \deg G - (2g-2)$*. Si* $\deg G > 2g - 2$ *entonces* $k' = \ell(C - (G - D)) \geq n + g - 1 - \deg G$ *y si* $2g - 2 < \deg G < n$*, entonces* $k' = n + g - 1 - \deg G$.

*Demostración.* La primera afirmación es una consecuencia del Teorema 2.40 y del hecho de que $\mathcal{C}(D,G)^\perp = \mathcal{C}(D, D - G + C)$. Las otras afirmaciones se obtienen de la primera y del teorema de Riemann-Roch. □

Observe que $\dim \mathcal{C}(D,G) + \dim \mathcal{C}(D, D - G + C) = \ell(G) - \ell(G - D) + \ell(D - G + C) - \ell(C - G) = \ell(G) - \ell(C - G) - (\ell(G - D) - \ell(C - (G - D))) = \deg G + 1 - g - (\deg G - \deg D + 1 - g) = \deg D = n$, que es el resultado esperado.

En estas notas llamaremos al número $\deg G - (2g - 2)$ *la cota de Goppa para la distancia mínima de* $\mathcal{C}(D,G)^\perp$.

**2.4. Semigrupos de Weierstrass y códigos de Goppa.** En esta sección pretendemos mostrar cómo utilizar la información de los semigrupos de Weierstrass en varios puntos para construir códigos de Goppa que tengan cotas para la distancia mínima mejor que la cota de Goppa.

Sea $F \mid K$ un cuerpo de funciones de una variable; en esta sección, $K$ es siempre un cuerpo finito. Comenzamos con una aplicación del semigrupo de Weierstrass habitual para construir códigos con una distancia mínima "grande" y que apareció en un trabajo de García, Kim y Lax (ver [18]).

**Teorema 2.43.** *Sean* $Q, P_1, \ldots, P_n$ *lugares racionales distintos de* $F \mid K$*. Supongamos que hay un conjunto de* $t + 1$ *lagunas consecutivas en* $H(Q)$*, digamos* $\gamma - t, \ldots, \gamma - 1, \gamma$ *(donde* $t$ *es un entero no negativo). Para* $G := \gamma Q$*, si el código* $\mathcal{C}(D,G)$ *tiene una dimensión positiva, entonces* $n - \deg G + t + 1$ *es una cota inferior para su distancia mínima.*

*Demostración.* Supongamos que existe $f \in L(G)$ tal que $w := d(\varphi(f), 0) \leq n - \deg G + t$ (recuerde que $\varphi$ es la función que aparece en la construcción de $\mathcal{C}(D,G)$ —vea la página 25); esto significa que existe un subconjunto $\Lambda \subset \{1, \ldots, n\}$ tal que $\#\Lambda = n - w$ y $v_{P_i}(f) > 0$ para todo $i \in \Lambda$ (y por supuesto, $v_{P_j}(f) = 0$ para $j \in \{1, \ldots, n\} \setminus \Lambda$). Luego $\operatorname{div}(f) + G > \sum_{i \in \Lambda} P_i$ por lo tanto $E := \operatorname{div}(f) +$

$G - \sum_{i\in\Lambda} P_i \geq 0$, además $\deg E = \deg G - n + w \leq t$, entonces escribimos $E = \lambda Q + E'$, donde $E' \geq 0$ y $Q \notin \operatorname{supp} E'$; observe que $0 \leq \lambda \leq t$. Por lo tanto, $\operatorname{div}(f) = \lambda Q + E' - G + \sum_{i=1}^{n-w} P_i = -(\gamma - \lambda)Q + E' + \sum_{i\in\Lambda} P_i$, de modo que $\gamma - \lambda \in H(Q)$. □

El teorema anterior muestra que la existencia de $t+1$ lagunas de Weierstrass consecutivas en el semigrupo $H(Q)$ de un lugar racional $Q$ permite la construcción de un código de Goppa cuya cota para la distancia mínima puede mejorarse mediante $t+1$ en comparación con la cota de Goppa. El caso $t = 0$ de este teorema fue probado por Janwa (ver [34]).

Presentamos ahora un teorema de García y Lax que muestra cómo construir un código dual con una cota inferior mejor que la cota de Goppa para la distancia mínima (ver [19]).

**Teorema 2.44.** *Sean $Q, P_1, \ldots, P_n$ lugares racionales distintos de $F \mid K$ y sean $\alpha$ y $\beta$ lagunas de Weierstrass en $Q$. Sean $D := P_1 + \cdots + P_n$ y $G := (\alpha + \beta - 1)Q$. Si $\dim \mathcal{C}(D,G)^\perp > 0$, entonces $\deg G - (2g-2) + 1$ es una cota inferior para la distancia mínima de $\mathcal{C}(D,G)^\perp$.*

*Demostración.* Recordemos que $\mathcal{C}(D,G)^\perp = \mathcal{C}(D, D-G+C)$, donde $C$ es un divisor canónico tal que $\operatorname{supp}(D-G+C) \cap \operatorname{supp}(D) = \emptyset$. Sea $\varphi : L(G) \to \mathcal{C}(D, D-G+C)$ como en la sección anterior. Suponga que hay un elemento $f \in L(G)$ tal que $w := d(\varphi(f), 0) \leq \deg G - (2g-2)$. Como $\deg G - (2g-2)$ es una cota inferior para la distancia mínima, debemos tener $w = \deg G - (2g-2)$. Entonces existe $\Gamma \subset \{1, \ldots, n\}$ tal que $\#\Gamma = w$ y $v_{P_i}(f) = 0$ para todo $i \in \Gamma$ (y $v_{P_j}(f) > 0$ para todo $j \in \{1, \ldots, n\} \setminus \Gamma$). Por lo tanto, $\operatorname{div}(f) + C + D - G \geq \sum_{j\in\{1,\ldots,n\}\setminus\Gamma} P_j$ y así $\operatorname{div}(f) + C \geq G - \sum_{j\in\Gamma} P_j$. Observe que ambos lados de esta desigualdad tienen el mismo grado, luego $\operatorname{div}(f) + C = G - \sum_{j\in\Gamma} P_j$. Dado que $\alpha$ es una laguna en $Q$ tenemos, del Lema 2.28, que existe $h \in F$ tal que $\operatorname{div}(h) + C = (\alpha - 1)Q + E$, con $E \geq 0$ y $Q \notin \operatorname{supp} E$. Por lo tanto, $\operatorname{div}(h/f) = \operatorname{div}(h) - \operatorname{div}(f) = \sum_{j\in\Gamma} P_j + E - \beta Q$ y tenemos $\beta \in H(Q)$, contradiciendo una hipótesis. □

En [18] los autores presentan la siguiente generalización del resultado anterior (que no demostraremos aquí).

**Teorema 2.45.** *Sean $Q, P_1, \ldots, P_n$ lugares racionales distintos de $F \mid K$. Supongamos que $\alpha, \ldots, \alpha + t, \beta - (t-1), \ldots, \beta - 1, \beta$ son lagunas de Weierstrass en $Q$, con $\alpha + t \leq \beta$ y $t \geq 1$. Sean $D := P - 1 + \cdots + P_n$ y $G := (\alpha + \beta - 1)Q$; si $\dim \mathcal{C}(D,G)^\perp > 0$, entonces $\deg G - (2g-2) + t + 1$ es una cota inferior para la distancia mínima de $\mathcal{C}(D,G)^\perp$.*

**Corolario 2.46.** *Sean $Q, P_1, \ldots, P_n$ lugares racionales distintos de $F \mid K$. Supongamos que $\alpha, \ldots, \alpha + t$ son $t+1$ lagunas de Weierstrass consecutivas en $Q$. Sean $D := P_1 + \cdots + P_n$ y $G := (2\alpha + t - 1)Q$; si $\dim \mathcal{C}(D,G)^\perp > 0$, entonces $\deg G - (2g-2) + t + 1$ es una cota inferior para la distancia mínima de $\mathcal{C}(D,G)^\perp$.*

En 2001, en su tesis doctoral, escrita bajo la supervisión de R.F. Lax, Gretchen Matthews dio la primera aplicación de los semigrupos de Weierstrass en más de un punto a la teoría de codificación (ver [27]). Uno de sus principales resultados en ese trabajo es el siguiente.

**Teorema 2.47.** *Sean $Q_1, Q_2, P_1, \ldots, P_n$ lugares racionales distintos de $F \mid K$. Suponga que $(\alpha_1, \alpha_2) \in G(Q_1, Q_2)$ con $\alpha_1 > 1$ y $\ell(\alpha_1 Q_1 + \alpha_2 Q_2) = \ell((\alpha_1 -$*

$1)Q_1 + \alpha_2 Q_2)$. *Suponga además que* $(\gamma_1, \gamma_2 - t - 1) \in G(Q_1, Q_2)$ *para todo* $t$, $0 \leq t \leq min\{\gamma_2 - 1, 2g - 1 - (\alpha_1 + \alpha_2)\}$. *Sean* $G = (\alpha_1 + \gamma_1 - 1)Q_1 + (\alpha_2 + \gamma_2 - 1)Q_2$, *y* $D = P_1 + \cdots + P_n$. *Si la dimensión de* $\mathcal{C}(D, G)^\perp$ *es positiva, entonces la distancia mínima de este código es al menos* $\deg G - (2g - 2) + 1$.

Poco después de la publicación del trabajo de Matthews, un trabajo de Homma y Kim introdujo el concepto de "lagunas puras", que vimos en la Sección 2.2 y lo usaron para obtener códigos con cotas inferiores mejores para la distancia mínima. Homma y Kim trabajaron con el semigrupo de Weierstrass en dos puntos, y sus resultados han sido generalizados al caso de los semigrupos de Weierstrass en $m$ puntos, $m$ cualquier entero positivo, por C. Carvalho y F. Torres (ver [10]). Para presentar los principales resultados de [10], comenzamos con un resultado de Homma y Kim (ver [31]), ligeramente modificado para estas notas.

**Lema 2.48.** *Sean* $B$, $E$, $N$ *y* $M$ *divisores de* $F \,|\, K$, *con* $B \geq 0$. *Si*

(1) $N + M + E$ *es un divisor canónico;*
(2) $\ell(M - B) = \ell(M)$;
(3) $\ell(N) \leq \ell(N + E)$;
(4) $\ell(N) = \ell(N + B)$;

*entonces* $\deg B \leq \deg E$.

*Demostración.* Por supuesto, tenemos que $N + M + E = (N + B + E) + (M - B)$ es un divisor canónico, por lo tanto, del teorema de Riemann-Roch obtenemos $\ell(N + B + E) = \deg N + \deg B + \deg E + 1 - g + \ell(M - B)$ y $\ell(N + E) = \deg N + \deg E + 1 - g + \ell(M)$, donde $g$ es el género de $F \,|\, K$. Usando (2) obtenemos $\ell(N + B + E) - \ell(N + E) = \deg B$. Ahora, de (3), (4) y el teorema de Riemann-Roch, obtenemos $\deg B = \ell(N + B + E) - \ell(N + E) \leq \ell(N + B + E) - \ell(N) = \ell(N + B + E) - \ell(N + B) \leq \deg E$. □

**Lema 2.49.** [31, Lema 3.1] *Sean* $M$ *y* $B$ *divisores de* $F \,|\, K$ *con* $B \geq 0$ *y* $\ell(M) = \ell(M - B)$. *Si* $R \geq 0$ *es un divisor que satisface* $\operatorname{supp} R \cap \operatorname{supp} B = \emptyset$, *entonces* $\ell(M - R) = \ell(M - R - B)$.

*Demostración.* Observe que $L(M - B) \subset L(M)$ y $L(M - R) \subset L(M)$; de $\operatorname{supp} R \cap \operatorname{supp} B = \emptyset$ obtenemos $L(M - B) \cap L(M - R) = L(M - R - B)$, por lo tanto debemos tener $L(M - R) = L(M - R - B)$ ya que $L(M) = L(M - B)$. □

Uno de los principales resultados de [10] es el siguiente.

**Teorema 2.50.** [10, Teorema 3.3] *Sean* $Q_1, \ldots, Q_m, P_1, \ldots, P_n$ *lugares racionales distintos de* $F \,|\, K$. *Supongamos que* $(\alpha_1, \ldots, \alpha_m)$ *y* $(\beta_1, \ldots, \beta_m)$ *son lagunas puras en* $Q_1, \ldots, Q_m$. *Sea* $G = \sum_{i=1}^{m} (\alpha_i + \beta_i - 1)Q_i$ *y* $D = P_1 + \cdots + P_n$. *Si la dimensión de* $\mathcal{C}(D, G)^\perp$ *es positiva, entonces a distancia mínima de este código es al menos* $\deg G - (2g - 2) + m$.

*Demostración.* Recuerde que $\mathcal{C}(D, G)^\perp = \mathcal{C}(D, D - G + C)$, donde $C$ es un divisor canónico tal que $\operatorname{supp}(D - G + C) \cap \operatorname{supp}(D) = \emptyset$ y sea $\varphi : L(D - G + C) \to \mathcal{C}(D, D - G + C)$ la función definida en la página 25 (es decir, $\varphi(h) = (h(P_1), \ldots, h(P_n))$ para todo $h \in L(D - G + C)$). Sea $f \in L(G)$ y sea $w = d(\varphi(f), 0)$, luego existe un subconjunto $\Lambda \subset \{1, \ldots, n\}$ tal que $\#\Lambda = w$, $v_{P_i}(f) = 0$ si $i \in \Lambda$ y $v_{P_i}(f) > 0$ si $i \in \{1, \ldots, n\} \setminus \Lambda$. Por lo tanto, $\operatorname{div}(f) + D - G + C \geq \sum_{i \in \{1, \ldots, n\} \setminus \Lambda} P_i$ (aquí estamos usando que $\operatorname{supp}(D - G + C) \cap \operatorname{supp}(D) = \emptyset$ y que $\operatorname{supp}(G) \cap \operatorname{supp}(D) = \emptyset$) por lo

tanto $\operatorname{div}(f) + C - G + \sum_{i\in\Lambda} P_i \geq 0$. Sean $B := \sum_{i=1}^{m} Q_i$, $E := C - G + \sum_{i\in\Lambda} P_i$, $N := \sum_{i=1}^{m}(\alpha_i - 1)Q_i$ y $M := \sum_{i=1}^{m} \beta_i Q_i - \sum_{i\in\Lambda} P_i$. Entonces $E + N + M$ es el divisor canónico $C$, $\ell(N) = \ell(N + B)$ por Lema 2.36 y $\ell(M - B) = \ell(M)$ por Lema 2.36 y el lema anterior. También tenemos $\ell(N) \leq \ell(N + \operatorname{div}(f) + E)$, porque $\operatorname{div}(f) + E \geq 0$; de $N \geq 0$ tenemos $\operatorname{div}(f) + E + N \geq 0$, por lo tanto, $f \in L(E + N)$ y $\ell(\operatorname{div}(f) + E + N) = \ell(E + N)$, luego $\ell(N) \leq \ell(N + E)$. Por lo tanto, del Lema 2.48 obtenemos $\deg B \leq \deg E$, es decir, $m \leq 2g - 2 - \deg G + w$, de modo que $w \geq \deg G - (2g - 2) + m$. □

Otro resultado principal de [10] es el siguiente.

**Teorema 2.51.** [10, Teorema 3.4] *Sean $Q_1, \ldots, Q_m, P_1, \ldots, P_n$ lugares racionales distintos de $F \mid K$. Sean $(\alpha_1, \ldots, \alpha_m), (\beta_1, \ldots, \beta_m) \in \mathbb{N}_0^m$ tales que $\alpha_i \leq \beta_i$ para todo $i \in \{1, \ldots, m\}$ y asuma que cada $m$-tupla $(\gamma_1, \ldots, \gamma_m) \in \mathbb{N}_0^m$ con $\alpha_i \leq \gamma_i \leq \beta_i$ para todo $i \in \{1, \ldots, m\}$ es una laguna pura en $Q_1, \ldots, Q_m$. Sea $G = \sum_{i=1}^{m}(\alpha_i + \beta_i - 1)Q_i$ y $D = P_1 + \cdots + P_n$. Si la dimensión de $\mathcal{C}(D, G)^{\perp}$ es positiva, entonces la distancia mínima de este código es al menos* $\deg G - (2g - 2) + m + \sum_{i=1}^{m}(\beta_i - \alpha_i)$.

*Demostración.* Sean $\varphi$, $f$, $w$ y $\Lambda$ como en el comienzo de la prueba anterior. Entonces, como arriba, tenemos $\operatorname{div}(f) + C - G + \sum_{i\in\Lambda} P_i \geq 0$. Ahora sea $B := \sum_{i=1}^{m}(\beta_i - \alpha_i + 1)Q_i$ y como arriba, tome $E := C - G + \sum_{i\in\Lambda} P_i$, $N := \sum_{i=1}^{m}(\alpha_i - 1)Q_i$ y $M := \sum_{i=1}^{m} \beta_i Q_i - \sum_{i\in\Lambda} P_i$. Entonces $E + N + M$ es un divisor canónico y $N + B = \sum_{i=1}^{m} \beta_i Q_i$. Así, a partir de la definición de lagunas puras y del Lema 2.36, obtenemos $\ell(N + B) = \ell(N)$. Además, como $\ell(\sum_{i=1} \beta_i Q_i) = \ell(\sum_{i=1}(\alpha_i - 1)Q_i) = \ell(\sum_{i=1} \beta_i Q_i - B)$ y $\operatorname{supp} B \cap \operatorname{supp}(\sum_{i\in\Lambda} P_i) = \emptyset$ obtenemos del Lema 2.49 que $\ell(\sum_{i=1}^{m} \beta_i Q_i - \sum_{i\in\Lambda} P_i) = \ell(\sum_{i=1} \beta_i Q_i - B - \sum_{i\in\Lambda} P_i)$, es decir, $\ell(M) = \ell(M - B)$. Como en la prueba anterior, tenemos $\ell(N) \leq \ell(N + E)$, así que del Lema 2.48 obtenemos $\deg B \leq \deg E$, es decir, $\sum_{i=1}^{m}(\beta_i - \alpha_i + 1) \leq 2g - 2 - \deg G + w$, luego $w \geq \deg G - (2g - 2) + m + \sum_{i=1}^{m}(\beta_i - \alpha_i)$. □

Observe que, de alguna manera, este resultado extiende los resultados 2.43 y 2.44, si pensamos que todos las lagunas puras en la hipótesis son "¡lagunas consecutivas!".

## 3. Curvas maximales

Al escribir el material de esta sección, intentamos presentar algunos de los problemas que se estudian hoy en esta área. La selección de los temas es algo personal y solo muestra una pequeña parte del panorama completo. Esperamos que sea un buen punto de partida para aquellos estudiantes interesados en este tema.

El propósito de esta sección es estudiar curvas definidas sobre cuerpos finitos, interesados particularmente en curvas con "muchos" puntos racionales. Tales curvas resultan muy útiles en Teoría de Códigos ya que para códigos fabricados a partir de tales curvas, la distancia mínima es grande.

Algunas veces, para probar los resultados es más fácil usar el lenguaje de cuerpos de funciones y a veces, para trabajar con ejemplos concretos es más fácil usar el lenguaje de curvas, así que comenzaremos mostrando que existe un diccionario que nos permite pasar de un lenguaje al otro sin problemas.

**3.1. Definiciones básicas.** Los detalles y pruebas de la parte de curvas pueden ser encontrados en [28], [15], [45] y [46], mientras que los detalles sobre cuerpos de funciones en [47].

Sea $\mathbb{F}_q$ un cuerpo finito con $q$ elementos, denotamos por $\bar{\mathbb{F}}_q$ el cierre galoisiano. Podemos pensar que el espacio proyectivo puede ser obtenido a partir del espacio afín, agregando puntos en el infinito.

**Definición 3.1.** El conjunto $\mathbb{A}^n(\bar{\mathbb{F}}_q) = \{(a_1, \ldots, a_n) : a_i \in \bar{\mathbb{F}}_q \ \forall i\}$ es llamado de *n-espacio afín.*

El *espacio proyectivo*, que denotaremos por $\mathbb{P}^n(\bar{\mathbb{F}}_q)$ es el conjunto formado por las clases de equivalencia cuando consideramos la siguiente relación definida sobre $\mathbb{A}^{n+1}$:

$$(a_0, \ldots, a_n) \sim (b_0, \ldots, b_n) \Leftrightarrow \exists \lambda \in \bar{\mathbb{F}}_q^* \text{ tal que } a_i = \lambda b_i \ \forall i.$$

Así, un punto $P \in \mathbb{P}^n(\bar{\mathbb{F}}_q)$ será denotado por $P = (a_0 : \cdots : a_n)$.

Note que si $\sigma \in G = Gal(\bar{\mathbb{F}}_q, \mathbb{F}_q)$ entonces $\sigma$ actúa sobre $\mathbb{P}^n(\bar{\mathbb{F}}_q)$ como sigue:

$$\sigma((a_0 : \cdots : a_n)) = (\sigma(a_0) : \cdots : \sigma(a_n)).$$

En particular, el automorfismo de Frobenius nos permite caracterizar el conjunto $\mathbb{P}^n(\mathbb{F}_q)$ de puntos $\mathbb{F}_q$-racionales (o sea, puntos cuyas coordenadas pertenecen a $\mathbb{F}_q$) ya que estos puntos son dejados fijos por este automorfismo.

Otra consecuencia de este hecho es que el número de puntos racionales de $\mathbb{P}^n(\bar{\mathbb{F}}_q)$ es $q^{n+1}/(q-1)$.

**Definición 3.2.** Considere $P \in \mathbb{P}^n(\bar{\mathbb{F}}_q)$, entonces

1. El conjunto $\mathbf{P} = \{\sigma(P) : \sigma \in G\}$ es un *punto cerrado* sobre $\mathbb{F}_q$.
2. La cardinalidad del conjunto $\{\sigma(P) : \sigma \in G\}$ es el *grado* de $\mathbf{P}$.

Si $P = (a_0 : \cdots : a_n)$ es un elemento de $\mathbf{P}$ con $a_i \neq 0$ entonces

$$\mathbf{P} = \{\sigma(P) : \sigma \in G\} = \{\sigma(P) : \sigma \in Gal(\mathbb{F}_q(a_0/a_i, \ldots, a_n/a_i); \mathbb{F}_q)\}$$

y el grado de $\mathbf{P}$ es igual al grado de la extensión $\mathbb{F}_q(a_0/a_i, \ldots, a_n/a_i)|\mathbb{F}_q)$.

Vamos a definir qué es una variedad afín. Sea $S$ un subconjunto de $\bar{\mathbb{F}}_q[x_1, \ldots, x_n] = \bar{\mathbb{F}}_q[X]$. Definimos el conjunto $V(S)$ de ceros de $S$ por

$$V(S) := \{P \in \mathbb{A}^n(\bar{\mathbb{F}}_q) : f(P) = 0 \ \forall f \in S\}.$$

**Definición 3.3.** Un *conjunto algebraico afín* es cualquier conjunto de la forma $V(S)$ para algún $S \subseteq \bar{\mathbb{F}}_q[X]$.

Diremos que el conjunto $V(S)$ está definido sobre $\mathbb{F}_q$ si $S \subseteq \mathbb{F}_q[X]$ y será denotado en este caso por $V|\mathbb{F}_q$.

El conjunto de puntos racionales de $V|\mathbb{F}_q$ es

$$V(\mathbb{F}_q) = V \cap \mathbb{A}^n(\mathbb{F}_q).$$

**Definición 3.4.** Un conjunto algebraico afín $V$ es *reducible* si existen dos conjuntos algebraicos afines $V_1$ y $V_2$ tales que $V_i \neq V$ para $i = 1, 2$ y $V = V_1 \cup V_2$. Si $V$ es no vacío y no reducible, diremos en este caso que $V$ es *irreducible.*

Un conjunto afín $V|\mathbb{F}_q$ tiene asociado los ideales

$$I(V) = \{f \in \bar{\mathbb{F}}_q[X] : f(P) = 0 \ \forall P \in V\},$$
$$I(V|\mathbb{F}_q) = I(V) \cap \mathbb{F}_q[X].$$

*Observación* 3.5. Un conjunto algebraico afín es irreducible si y solamente si $I(V)$ es un ideal primo de $\bar{\mathbb{F}}_q[X]$.

**Definición 3.6.** Un conjunto algebraico afín $V|\mathbb{F}_q$ es una *variedad afín* si $I(V|\mathbb{F}_q)$ es un ideal primo de $\mathbb{F}_q[X]$.

Para una variedad afín $V|\mathbb{F}_q$ definimos:

1. El *anillo de coordenadas* de $V|\mathbb{F}_q$ como $\mathbb{F}_q[V] = \mathbb{F}_q[X]/I(V|\mathbb{F}_q)$;
2. El *anillo de coordenadas absoluto* de $V$ como $\bar{\mathbb{F}}_q[V] = \bar{\mathbb{F}}_q[X]/I(V)$;
3. El *cuerpo de funciones* $\mathbb{F}_q(V)$ de $V|\mathbb{F}_q$ como el cuerpo de fracciones de $\mathbb{F}_q[V]$;
4. El *cuerpo de funciones absoluto* $\bar{\mathbb{F}}_q(V)$ como el cuerpo de fracciones de $\bar{\mathbb{F}}_q[V]$.

**Definición 3.7.** Sean $V|\mathbb{F}_q$ una variedad afín y $P \in V$. El *anillo local* de $V$ en $P$ es

$$\bar{\mathbb{F}}_q(V)_P = \{h \in \bar{\mathbb{F}}_q(V) : \ h \text{ está bien definida en } P\}.$$

El único ideal maximal de este anillo está dado por

$$\bar{M}_P = \{h \in \bar{\mathbb{F}}_q(V) : h(P) = 0\}.$$

*Observación* 3.8. Si $P$ es un punto $\mathbb{F}_q$-racional, entonces

$$\mathbb{F}_q(V)_P = \bar{\mathbb{F}}_q(V)_P \cap \mathbb{F}_q(V) \quad \text{y} \quad M_P = \bar{M}_P \cap \mathbb{F}_q(V).$$

**Definición 3.9.** La *dimensión* de una variedad afín $V|\mathbb{F}_q$ es el grado de trascendencia de la extensión $\bar{\mathbb{F}}_q(V)$ sobre $\bar{\mathbb{F}}_q$.

**Definición 3.10.** Sea $V|\mathbb{F}_q$ una variedad afín y sea $\{f_1, \ldots, f_m\} \subseteq \mathbb{F}_q[X]$ un conjunto que genera $I(V)$, entonces $V$ será una variedad *suave* en $P \in V$ si la matriz $m \times n$

$$\left(\frac{\partial f_i}{\partial x_j}(P)\right)_{1 \le i \le m, 1 \le j \le n}$$

tiene rango igual a $n - \dim(V)$.

En el caso proyectivo los conceptos son definidos de manera análoga.

**Definición 3.11.** $V \subseteq \bar{\mathbb{F}}_q$ es un *conjunto algebraico proyectivo* si existe un conjunto de polinomios homogéneos $S \subseteq \bar{\mathbb{F}}_q[X_0, X_1, \ldots, X_n] = \bar{\mathbb{F}}_q[X]$ tales que $V = V(S) = \{P \in \mathbb{P}^n(\bar{\mathbb{F}}_q) : f(P) = 0 \ \forall f \in S\}$.

Como antes, diremos que $V$ está definido sobre $\mathbb{F}_q$ si $S \subseteq \mathbb{F}_q[X]$.

**Definición 3.12.** Un conjunto algebraico proyectivo $V$ es *reducible* se existen dos conjuntos algebraicos proyectivos $V_1$ y $V_2$ tales que $V_i \neq V$ para $i = 1, 2$ y $V = V_1 \cup V_2$. Si $V$ es no vacío y no reducible, diremos en este caso que $V$ es *irreducible*.

También podemos asociarle los ideales

$$I(V) = \langle \{f \in \bar{\mathbb{F}}_q[X] : f \text{ es homogéneo y } f(P) = 0 \ \forall P \in V\} \rangle,$$
$$I(V|\mathbb{F}_q) = I(V) \cap \mathbb{F}_q[X].$$

*Observación* 3.13. Un conjunto algebraico proyectivo es irreducible si y solamente si $I(V)$ es un ideal primo de $\bar{\mathbb{F}}_q[X]$.

**Definición 3.14.** Un conjunto algebraico proyectivo es una *variedad proyectiva* cuando $I(V)$ es un ideal primo homogéneo de $\bar{\mathbb{F}}_q[X]$.

Para cada $i = 1, \ldots, n$ considere las aplicaciones

$$\begin{aligned} \varphi_i : \mathbb{A}^n(\bar{\mathbb{F}}_q) &\to \mathbb{P}^n(\bar{\mathbb{F}}_q) \\ (a_1 \ldots, a_n) &\mapsto (a_1 : \cdots : a_{i-1} : 1 : a_i : \cdots : a_n) \end{aligned}$$

**Definición 3.15.** Sea $V|\mathbb{F}_q$ una variedad proyectiva y sea $i$ tal que $\varphi_i^{-1}(V \cap U_i) \neq \emptyset$ donde $U_i = \varphi_i(\mathbb{A}^n(\bar{\mathbb{F}}_q))$. Entonces definimos:

1. La *dimensión* de $V$ como la dimensión de $\varphi_i^{-1}(V \cap U_i)$.
2. El *cuerpo de funciones* $\mathbb{F}_q(V)$ de $V$ como el cuerpo de funciones $\mathbb{F}_q(\varphi_i^{-1}(V \cap U_i))$.
3. El *cuerpo de funciones absoluto* $\bar{\mathbb{F}}_q(V)$ de $V$ como el cuerpo de funciones absoluto $\bar{\mathbb{F}}_q(\varphi_i^{-1}(V \cap U_i))$.

**Definición 3.16.** Sea $P = (a_0 : \cdots : a_n) \in V$ tal que $a_i \neq 0$, entonces diremos que $V$ es *suave* en $P$ si $\varphi_i^{-1}(V \cap U_i)$ es suave en $(a_0/a_i, \ldots, a_{i-1}/a_i, a_{i+1}/a_i, \ldots, a_n/a_i)$.

**Definición 3.17.** Una variedad proyectiva de dimensión uno, definida sobre $\mathbb{F}_q$ será una *curva proyectiva* sobre $\mathbb{F}_q$.

Sea $f(X,Y) \in \mathbb{F}_q[X,Y]$ un polinomio absolutamente irreducible de grado $d$, o sea, es irreducible sobre $\bar{\mathbb{F}}_q$ y sea $F(X,Y,Z)$ su homogeneización. Considere la curva proyectiva asociada $\mathcal{C} = \{(x : y : z) \in \mathbb{P}^2(\bar{\mathbb{F}}_q) : F(x,y,z) = 0\}$; entonces el género de $\mathcal{C}$ satisface

$$g \leq \frac{(d-1)(d-2)}{2},$$

donde la igualdad vale si $\mathcal{C}$ es suave.

**Proposición 3.18.** *Sea $f(X,Y) \in \mathbb{F}_q[X,Y]$ un polinomio del siguiente tipo:*

$$f(X,Y) = a_0 Y^n + a_1(X)Y^{n-1} + \cdots + a_n(X)$$

*donde $a_0 \in \mathbb{F}_q^*$ y $a_1(X), \ldots, a_n(X) \in \mathbb{F}_q[X]$. Suponga que $\deg(a_n) = m$ con $\mathrm{mcd}(n,m) = 1$ y que*

$$\frac{m}{n} > \frac{\deg(a_i)}{i} \qquad \text{para cada } i \text{ tal que } a_i(X) \neq 0.$$

*Entonces $f(X,Y)$ es absolutamente irreducible sobre $\mathbb{F}_q$.*

Dada una curva proyectiva, no singular e irreducible $\mathcal{C}$, podemos asociarle su cuerpo de funciones $\mathbb{F}_q(\mathcal{C})$ definido anteriormente.

Ahora, si comenzamos con un cuerpo de funciones $F = \mathbb{F}_q(x,y)|\mathbb{F}_q(x)$ sabemos que existe un polinomio irreducible $g(X,Y) \in \mathbb{F}_q[X,Y]$ tal que $g(x,y) = 0$. Así, podemos asociar a $F$ la curva proyectiva no singular cuyo modelo plano está dado por $g(X,Y) = 0$.

**Teorema 3.19.** *Sea $\mathcal{C}$ sobre $\mathbb{F}_q$ una curva proyectiva, no singular y sea $F$ su cuerpo de funciones, entonces existe una correspondencia biunívoca entre los puntos $P \in \mathcal{C}$ y los lugares de $F|\mathbb{F}_q$ dada por*

$$P \mapsto M_P,$$

*donde $M_P$ es el ideal maximal del anillo local de $\mathcal{C}$ en $P$.*

*3.1.1. Extensiones de cuerpos de funciones.* Antes de adentrarnos en el problema de contar puntos racionales, vamos a echarle una mirada rápida a algunas extensiones de cuerpos de funciones que serán útiles en lo que sigue.

**Definición 3.20.** Un cuerpo de funciones $F'|K'$ es una extensión *algebraica finita* del cuerpo de funciones $F|K$ si $F' \supseteq F$, $K' \supseteq K$ y $[F' : F] < \infty$.

**Definición 3.21.** Un lugar $P' \in \mathbb{P}_{F'}$ está sobre un lugar $P \in \mathbb{P}_F$ si $P \subseteq P'$ o, equivalentemente, si existe un número entero $e \geq 1$ tal que $v_{P'}(x) = e \cdot v_P(x)$ para todo $x \in F$.

En este caso denotaremos por $P'|P$ y diremos que el lugar $P'$ es una *extensión* del lugar $P$.

• El número entero $e$, que pasaremos a denotar por $e(P'|P)$, es llamado *índice de ramificación* de $P'$ sobre $P$.

• La extensión $P'|P$ es ramificada si $e > 1$.

• Para $P'|P$, el número entero $f(P'|P) := [F'_{P'} : F_P]$ es llamado de *grado relativo* de $P'$ sobre $P$.

**Proposición 3.22.** *Sea $F'|F$ una extensión algebraica finita de cuerpos de funciones, entonces:*

1. *Para todo lugar $P' \in \mathbb{P}_{F'}$, existe un único lugar $P \in \mathbb{P}_F$ tal que $P'|P$.*
2. *Para todo lugar $P \in \mathbb{P}_F$, existe por lo menos una extensión $P' \in \mathbb{P}_{F'}$. Más aún, el número de tales extensiones es finito.*
3. *Si $P \in \mathbb{P}_F$ y $P_1, \ldots, P_m$ son todas las extensiones de $P$ en $\mathbb{P}_{F'}$, entonces*

$$\sum_{i=1}^{m} e(P_i|P) f(P_i|P) = [F' : F] = n.$$

**Definición 3.23.** Sea $F'|F$ una extensión algebraica de cuerpos de funciones y sean $P \in \mathbb{P}_F$ y $P' \in \mathbb{P}_{F'}$. Entonces

1. $P$ es *ramificada* si existe $P' \in \mathbb{P}_{F'}$ tal que $P'|P$ y $e(P'|P) > 1$.
2. $P'|P$ es de *ramificación moderada* si $e(P'|P) > 1$ y $e(P'|P)$ no es divisible por la característica del cuerpo de constantes.
3. $P'|P$ es de *ramificación salvaje* si $e(P'|P) > 1$ y $e(P'|P)$ es divisible por la característica del cuerpo de constantes.
4. $P$ es de *ramificación moderada* si para todo lugar $P' \in \mathbb{P}_{F'}$ con $P'|P$, la extensión $P'|P$ es de ramificación moderada.
5. $P$ es de *ramificación salvaje* si existe un lugar $P' \in \mathbb{P}_{F'}$ con $P'|P$ tal que la extensión $P'|P$ es de ramificación salvaje.
6. $P$ *se descompone totalmente* si para todo lugar $P' \in \mathbb{P}_{F'}$ con $P'|P$, vale que $e(P'|P) = f(P'|P) = 1$.
7. $P$ es *totalmente ramificado* si existe un único lugar $P' \in \mathbb{P}_{F'}$ con $P'|P$ y $e(P'|P) = [F' : F]$.

A partir de ahora, vamos a suponer que el cuerpo de constantes $K$ es un cuerpo perfecto, el cual es el caso de un cuerpo finito por ejemplo.

Consideremos $F'|F$ una extensión separable de cuerpos de funciones. Para $P \in \mathbb{P}_F$ definimos $\mathcal{O}'_P$ el cierre integral en $F'$ del anillo de valoración $\mathcal{O}_P$. El módulo complementar $\mathcal{C}_P$ sobre $\mathcal{O}_P$ es dado por

$$\mathcal{C}_P = \{z \in F' : \mathcal{T}_{F'|F}(z \cdot \mathcal{O}'_P) \subseteq \mathcal{O}_P\},$$

donde $\mathcal{T}_{F'|F}$ es la función traza asociada a la extensión separable $F'|F$.

*Observación* 3.24. Existe un elemento $t \in F'$ que depende de $P$ tal que $\mathcal{C}_P = t \cdot \mathcal{O}'_P$. Más aún, $v_{P'}(t) \leq 0$ para todo $P'|P$.

**Definición 3.25.** Considere un lugar $P \in \mathbb{P}_F$ y sea $\mathcal{C}_P = t \cdot \mathcal{O}'_P$ el módulo complementar. Para $P'|P$ definimos el *exponente de la diferente* por $d(P'|P) = -v_{P'}(t)$.

**Definición 3.26.** La *diferente* de $F'|F$ es el divisor dado por

$$\mathrm{Diff}(F'|F) := \sum_{P\in\mathbb{P}_F}\sum_{P'|P} d(P'|P)\cdot P'.$$

**Proposición 3.27.** *(Fórmula del género de Hurwitz) Sea $F|K$ un cuerpo de funciones de género $g$ y sea $F'|F$ una extensión separable finita. Denotamos por $K'$ el cuerpo de constantes de $F'$ y por $g'$ el género de $F'|K'$. Entonces*

$$2g' - 2 = \frac{[F' : F]}{[K' : K]}(2g - 2) - \deg(\mathrm{Diff}(F'|F)).$$

**Proposición 3.28.** *(Teorema de la Diferente de Dedekind) Con las notaciones anteriores, para $P'|P$ vale que*

1. *$d(P'|P) \geq e(P'|P) - 1$.*
2. *$d(P'|P) = e(P'|P) - 1$ si y solamente si la característica de $K$ no divide $e(P'|P)$.*

En general, es difícil calcular el género de un cuerpo de funciones. Vamos a enunciar dos resultados que suelen ser útiles a la hora de calcularlo.

**Teorema 3.29.** *Suponga que $F' = F(y)$ es una extensión de grado $n$, separable del cuerpo de funciones $F$. Sea $P \in \mathbb{P}_F$ tal que el polinomio mínimo $\varphi(T)$ de $y$ sobre $F$ tiene sus coeficientes en $\mathcal{O}_P$ y sean $P_1, \ldots, P_r$ todos los lugares de $F'$ sobre $P$. Entonces*

1. *$d(P_i|P) \leq v_{P_i}(\varphi'(y))$;*
2. *$\{1, y, \ldots, y^{n-1}\}$ es una base entera de $F'|F$ en el lugar $P$ si y solamente si $d(P_i|P) = v_{P_i}(\varphi'(y))$ para todo $i = 1, \ldots, r$,*

*donde $(\varphi'(T))$ denota la derivada usual del polinomio $\varphi(T)$.*

**Proposición 3.30.** *Sean $F'|F$ una extensión separable de cuerpos de funciones, $P \in \mathbb{P}_F$ y $P' \in \mathbb{P}_{F'}$ con $P'|P$. Suponga que $P'|P$ sea totalmente ramificada y sea $t \in F'$ un elemento $P'$-primo (o sea, $P' = t\mathcal{O}_{P'}$). Considere $\varphi(T) \in F[T]$ el polinomio mínimo de $t$ sobre $F$. Entonces $d(P'|P) = v_{P'}(\varphi'(t))$ y $\{1, t, \ldots, t^{n-1}\}$ es una base entera de $F'|F$ en el lugar $P$.*

*3.1.2. Extensiones de Kummer y Artin-Schreier.*

**Definición 3.31.** $F'|F$ es una *extensión de Galois* de cuerpos de funciones si $F'|F$ es una extensión de Galois de cuerpos.

Tenemos dos clases importantes de extensiones galoisianas: extensiones de Kummer y de Artin-Schreier. Estas extensiones tienen dos ventajas importantes: podemos explicitar las ecuaciones que las definen y el género puede ser calculado fácilmente.

**Definición 3.32.** Sea $F|K$ un cuerpo de funciones en donde $K$ contiene una raíz primitiva de orden $n \geq 1$ de la unidad y tal que $\mathrm{mcd}(n, \mathrm{char}\, K) = 1$. Suponga que existe un elemento $u \in F$ tal que $u \neq w^d$ para todo $w \in F$ y para todo $d > 1$ con $d|n$. Entonces el cuerpo de funciones dado por

$$F' := F(y), \quad \text{con } y^n = u,$$

es una *extensión de Kummer* de $F$.

*Observación* 3.33. Una extensión de Kummer $F'|F$ satisface las siguientes propiedades:

1. $F'|F$ es una extensión cíclica y su grupo de Galois está generado por $\sigma(y) = \zeta y$ donde $\zeta \in K$ es una raíz primitiva de la unidad de orden $n$.
2. Sea $P \in \mathbb{P}_F$ y sea $P' \in \mathbb{P}_{F'}$ una extensión de $P$, entonces
$$e(P'|P) = \frac{n}{r_P} \quad \text{y} \quad d(P'|P) = e(P'|P) - 1,$$
donde $r_P := \text{mcd}(n, v_P(u)) > 0$.
3. Si $K'$ es el cuerpo de constantes de $F'$ y $g'$ es el género de $F'$, entonces
$$g' = 1 + \frac{n}{[K' : K]}\left(g - 1 + \frac{1}{2}\sum_{P \in \mathbb{P}_F}\Big(1 - \frac{r_P}{n}\Big)\deg(P)\right).$$

*Observación* 3.34. Sea $F|K$ un cuerpo de funciones de género $g$ y sea $F'$ dado por $F' := F(y)$ con $y^n = u \in F$, donde $n \not\equiv 0 \mod \text{char(K)}$. Suponga que $K$ contiene una raíz primitiva de orden $n$ de la unidad y que existe un lugar $Q \in \mathbb{P}_F$ tal que $\text{mcd}(v_Q(u), n) = 1$. Entonces $K$ es el cuerpo de constantes de $F'$ y la extensión $F'|F$ es cíclica de grado $n$. El género $g'$ está dado por
$$g' = 1 + n(g-1) + \frac{1}{2}\sum_{P \in \mathbb{P}_F}\Big(n - r_P\Big)\deg(P).$$

**Ejemplo 3.35.** Considere un cuerpo $K$ de característica diferente de 2. Sea
$$F := K(x, y) \quad \text{con } y^2 = f(x) = \prod_{i=1}^{s} p_i(x) \in K[x],$$
donde $p_i(x)$ son polinomios mónicos, irreducibles distintos y $s \geq 1$.

Denotando por $m$ el grado del polinomio $f(x)$ tenemos que $K$ es el cuerpo de constantes de $F$ y que el género está dado por
$$g = \begin{cases} (m-1)/2 \text{ si } m \equiv 1 \mod 2, \\ (m-2)/2 \text{ si } m \equiv 0 \mod 2. \end{cases}$$

De hecho, considere $F = F_0(y)$ donde $F_0 := K(x)$ es el cuerpo de funciones racionales. Denotando por $P_i \in \mathbb{P}_{K(x)}$ el cero de $p_i(x)$ y $P_\infty$ el polo de $x$ en $K(x)$ , entonces $v_{P_i}(f(x)) = 1$ y $v_{P_\infty}(f(x)) = -m$. Así $F|F_0$ es una extensión cíclica de grado 2. Vale también que

1. $r_{P_j} = 1$ para $j = 1, \ldots, s$;
2. $r_{P_\infty} = 1$ si $m \equiv 1 \mod 2$;
3. $r_{P_\infty} = 2$ si $m \equiv 0 \mod 2$.

**Definición 3.36.** Sea $F|K$ un cuerpo de funciones de característica $p > 0$. Suponga que $u \in F$ es tal que $u \neq w^p - w$ para todo $w \in F$. El cuerpo de funciones $F'$ dado por
$$F' := F(y), \quad \text{con } y^p - y = u$$
es una *extensión de Artin-Schreier* de $F$.

*Observación* 3.37. Una extensión de Artin-Schreier satisface las siguientes propiedades:

1. $F'|F$ es una extensión cíclica de grado $p$ y los automorfismos de $F'|F$ están dados por $\sigma(y) = y + \nu$ donde $\nu = 0, 1, \ldots, p-1$.

2. Para $P \in \mathbb{P}_F$ y $P'$ una extensión de $P$, el número a continuación está bien definido

$$m_P := \begin{cases} m & \text{si existe } z \in F \text{ tal que } v_P(u-(z^p-z)) = -m \text{ y } m \not\equiv 0 \mod p, \\ -1 & \text{si } v_P(u-(z^p-z)) \geq 0 \text{ para algún } z \in F. \end{cases}$$

3. $P$ es no ramificado si y solamente si $m_P = -1$.
4. $P$ es totalmente ramificado si y solamente si $m_P \geq 0$. En este caso tenemos que

$$d(P'|P) = (p-1)(m_P+1), \quad \text{donde } P' \text{ es el único lugar sobre } P.$$

5. Si por lo menos un lugar $Q \in \mathbb{P}_F$ es totalmente ramificado, entonces $K$ es el cuerpo de constantes de $F'$ y el género $g'$ de $F'$ está dado por

$$g' = p \cdot g + \frac{p-1}{2}\left( -2 + \sum_{P \in \mathbb{P}_F} \Big(m_P + 1\Big) \cdot \deg(P)\right).$$

**Definición 3.38.** Un polinomio $a(x) \in K[x]$ del tipo $a(x) = a_n x^{p^n} + a_{n-1}x^{p^{n-1}} + \cdots + a_1 x^p + a_0 x$ donde $p$ es la característica del cuerpo $K$, es llamado polinomio aditivo sobre $K$.

**Proposición 3.39.** *Sean $F|K$ un cuerpo de funciones de característica $p \geq 0$, $a(x)$ un polinomio aditivo de grado $p^n$ que tiene todas sus raíces en $K$ y $u \in F$. Suponga que para todo lugar $P \in \mathbb{P}_F$ existe un elemento $z \in F$ ($z$ depende de $P$) tal que*

$$\begin{cases} v_P(u-a(z)) \geq 0 & o \\ v_P(u-a(z)) = -m & \text{con } m > 0 \text{ y } m \not\equiv 0 \mod p. \end{cases}$$

*Definimos $m_P = -1$ en el primer caso y $m_p = -m$ en el segundo caso.*

*Consideremos ahora la extensión*

$$F' := F(y) \quad \text{con } a(y) = u.$$

*Si existe un lugar $Q \in \mathbb{P}_F$ con $m_Q > 0$, entonces*

1. *$F'|F$ es una extensión de Galois de grado $p^n$, cuyo grupo de Galois es isomorfo a $(\mathbb{Z}/p\mathbb{Z})^n$.*
2. *$K$ es algebraicamente cerrado en $F'$.*
3. *Si $P \in \mathbb{P}_F$ es tal que $m_P = -1$, entonces $P$ es no ramificado en $F'|F$.*
4. *Si $P \in \mathbb{P}_F$ es tal que $m_P > 0$, entonces $P$ es totalmente ramificado en $F'|F$. En este caso*

$$d(P'|P) = (p^n-1)(m_P+1), \quad \text{donde } P' \text{ es el único lugar sobre } P.$$

5. *El género $g'$ de $F'$ está dado por*

$$g' = p^n \cdot g + \frac{p^n-1}{2}\left( -2 + \sum_{P \in \mathbb{P}_F} \Big(m_P + 1\Big) \cdot \deg(P)\right).$$

**3.2. ¿Cuántos puntos podemos esperar?** Nuestro objetivo es establecer un límite superior para el número de puntos racionales de una curva definida sobre un cuerpo finito $\mathbb{F}_q$, o lo que es equivalente, para el número de lugares racionales de un cuerpo de funciones $F|\mathbb{F}_q$. Vamos a hacer una serie de afirmaciones, la mayor parte sin demostraciones. Las pruebas se encuentran en [47] o [42]

**Afirmación 3.40.** *Para todo $n \geq 0$ existe únicamente un número finito de divisores positivos de grado $n$*

Consideremos los conjuntos

$$\mathcal{D}_F^0 := \{A \in \mathcal{D}_F; \deg(A) = 0\},$$
$$\mathcal{C}_F^0 := \{[A] \in \mathcal{C}_F; \deg([A]) = 0\},$$

donde $\mathcal{P}_F$ es el grupo de divisores principales (vea pág. 16) y $\mathcal{C}_F := \mathcal{D}_F/\mathcal{P}_F$ es llamado de grupo de clases de divisores de $F|\mathbb{F}_q$.

**Afirmación 3.41.** *$\mathcal{C}_F^0$ es un grupo finito de orden $h$.*

Definimos $\partial > 0$ por

$$\partial = \text{mín}\{\deg(A) : A \in \mathcal{D}_F \text{ y } \deg(A) > 0\}.$$

*Observación* 3.42. No es verdad en general que para un cuerpo de constantes arbitrario exista un divisor positivo de grado $\partial$.

Vamos ahora a estimar los números $A_n := \#\{A \in \mathcal{D}_F : A \geq 0 \text{ y } \deg(A) = n\}$.

Por ejemplo $A_0 = 1$ y $A_1$ es precisamente el número de lugares de grado uno que queremos limitar.

**Proposición 3.43.** *Con las notaciones anteriores vale que:*

1. *$A_n = 0$ si $\partial \nmid n$.*
2. *Para una clase de divisores fija $[C] \in \mathcal{C}_F$ tenemos*
$$\#\{A \in [C] : A \geq 0\} = \frac{h}{q-1}\Big(q^{\dim[C]} - 1\Big).$$
3. *Para cualquier entero $n > 2g - 2$ con $\partial | n$, tenemos*
$$A_n = \frac{h}{q-1}\Big(q^{n+1-g} - 1\Big).$$

**Definición 3.44.** La serie de potencias definida por

$$Z_F(t) = \sum_{n=0}^{\infty} A_n t^n \in \mathbb{C}[[t]]$$

es llamada la *Función Zeta* de $F|\mathbb{F}_q$.

**Proposición 3.45.** *La serie de potencias $Z(t) = Z_F(t)$ es convergente para $|t| < q^{-1}$ y en este caso tenemos que*

1. *Si $F|\mathbb{F}_q$ tiene género cero, entonces*
$$Z(t) = \frac{1}{q-1}\Big(\frac{q}{1-(qt)^{\partial}} - \frac{1}{1-t^{\partial}}\Big).$$
2. *Si $g \geq 1$, entonces $Z(t)$ puede ser escrita como $Z(t) = F(t) + G(t)$ donde*
$$F(t) = \frac{1}{q-1} \sum_{0 \leq \deg[C] \leq 2g-2} q^{\dim[C]} \cdot t^{\deg[C]},$$
$$G(t) = \frac{h}{q-1}\Big(q^{1-g}(qt)^{2g-2+\partial}\frac{1}{1-(qt)^{\partial}} - \frac{1}{1-t^{\partial}}\Big).$$

**Corolario 3.46.** *La función $Z(t)$ puede ser extendida a una función racional definida sobre $\mathbb{C}$ con polo simple en $t = 1$.*

La función Zeta también puede ser escrita para $|t| < q^{-1}$ como un producto (conocido como producto de Euler) de la forma

$$Z(t) = \prod_{P \in \mathbb{P}_F} \left(1 - t^{\deg P}\right)^{-1}.$$

En particular, $Z(t) \neq 0$ para $|t| < q^{-1}$ y $Z(t)$ satisface la ecuación funcional

$$Z(t) = q^{g-1}\, t^{2g-1}\, Z\Big(\frac{1}{qt}\Big).$$

Vamos a estudiar ahora el comportamiento de la función Zeta cuando hacemos extensiones del cuerpo $F$ por constantes.

Consideremos la siguiente extensión por constantes $F_r := F \cdot \mathbb{F}_{q^r}$ del cuerpo $F|\mathbb{F}_q$ y denotemos por $Z_r(t)$ la función Zeta asociada. Vale que

$$Z_r(t) = \prod_{\zeta^r = 1} Z(\zeta t) \quad \forall t \in \mathbb{C}, \qquad \text{y } \zeta \in \mathbb{C} \text{ tal que } \zeta^r = 1.$$

*Observación* 3.47. Usando la igualdad anterior, F.K. Schmidt mostró que $\partial = 1$.

Podemos asociar a $Z(t)$ un polinomio que va a jugar un papel importante para determinar un límite superior para el número de lugares racionales.

**Definición 3.48.** El polinomio $L(t) := (1-t)(1-qt)Z(t)$ es llamado el *L-polinomio* de $F|\mathbb{F}_q$ y tiene grado como máximo $2g$.

Como $L(t) = (1-t)(1-qt)\sum_{n=0}^{\infty} A_n t^n$, este polinomio guarda toda la información sobre los $A_n'^s$ para todo $n \geq 0$. Este hecho nos motiva a estudiarlo un poco más. Algunas de sus propiedades están resumidas en el siguiente teorema.

**Teorema 3.49.** *Sea $L(t)$ como antes. Entonces:*

1. *$L(t) \in \mathbb{Z}[t]$ y tiene grado $2g$;*
2. *$L(t) = q^g\, t^{2g} L(1/qt)$;*
3. *$L(1) = h$;*
4. *si $L(t) = \sum_{i=0}^{2g} a_i t^i$, entonces:*
   a) *$a_0 = 1$ y $a_{2g} = q^{2g}$;*
   b) *$a_{2g-i} = q^{g-i} a_i$ para $0 \leq i \leq g$;*
   c) *$a_1 = N - (q-1)$ donde $N$ es el número de lugares de grado uno;*
5. *$L(t)$ se factoriza en $\mathbb{C}[t]$ como:*

$$L(t) = \prod_{i=1}^{2g} (1 - \alpha_i\, t)$$

*y $\alpha_1, \ldots, \alpha_{2g}$ pueden ser ordenados tal que $\alpha_i\, \alpha_{g+i} = q$ para $i = 1, \ldots, g$.*

**Definición 3.50.** Los números $\alpha_1, \ldots, \alpha_{2g}$ son llamados las *raíces recíprocas* de $L(t)$.

**Corolario 3.51.** *Para cada $r \geq 1$ tenemos*

$$N_r := N(F_r) = q^r + 1 - \sum_{i=1}^{2g} \alpha_i^r,$$

*donde $\alpha_1, \ldots, \alpha_{2g}$ son las raíces recíprocas de $L(t)$.*

Así, para estimar $N$ basta estimar $|\alpha_i|$.

**Teorema 3.52.** *(Hasse-Weil) Los números $\alpha_i$ satisfacen*

$$|\alpha_i| = \sqrt{q}, \quad \forall i = 1, \ldots, 2g.$$

**Teorema 3.53.** *(Cota de Hasse-Weil) El número $N$ de lugares de grado uno puede ser estimado por*

$$|N - (q+1)| \leq 2g\sqrt{q}.$$

Esta desigualdad es equivalente a la validez de la hipótesis de Riemann para la función Zeta asociada al cuerpo de funciones $F|\mathbb{F}_q$. De hecho, si definimos la *norma absoluta* de un divisor $A$ como $\mathcal{N}(A) = q^{\deg A}$ y consideramos la función dada por

$$\zeta_F(s) := Z_F(q^{-s}),$$

entonces esta función puede ser escrita como

$$\zeta_F(s) = \sum_{n=0}^{\infty} A_n \, q^{-sn} = \sum_{A \in \mathcal{D}_F, A \geq 0} \mathcal{N}(A)^{-s},$$

que es análoga a la función Zeta de Riemann clásica $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$.

La hipótesis de Riemann para la función clásica afirma que, además de los ceros triviales, todos los ceros se encuentran localizados en la recta $Re(s) = 1/2$. En el caso de cuerpo de funciones, el teorema de Hasse-Weil muestra que

$$\zeta_F(s) = 0 \Rightarrow Z_F(q^{-s}) = 0 \Rightarrow |q^{-s}| = q^{1/2},$$

ya que $|q^{-s}| = q^{Re(s)}$, lo que implica trivialmente que $Re(s) = 1/2$.

Históricamente, los primeros estudios sobre ecuaciones sobre cuerpos finitos se centraban en las congruencias del tipo

$$Y^2 \equiv f(X) \mod p \quad (*)$$

donde $p$ es un número primo y $f(X)$ un polinomio (o una función racional) definido sobre $\mathbb{Z}$.

Artin introdujo la función Zeta para extensiones cuadráticas del cuerpo racional $\mathbb{F}_p(x)$ obtenido por la adjunción de las raíces de la congruencia anterior, basado en la función Zeta de Dedekind para extensiones cuadráticas de $\mathbb{Q}$.

Asumiendo como verdadera la hipótesis de Riemann para la función clásica, Artin llegó a conjeturar una cota superior para el número de soluciones de la congruencia (*). Esta conjetura fue probada por Hasse para polinomios de grado 3 o 4 definidos sobre un cuerpo finito arbitrario y después generalizada por Weil como vimos en el teorema anterior.

*Observación* 3.54. Ihara mostró en [32] que si $F|\mathbb{F}_q$ es un cuerpo de funciones de género $g$, entonces el número de lugares racionales $N$ satisface

$$N \leq q + 1 + \left\lfloor \frac{\sqrt{(8q+1)g^2 + 4(q^2-q)g} - g}{2} \right\rfloor,$$

donde $\lfloor x \rfloor$ denota la parte entera de $x$.

Note que si $g > (q - \sqrt{q})/2$, esta cota es mejor que la de Hasse-Weil.

**Ejemplo 3.55.** (Hermitiana) Considere la curva $\mathcal{H}$ definida sobre $\mathbb{F}_{q^2}$ asociada al polinomio $f(X,Y) = Y^q + Y - X^{q+1}$. Esta curva es no singular y tiene género $g = q(q-1)/2$.

Vamos a calcular el número de puntos racionales, o sea

$$\#\{(x:y:z)\,|\,x,y,z\in\mathbb{F}_{q^2}\text{ tales que }F(x,y,z)=0\},$$

donde $F(X,Y,Z)=ZY^q+Z^qY-X^{q+1}$ es la homogeneización de $f(X,Y)$.

Vamos a comenzar por la parte afín, o sea, cuando $z=1$, en este caso tenemos que calcular las soluciones de $f(x,y)$ en $\mathbb{F}_{q^2}\times\mathbb{F}_{q^2}$. Podemos pensar en calcular el número $N_a$ de elementos del conjunto

$$N_a=\#\{(x,y)\in\mathbb{F}_{q^2}\times\mathbb{F}_{q^2}:Tr_{\mathbb{F}_{q^2}|\mathbb{F}_q}(y)=N_{\mathbb{F}_{q^2}|\mathbb{F}_q}(x)\}$$

donde $Tr_{\mathbb{F}_{q^2}|\mathbb{F}_q}$ y $N_{\mathbb{F}_{q^2}|\mathbb{F}_q}$ son las funciones Traza y Norma de la extensión galoisiana $\mathbb{F}_{q^2}|\mathbb{F}_q$.

Como $\mathcal{T}$ es suryectiva, tenemos que $\#\mathcal{T}^{-1}(y)=q$ para todo $y\in\mathbb{F}_{q^2}$, obtenemos así

$$\begin{aligned}N_a &= \sum_{x\in\mathbb{F}_{q^2}}\#\{y\in\mathbb{F}_{q^2}:Tr_{\mathbb{F}_{q^2}|\mathbb{F}_q}(y)=x^{q+1}\}\\ &= \sum_{x\in\mathbb{F}_{q^2}}\#Tr^{-1}_{\mathbb{F}_{q^2}|\mathbb{F}_q}(x^{q+1})\\ &= \sum_{x\in\mathbb{F}_{q^2}}q\\ &= q^3.\end{aligned}$$

Falta calcular el número de puntos racionales en el infinito (en este caso $z=0$). Así se $(x:y:0)$ es una solución de $F(X,Y,Z)=0$, entonces $x=0$ lo que implica que $y=1$. Tenemos, por lo tanto, un único punto racional en el infinito. Finalmente tenemos que la curva $\mathcal{H}$ tiene $q^3+1$ puntos racionales alcanzando la cota de Hasse-Weil.

**Definición 3.56.** Un cuerpo de funciones $F|\mathbb{F}_q$ es llamada *maximal* sobre $\mathbb{F}_q$ si el número de lugares racionales de grado uno alcanza la cota de Hasse-Weil.

*Observación* 3.57. Si $F|\mathbb{F}_q$ es maximal, como $N(F|\mathbb{F}_q)=q+1+2g\sqrt{q}$ es un número entero, tenemos necesariamente que $q$ debe ser un cuadrado.

**3.3. Mejoras de la cota de Hasse-Weil.** Como vimos en la subsección anterior, la cota de Hasse-Weil sólo es alcanzada si la cardinalidad del cuerpo finito es un cuadrado, así una mejora trivial para la cota es

$$|N-(q+1)|\leq\lfloor 2g\sqrt{q}\rfloor.$$

Serre dio en [44] la siguiente mejora.

**Teorema 3.58.** *(Cota de Serre) Sea $F|\mathbb{F}_q$ un cuerpo de funciones de género $g$, entonces el número $N$ de lugares de grado uno está limitado por*

$$|N-(q+1)|\leq g\lfloor 2\sqrt{q}\rfloor.$$

La idea de la prueba es bastante simple, así que vamos hacer un esbozo.

*Demostración.* Sabemos que el $L$-polinomio puede ser escrito como

$$L(t)=\prod_{i=1}^{2g}(1-\alpha_i\,t)\in\mathbb{Z}[t]$$

y que los $\alpha_i$ pueden ser ordenados de tal forma que $\bar{\alpha}_i = \alpha_{g+i}$ para $i = 1, \ldots, g$, donde $\bar{\alpha}_i$ denota el complejo conjugado de $\alpha_i$. De hecho, como $|\alpha_i| = \sqrt{q}$ y $\alpha_i \alpha_{g+i} = q$, debemos tener que $\bar{\alpha}_i = \alpha_{g+i}$ para $i = 1, \ldots, g$.

Vamos a comenzar mostrando que

$$N - (q+1) \leq g\lfloor 2\sqrt{q} \rfloor.$$

Para cada $i = 1, \ldots, g$, definimos $\beta_i := \alpha_i + \bar{\alpha}_i + \lfloor 2\sqrt{q} \rfloor$. Los números así definidos son números reales positivos y son enteros algebraicos.

Sea $E = \mathbb{Q}(\alpha_1, \ldots, \alpha_{2g})$ una extensión $E|\mathbb{Q}$ galoisiana. Todo $\sigma \in Gal(E|\mathbb{Q})$ induce una permutación del conjunto $\{\beta_1, \ldots, \beta_g\}$, lo que implica que

$$\beta = \prod_{i=1}^{g} \beta_i$$

es dejado fijo por todos los automorfismos y como $\beta$ es un entero algebraico debemos tener que $\beta \in \mathbb{Z}$ y $\beta \geq 1$.

Tenemos entonces que

$$\begin{aligned} g \ \leq \ \sum_{i=1}^{g} \beta_i &= \sum_{i=1}^{g} (\alpha_i + \bar{\alpha}_i) + g\lfloor 2\sqrt{q} \rfloor + g \\ &= \Big( \sum_{i=1}^{2g} \alpha_i \Big) + g\lfloor 2\sqrt{q} \rfloor + g. \end{aligned}$$

Como $N = q + 1 - \sum_{i=1}^{2g} \alpha_i$ y usando la desigualdad entre la media aritmética y la geométrica, tenemos que

$$\frac{1}{g} \sum_{i=1}^{g} \beta_i \geq \Big( \prod_{i=1}^{g} \beta_i \Big)^{1/g},$$

lo que concluye la prueba de la desigualdad.

La prueba de que $N - (q+1) \geq -(g\lfloor 2\sqrt{q} \rfloor)$ es análoga y será dejada como ejercicio para el lector. □

**Ejemplo 3.59.** (La cuártica de Klein) Considere la curva $\mathcal{C}$ definida sobre $\mathbb{F}_8$ por el polinomio

$$f(X, Y) = Y^3 + X^3 Y + X$$

y sea $F(X, Y, Z) = ZY^3 + X^3 Y + Z^3 X \in \mathbb{F}_8[X, Y, Z]$ la homogeneización de $f$. La curva $\mathcal{C}$ es no singular y tiene género $3 = (d-1)(d-2)/2$, donde $d = 4$ es el grado de $f$.

Vamos a calcular los puntos racionales afines, o sea, cuando $z = 1$ y para los cuales $x \neq 0$ y $y \neq 0$.

Multiplicando $f(X, Y)$ por $X^6$ obtenemos la ecuación

$$W^3 + X^7 W + X^7,$$

donde $W = X^2 Y$.

Olvidándonos del origen, tenemos que

$$\begin{aligned} \# N_a - 1 &= \# \left\{ (x, w) \in \mathbb{F}_8^* \times \mathbb{F}_8 : x^7 = \frac{w^3}{w+1} \right\} \\ &= \# \Big\{ (x, w) \in \mathbb{F}_8^* \times \mathbb{F}_8 ; w^3 + w + 1 = 0 \Big\}. \end{aligned}$$

El polinomio $W^3 + W + 1$ es irreducible sobre $\mathbb{F}_2[W]$, así tiene todas sus raíces en $\mathbb{F}_8$. Por lo tanto $N_a - 1 = 7 \cdot 3 = 21$ y la cuártica de Klein tiene $N = 21 + 3 = 24$ puntos racionales alcanzando así la cota de Serre. Los 3 puntos que faltan son puntos en el infinito y la caracterización de los mismos es dejada para el lector.

Es lógico esperar que si tenemos un cuerpo de funciones $F$ definido sobre $\mathbb{F}_q$ y extendemos el cuerpo de constantes para algún $\mathbb{F}_{q^r}$, el número de lugares de grado uno de $F|\mathbb{F}_{q^r}$ sólo puede aumentar. Usando esta idea, Serre mostró el siguiente teorema.

**Teorema 3.60.** *Sea $\Psi(t) = \sum_{r=1}^{m} c_r t^r \in \mathbb{R}[t] \setminus \{0\}$ un polinomio con coeficientes no negativos. Considere la función racional dada por $f(t) = 1 + \Psi(t) + \Psi(t^{-1})$. Suponga que $f(\gamma) \geq 0$ para todo $\gamma \in \mathcal{C}$ con $|\gamma| = 1$. Entonces para $F$ un cuerpo de funciones de género $g$ definido sobre $\mathbb{F}_q$ vale que*

$$N_1 := N \leq \frac{g}{\Psi\left(q^{-\frac{1}{2}}\right)} + \frac{\Psi\left(q^{\frac{1}{2}}\right)}{\Psi\left(q^{-\frac{1}{2}}\right)} + 1.$$

*Demostración.* Claramente tenemos que $\Psi\left(q^{-\frac{1}{2}}\right) \geq 0$ y $\Psi\left(q^{\frac{1}{2}}\right) \geq 0$.

Reordenando $\alpha_1, \ldots, \alpha_{2g}$ tal que $\bar{\alpha}_i = \alpha_{g+i}$ para todo $i = 1, \ldots, g$ tenemos que

$$N_r = q^r + 1 - \sum_{i=1}^{g} (\alpha_i^r + \bar{\alpha}_i^r),$$

donde $N_r$ denota el número de lugares de grado uno de $F|\mathbb{F}_{q^r}$. Multiplicando la igualdad anterior por $q^{-r/2}$ obtenemos

$$q^{-r/2} N_r = q^{r/2} + q^{-r/2} - \sum_{i=1}^{g} [(\alpha_i q^{-1/2})^r + (\bar{\alpha}_i q^{-1/2})^r].$$

Por Hasse-Weil, sabemos que los números $\gamma_i := \alpha_i q^{-1/2}$ son complejos de norma uno para $i = 1, \ldots, g$. Así,

$$q^{-r/2} N_r = q^{r/2} + q^{-r/2} - \sum_{i=1}^{g} (\gamma_i{}^r + \gamma_i{}^{-r}).$$

Multiplicando esta igualdad por $c_r$ para cada $r = 1, \ldots, m$ y sumando, tenemos

$$0 = \Psi(q^{1/2}) + \Psi(q^{-1/2}) + g - \sum_{i=1}^{g} f(\gamma_i) - \sum_{r=1}^{m} N_r c_r q^{-r/2}.$$

Por lo tanto

$$N_1 \Psi(q^{1/2}) = \Psi(q^{1/2}) + \Psi(q^{-1/2}) + g - R,$$

donde $R = \sum_{i=1}^{g} f(\gamma_i) + \sum_{r=1}^{m} (N_r - N_1) c_r q^{-r/2}$.

Como $f(\gamma_i) \geq 0$ para todo $i$ y $c_r \geq 0$ para todo $r$, tenemos que $R \geq 0$ y el teorema sigue a partir de ahí. □

**Ejemplo 3.61.** Considere la curva no singular definida sobre $\mathbb{F}_q$ cuyo modelo plano está dado por

$$f(X, Y) = Y^q - Y - X^{q_0}(X^q - X)$$

donde $q = 2^{2e+1}$ y $q_0 = 2^{2e}$. Esta curva es llamada la Curva de Suzuki y tiene género $g = q_0(q - 1)$.

Tomando $\Psi(t) = \frac{1}{\sqrt{2}}t + \frac{1}{4}t^2$ y aplicando el teorema anterior, vemos que el número de puntos racionales es $1 + q^2$, y por lo tanto esta curva alcanza la cota de Serre.

Otra herramienta que podemos usar para mejorar la cota de Hasse-Weil es conocido como Método de Stöhr-Voloch [49]. Este método consiste en construir una función auxiliar que tenga ceros de orden alto en los puntos racionales de la curva y es una herramienta fundamental en el estudio de curvas maximales.

Vamos a ilustrar su funcionamiento en el caso de curvas no singulares planas. En este caso alto va a significar mayor o igual a dos.

Sea $\mathcal{C}$ la curva afín dada por el polinomio

$$f(X,Y) = \sum_{i,j} a_{ij} X^i Y^j \in \mathbb{F}_q[X,Y].$$

Para un punto de la curva $(x,y) \in \bar{\mathbb{F}}_q \times \bar{\mathbb{F}}_q$ podemos definir la recta tangente como

$$(X-x)f_X(x,y) + (Y-y)f_Y(x,y) = 0,$$

donde $f_X, f_Y$ son las derivadas parciales usuales.

Vamos ahora a definir un polinomio auxiliar $h(X,Y)$ dado por

$$h(X,Y) = (X - X^q)f_X(X,Y) + (Y^q - Y)f_Y(X,Y),$$

el grado del polinomio satisface $\deg(h) \leq \deg(f) + q - 1$.

**Proposición 3.62.** *Si* $h(X,Y) \equiv 0 \mod f$, *entonces*

$$f_{XX}f_Y^2 - 2f_{XY}f_Xf_Y + f_{YY}f_X^2 \equiv 0 \mod f.$$

*Demostración.* Si $h(X,Y) \equiv 0 \mod f$, implica que

$$(X - X^q)f_X(X,Y) \equiv -(Y^q - Y)F_Y(X,Y) \mod f.$$

Multiplicando $f_{XX}f_Y^2 - 2f_{XY}f_Xf_Y + f_{YY}f_X^2$ por $(X - X^q)^2$ tenemos que

$$\begin{aligned}&(X - X^q)^2(f_{XX}f_Y^2 - 2f_{XY}f_Xf_Y + f_{YY}f_X^2)\\ &\equiv \left[(X - X^q)^2(f_{XX} + 2(X - X^q)(Y - Y^q)f_{XY} + (Y - Y^q)f_{YY}\right]f_Y^2 \mod f.\end{aligned}$$

Sabemos también que existe un polinomio $g = g(X,Y)$ tal que $h = f \cdot g$, o sea

$$h(X,Y) = (X - X^q)f_X(X,Y) + (Y^q - Y)f_Y(X,Y) = f(X,Y)g(X,Y).$$

Derivando con respecto a $X$ la igualdad anterior y multiplicando el resultado por $X - X^q$, obtenemos que

$$(X - X^q)^2 f_{XX} + (X - X^q)(Y - y^q)f_{YX} \equiv (X - X^q)f_X(g-1) \mod f.$$

Haciendo lo mismo con $Y$ tenemos

$$(Y - Y^q)^2 f_{YY} + (X - X^q)(Y - y^q)f_{XY} \equiv (Y - Y^q)f_Y(g-1) \mod f.$$

Sumando estas dos congruencias se obtiene que

$$(X - X^q)^2 f_{XX} + 2(X - X^q)(Y - y^q)f_{XY} + (Y - Y^q)^2 f_{YY} \equiv h(g-1)0 \mod f.$$

El resultado sigue a partir de aquí ya que $(X - X^q)^2 \not\equiv 0 \mod f$. □

*Observación* 3.63.

1. Si la característica del cuerpo es dos, entonces $f_{XX}f_Y^2 - 2f_{XY}f_Xf_Y + f_{YY}f_X^2$ es idénticamente cero.
2. Si $h \equiv 0 \mod f$, entonces para cualquier punto $(x,y)$ que pertenezca a la parte afín de $\mathcal{C}$ tenemos que $(x^q, y^q)$ pertenece a la recta tangente en $(x,y)$.

**Teorema 3.64.** *Sea $\mathbb{F}_q$ un cuerpo finito de característica impar y sea $f(X,Y) \in \mathbb{F}_q[X,Y]$ un polinomio absolutamente irreducible de grado $d$. Suponga que $f$ no divide al polinomio $f_{XX}f_Y^2 - 2f_{XY}f_Xf_Y + f_{YY}f_X^2$, entonces*

$$\#\mathcal{C}(\mathbb{F}_q) \leq \frac{d(d+q-1)}{2},$$

*donde $\mathcal{C}$ es la curva proyectiva asociada a $f(X,Y)$.*

*Demostración.* Para simplificar, vamos a hacer la prueba para el caso afín.

Como $f$ no divide a $f_{XX}f_Y^2 - 2f_{XY}f_Xf_Y + f_{YY}f_X^2$, eso implica que tampoco divide a $h(X,Y)$ y por lo tanto no tiene ningún factor de grado positivo en común.

Si $P = (x,y) \in \mathbb{F}_q \times \mathbb{F}_q$ pertenece a $\mathcal{C}$, entonces $(x,y)$ es un cero de $h(X,Y)$ por definición. Más aún, como $\mathcal{C}$ y la curva asociada a $h(X,Y)$ comparten la misma recta tangente por $P$, concluimos que el índice de intersección $I(P;f,h)$ entre las dos curvas en $P$ es de por lo menos 2. Así,

$$\#\mathcal{C}(\mathbb{F}_q) = \#\{(x,y) \in \mathbb{F}_q \times \mathbb{F}_q : f(x,y) = 0\} \leq \frac{1}{2}\sum_P \mathcal{I}(P;f,h),$$

donde $P$ se mueve sobre todos los puntos de $\mathbb{F}_q \times \mathbb{F}_q$. Por el Teorema de Bezout, tenemos que

$$\frac{1}{2}\sum_P \mathcal{I}(P;f,h) \leq \frac{1}{2}\deg(f)\cdot\deg(h) \leq \frac{d(d+q-1)}{2}.$$

□

**Ejemplo 3.65.** Considere la curva de Fermat sobre $\mathbb{F}_{13}$ cuyo modelo plano está dado por

$$f(X,Y) = w^2X^4 + Y^4 + w,$$

donde $w \in \mathbb{F}_{13} \setminus \{1\}$ es una raíz tercera de la unidad. La curva asociada a este polinomio tiene género $g = 3$. Esta curva no tiene ningún punto racional en el infinito y tampoco tiene puntos racionales afines donde una de las coordenadas sea cero.

Vamos a calcular los puntos racionales afines. Sea $H = \{1, w, w^2\}$ el único subgrupo de $\mathbb{F}_{13}^*$ de orden 3. Considere el homomorfismo suryectivo

$$\begin{aligned} \phi : \mathbb{F}_{13}^* &\to H \\ x &\mapsto x^4. \end{aligned}$$

El núcleo es el único subgrupo de $\mathbb{F}_{13}^*$ de orden 4, así $\phi^{-1}(z) = 4$ para todo $z \in H$ $(*)$.

Sea $(x,y) \in \mathcal{C}_a(\mathbb{F}_{13})$ un punto racional afín. Como $x \in \mathbb{F}_{13}^*$, entonces $x^4 \in H$.

1. Si $x^4 = w^2$, entonces

   $$f(x,Y) = w^4 + Y^4 + w = Y^4 + 2w.$$

   Como $y \in H$ y $1 + w + w^2 = 0$, tenemos que no existen puntos racionales afines con primera coordenada satisfaciendo $x^4 = w^2$.
2. Si $x^4 = 1$ entonces $y^4 = 1$.
3. Si $x^4 = w$ entonces $y^4 = w^2$.

Resumiendo, tenemos que

$$\mathcal{C}(\mathbb{F}_{13}) = \underbrace{\{(x,y) : x^4 = y^4 = 1\}}_{\mathcal{C}_1} \cup \underbrace{\{(x,y) : x^4 = w, y^4 = w^2\}}_{\mathcal{C}_w}.$$

Por lo tanto

$$\#\mathcal{C}(\mathbb{F}_{13}) = \#\mathcal{C}_1 + \#\mathcal{C}_w.$$

Ahora, por $(*)$, tenemos que $\#\mathcal{C}_1 = \#\mathcal{C}_w = 16$.

Falta mostrar que las hipótesis del teorema son satisfechas y dejamos esto para el lector.

**3.4. Algunas construcciones de curvas con muchos puntos racionales.** Curvas con muchos puntos racionales son importantes desde el punto de vista de las aplicaciones. Los tres métodos descritos aquí tienen como objetivo construir curvas de manera que el número de puntos racionales se aproxime a los mejores registros existentes.

*3.4.1. Primer método.* Este método apareció en [35]. Sea $H$ un subgrupo de orden $v$ del grupo multiplicativo $\mathbb{F}_{q^r}^*$ y considere un polinomio separable $f_1(t) \in \mathbb{F}_{q^r}[t]$ tal que $\{\alpha \in \mathbb{F}_q : f_1(\alpha) = 0\} \subseteq H$.

Dado un par $(k,l) \in \mathbb{N} \times N$, definimos el polinomio

$$f(X,Y) = X^k \cdot f_1(X^l \cdot Y).$$

Podemos reducir el grado de la variable $X$ haciendo $X^v = 1$ y vamos a denotar por $\tilde{f}(X,Y)$ el polinomio reducido. El grado de $X$ del nuevo polinomio es como máximo $v-1$.

Las principales ventajas de la reducción son:

1. El género de la curva $\tilde{\mathcal{C}}$ asociada a $\tilde{f}$ es menor que el género de la curva $\mathcal{C}$ asociada a $f$.
2. La probabilidad de $\tilde{f}$ de ser absolutamente irreducible es mayor.

Dado un punto $(a,b) \in \bar{\mathbb{F}}_q \times \bar{\mathbb{F}}_q$ con $a \in H$, por construcción de $\tilde{f}$, tenemos que $\tilde{f}(a,b) = f(a,b)$. Así, los puntos racionales de $\mathcal{C}$ sobre $\mathbb{F}_{q^r}$ cuya primera coordenada pertenece a $H$, son puntos racionales de la curva $\tilde{\mathcal{C}}$.

Entonces, tenemos claramente que

$$\#\tilde{\mathcal{C}}(\mathbb{F}_{q^r}) \geq v \cdot \deg(f_1).$$

De hecho, sea $d = \deg(f_1)$ y sean $\alpha_1, \ldots, \alpha_d \in H$ todas las raíces de $f_1$. Dado $a \in H$ considere $b_j = \alpha_j / a^l$ para $j = 1, \ldots, d$. Entonces $\tilde{f}(a, b_j) = f(a, b_j) = a^k f_1(a^l b_j) = a^k f_1(\alpha_j) = 0$.

**Ejemplo 3.66.** Tome $f_1(t) = t^5 + t^2 + 1 \in \mathbb{F}_{32}[t]$. Este polinomio tiene 5 raíces en $\mathbb{F}_{32}^*$, así podemos tomar $v = 31$. Considerando

$$f(X,Y) = X^7 f_1(X^{12}Y) = X^{67}Y^5 + X^{31}Y^2 + X^7,$$

tenemos que

$$\tilde{f}(X,Y) = X^5Y^5 + Y^2 + X^7$$

es absolutamente irreducible y así

$$\#\tilde{\mathcal{C}}(\mathbb{F}_{32}) \geq 5 \cdot 31 = 155.$$

En realidad faltan solo 3 puntos racionales: el origen y dos puntos en el infinito, así $\tilde{g} = 15$ y $\#\tilde{\mathcal{C}}(\mathbb{F}_{32}) = 158$.

*3.4.2. Segundo método.* Este método creado por van der Geer-van der Vlugt está descrito en [24].

Sea $R(X) \in \mathbb{F}_q[X]$ un polinomio separable que tiene todas sus raíces en $\mathbb{F}_q$. Consideremos ahora dos polinomios $R_1(X), R_2(X) \in \mathbb{F}_q[X]$ tales que

$$R(X) = R_1(X) + R_2(X).$$

La curva $\mathcal{C}$ asociada al polinomio

$$f(X,Y) = Y^{q-1} + \frac{R_1(X)}{R_2(X)}$$

tiene muchos puntos racionales ya que

$$\{(a,b) : R(a) = 0, R_1(a) \neq 0 \text{ y } b \in \mathbb{F}_q^*\} \subseteq \mathcal{C}(\mathbb{F}_q),$$

lo que nos da

$$\#\mathcal{C}(\mathbb{F}_q) \geq (q-1)\#\{a \in \mathbb{F}_q : R(a) = 0, R_1(a) \neq 0\}.$$

Para mantener el género bajo, es deseable que el producto $R_1(X) \cdot R_2(X)$ sea altamente inseparable. Vamos a ver un ejemplo.

**Ejemplo 3.67.** Sea $R(X) = X^{16} + X \in \mathbb{F}_{16}[X]$ y tomemos $R_1(X) = X^{16} + X^2$ y $R_2(X) = X^2 + X$. Con esta elección tenemos que

$$f(X,Y) = Y^{15} + \frac{(X^8 + X)^2}{X^2 + X} \quad \text{y} \quad \#\mathcal{C}(\mathbb{F}_{16}) = 213.$$

De hecho

$$\#\{a \in \mathbb{F}_{16} : R(a) = 0, R_1(a) \neq 0\} = 14,$$

luego $\#\mathcal{C}(\mathbb{F}_{16}) \geq 15 \cdot 14 = 210$. Los tres puntos que faltan son $(0,0), (0,1)$ y un punto en el infinito. La parte difícil aquí es mostrar que el género es 49.

Para calcular el género necesitamos introducir algunas propiedades de las extensiones de Kummer. Vamos a comenzar concentrándonos en extensiones del tipo

$$Y^m = f(X) \in \mathbb{F}_q(X), \quad \text{donde } m \text{ divide } q-1.$$

Suponga que la función racional $f(X)$ pueda ser escrita como $f(X) = \frac{g(X)}{h(X)}$ donde $g$ y $h$ son polinomios primos entre sí y definidos sobre $\mathbb{F}_q$. Digamos, para fijar la notación, que

$$g(X) = \prod_{i=1}^{r} (X - \alpha_i)^{m_i} \quad \text{y} \quad h(X) = \prod_{j=1}^{s} (X - \beta_j)^{n_j}$$

donde $\alpha_i, \beta_j \in \bar{\mathbb{F}}_q$ son distintos.

La fórmula de Riemann-Hurwitz afirma que

$$2g(\mathcal{C}) - 2 = m(-2) + D_{\text{ceros}} + D_{\text{polos}} + D_\infty$$

donde $\mathcal{D}_{\text{ceros}}, \mathcal{D}_{\text{polos}}$ y $\mathcal{D}_\infty$ son ciertos divisores y $D_{\text{ceros}}, D_{\text{polos}}, D_\infty$ son los grados respectivos. Hasse mostró que, si denotamos

$$\begin{cases} m_\infty = |\deg(g) - \deg(h)|, \\ d_\infty = \text{mcd}(m_\infty, m), \end{cases}$$

entonces

$$D_\infty = d_\infty(e_\infty - 1) = m - d_\infty, \quad \text{donde } e_\infty = \frac{m}{d_\infty}.$$

El conjunto $\{\alpha_i, i = 1, \ldots, r\}$ contribuirá con $D_{\text{ceros}}$ con

$$D_{\text{ceros}} = \sum_{i=1}^{r} d(\alpha_i)(e(\alpha_i) - 1) = rm - \sum_{i=1}^{r} d(\alpha_i),$$

ya que

$$\begin{cases} d(\alpha_i) = \text{mcd}(m_i, m), \\ e(\alpha_i) = \frac{m}{d(\alpha_i)}. \end{cases}$$

Análogamente, el conjunto $\{\beta_j : j = 1, \ldots, s\}$ contribuirá con $D_{\text{polos}}$ con

$$D_{\text{polos}} = \sum_{j=1}^{s} d(\beta_j)(e(\beta_j) - 1) = sm - \sum_{j=1}^{s} d(\beta_j).$$

Así, en el ejemplo anterior reescribiendo las ecuaciones como

$$Y^{15} = \left(\frac{X^8 + X}{X^2 + X}\right)^2 (X^2 + X),$$

donde $\frac{X^8+X}{X^2+X}$ es un polinomio mónico de grado 6 cuyas seis raíces son exactamente los elementos de $\mathbb{F}_8 \setminus \mathbb{F}_2$. Tenemos que $f$ tiene seis raíces duplas y dos raíces simples (las raíces de $X^2 + X$). Con las notaciones anteriores

$$\begin{cases} m_\infty = 15, \\ D_\infty = 14, \\ D_{\text{ceros}} = 112. \end{cases}$$

Lo que nos da que el género es igual a 49 como fue afirmado.

*3.4.3. Tercer método.* Este método fue inspirado por el método anterior, así que continuaremos trabajando con extensiones de Kummer. Fue propuesto por [17] como una generalización de la idea de [16].

Considere dos polinomios $f(X), \ell(X) \in \mathbb{F}_q[X]$ tales que

1. $\deg(f) \geq \deg(\ell)$
2. $\ell(X) \nmid f(X)$.

Así, existe un polinomio $r(X) \in \mathbb{F}_q[X]$ tal que $f(X) = h(X)\ell(X) + r(X)$ y $\deg(r) < \deg(\ell)$.

Sea $\mathcal{C}$ la curva proyectiva no singular asociada a

$$Y^m = \frac{f(X)}{r(X)}, \quad \text{donde } m \text{ divide } q - 1.$$

La idea es nuevamente que $f(X)r(X)$ sea altamente inseparable para garantizar un género bajo y que $\ell(X)$ tenga todas sus raíces en $\mathbb{F}_q$ para que tengamos muchos puntos racionales. De hecho, considere el conjunto

$$S = \{\alpha \in \mathbb{F}_q : h(\alpha)\ell(\alpha) = 0 \text{ y } f(\alpha) \neq 0\}.$$

Así, para todo $\alpha \in S$ tenemos que $f(\alpha)/r(\alpha) = 1$. Esto nos da la siguiente cota para el número de puntos racionales:

$$\#\mathcal{C}(\mathbb{F}_q) \geq m \cdot \#S.$$

Ahora, el género puede ser calculado sabiendo que la extensión es de Kummer.

**Ejemplo 3.68.** Sean $f(X) = (X^3 + X^2 + 1)^4$ y $\ell(X) = \frac{X^{16}+X}{X^4+X}$ definidos sobre $\mathbb{F}_{16}$. En este caso $r(X) = X^3(X+1)^3(X^3+X+1)$. La curva proyectiva asociada a

$$Y^3 = \frac{f(X)}{r(X)}$$

tiene género 4 y 45 puntos racionales sobre $\mathbb{F}_{16}$.

**3.5. Curvas maximales.** Como definimos antes, una curva $\mathcal{C}$ proyectiva, geométricamente irreducible, no singular de género $g$ definida sobre $\mathbb{F}_{q^2}$ es maximal cuando el número de puntos racionales alcanza la cota de Hasse-Weil, o sea

$$\#\mathcal{C}(\mathbb{F}_{q^2}) = q^2 + 1 + 2gq.$$

Así, tenemos un número finito de posibilidades para el género de una curva maximal ya que Ihara mostró en [32] que

$$g \leq \frac{q(q-1)}{2}.$$

De hecho, si $\mathcal{C}$ es maximal sobre $\mathbb{F}_{q^2}$, entonces $\alpha_i = -q$ para todo $i = 1, \ldots, 2g$.

Como $N_2 \geq N_1$, tenemos que

$$N_2 := q^2 + 1 - \sum_{i=1}^{2g} \alpha_i^2 \leq N_1 = q + 1 - \sum_{i=1}^{2g} \alpha_i,$$

lo que es equivalente a

$$q^2 - 2gq^2 \geq q + 2gq,$$

y a partir de esta desigualdad podemos mostrar la cota de Ihara.

La curva Hermitiana del Ejemplo 3.55 definida sobre $\mathbb{F}_{q^2}$ es una curva de género máximo y es la única curva, a menos de isomorfismo, con esta propiedad. Este resultado fue probado por [43].

Vamos a considerar dos problemas sobre curvas $\mathbb{F}_{q^2}$-maximales:

1. Determinar todos los géneros posibles para una curva maximal.
2. Clasificar las curvas maximales de un dado género.

*3.5.1. Géneros posibles.* Stichtenoth y Xing conjeturaron en [48] que el género de una curva maximal sobre $\mathbb{F}_{q^2}$ debe satisfacer

$$g = g_1 = \frac{q(q-1)}{2} \quad \text{o} \quad g \leq \frac{(q-1)^2}{4}.$$

Esta conjetura fue mostrada por Fuhrmann y Torres en [14] usando un caso particular de la Teoría de ordenes de Frobenius asociada a un sistema linear. Tal herramienta fue creada por Stöhr-Voloch en [49].

El punto clave de la prueba de Fuhrmann-Torres es el hecho que existe un punto $\mathbb{F}_{q^2}$-racional $P_0 \in \mathcal{C}$ tal que $q$ y $q+1$ son no lagunas en $P$, o sea, existen funciones $x, y$ tales que $\mathrm{div}_\infty(x) = qP_0$ y $\mathrm{div}_\infty(y) = (q+1)P_0$.

La conjetura fue probada aplicando las técnicas de Stöhr-Voloch al sistema linear definido por $\mathcal{D} := |(Q+1)P_0|$ Así, existe un segundo mayor género $g_2$ dado por

$$g_2 := \begin{cases} \frac{(q-1)^2}{4} & \text{si } q \text{ es impar,} \\ \frac{q(q-2)}{4} & \text{si } q \text{ es par.} \end{cases}$$

Nuevamente existe una única curva maximal, a menos de isomorfismo, de género $g_2$.

En el caso de característica impar, fue mostrado en [13] que la curva maximal de $g_2$ es isomorfa a la curva cuyo modelo plano está dado por

$$Y^q + Y = X^{(q+1)/2}.$$

Esta curva es maximal ya que está cubierta por la curva Hermitiana y todo cubrimiento de una curva maximal es maximal también [39].

En el caso de característica par, en [5] fue mostrado que la curva cuyo modelo plano está dado por

$$Y^{q/2} + Y^{q/4} + \cdots + Y = X^{q+1}$$

es maximal y tiene género $g_2$, y es única si $q/2$ es una no laguna en algún punto. Finalmente la unicidad fue mostrada en [38], ya que $q/2$ es siempre una no laguna en algún punto. Más aún, ellos mostraron que existe un tercer mayor género $g_3$ dado por

$$g_3 = \lfloor (q^2 - q + 4)/6 \rfloor.$$

Resumiendo, tenemos el siguiente teorema.

**Teorema 3.69.** *Si $\mathcal{C}$ es una curva maximal definida sobre $\mathbb{F}_{q^2}$ de género $g$, entonces*

$$g = g_1, \quad g = g_2 \quad o \quad g \leq g_3.$$

*Observación* 3.70. El resultado es el mejor posible ya que existe una curva maximal de género $g_3$.

*Observación* 3.71. Si $q \equiv 2 \mod 3$, entonces $g_4 = g_3 - 1$.

Existen solamente 12 ejemplos de curvas maximales salvo isomorfismos, cuyo género satisface

$$\lfloor (q-1)(q-3)/8 \rfloor \leq g < \lfloor (q-1)^2/4 \rfloor,$$

a saber:

1. $g = \lfloor (q^2 - q + 4)/6 \rfloor$ para $q \equiv 0, 1, 2 \mod 3$ [38];
2. $g = \lfloor (q^2 - q - 2)/6 \rfloor$ para $q \equiv 2 \mod 3$ [22], [12];
3. $g = \lfloor (q-1)(q-2)/6 \rfloor$ para $q \equiv 0, 2 \mod 3$ [38];
4. $g = \lfloor (q^2 - 2q + 5)/8 \rfloor$ para $q \equiv 0, 1, 3 \mod 4$ [38];
5. $g = \lfloor (q-1)(q-3)/8 \rfloor$ para $q \equiv 0, 1, 3 \mod 4$ [38].

El objetivo de todas las investigaciones realizadas inicialmente (vea [22], [11], [4], [2]) para encontrar nuevos valores para el género de un cuerpo de funciones maximal $F|\mathbb{F}_{q^2}$ se concentraba en estudiar subcuerpos del cuerpo de funciones Hermitiano $H = \mathbb{F}_{q^2}(x, y)$ donde $y^q + y = x^{q+1}$. Giulietti y Korchmáros construyeron el primer ejemplo de un cuerpo de funciones maximal que no puede ser obtenido como subcuerpo del Hermitiano [25]. Otros ejemplos pueden ser vistos en [50].

En general, en [22] y [4], los autores consideran $H|\mathbb{F}_{q^2}$ como una extensión galoisiana de $\mathbb{F}_{q^2}(x)$ y para algunos subgrupos no moderados del grupo de automorfismos de la Hermitiana computan el género del cuerpo fijo por ese subgrupo.

A menos de conjugación por $PGU(3, \mathbb{F}_{q^2})$ los grupos no moderados están contenidos en el grupo de descomposición $\mathcal{A}(P_\infty)$ del único lugar $P_\infty$ sobre el polo de la función $x$.

En nuestro caso, el grupo de descomposición $\mathcal{A}(P_\infty)$ consiste en todos los automorfismos que satisfacen

$$\begin{cases} \sigma(x) = ax + b, \\ \sigma(y) = a^{q+1}y + ab^q x + c, \end{cases}$$

donde $a \in \mathbb{F}_{q^2}^*, b \in \mathbb{F}_{q^2}$ y $c^q + c = b^{q+1}$.

El grupo $\mathcal{A}(P_\infty)$ tiene orden $q^3(q^2-1)$ y contiene un número grande de subgrupos.

*Observación* 3.72. Suponga que la característica es impar. Sean $\mathcal{G}$ un subgrupo de $\mathcal{A}(P_\infty)$,

$$V = \{b \in \mathbb{F}_{q^2} : \text{ existe } c \in \mathbb{F}_{q^2} \text{ tal que } [1, b, b^{q+1}/2] \in \mathcal{G}\}$$

un subgrupo de $\mathbb{F}_{q^2}$ de orden $p^v$ y

$$W = \{c \in \mathbb{F}_{q^2} : \text{ tal que } [1, 0, c] \in \mathcal{G}\}$$

un subgrupo de $\mathbb{F}_{q^2}$ de orden $p^w$.

Con estas notaciones tenemos que

$$ord(\mathcal{G}) = mp^{v+w} \quad \text{con } m \mid q^2 - 1.$$

Ahora, el género del cuerpo fijo por $\mathcal{G}$ está dado por ([22])

$$g(H^{\mathcal{G}}) = (p^{n-w} - 1)(p^{n-v} + 1 - d)/2m, \quad \text{donde } d = \text{mcd}(m, q+1).$$

Ahora, para $m$ un divisor de $q-1$, definimos

$$s = \text{orden de } p \text{ en el grupo multiplicativo } (\mathbb{Z}/m\mathbb{Z})^*,$$

$$r = \begin{cases} \text{orden de } p \text{ en } (\mathbb{Z}/\frac{m}{2}\mathbb{Z})^* \text{ si } m \equiv 0 \mod 2, \\ s \text{ caso contrario.} \end{cases}$$

**Teorema 3.73.** [4] *Con las notaciones anteriores, para cada divisor $m$ de $q-1$, para todos los múltiplos $v$ de $s$ satisfaciendo $0 \le n \le s$ y para todos los múltiplos $w$ de $r$ satisfaciendo $0 \le w \le n-1$, existe un subgrupo $\mathcal{G}$ de $\mathcal{A}(P_\infty)$ de orden $mp^{v+w}$ tal que el género del cuerpo fijo por $H^{\mathcal{G}}$ es*

$$g = \begin{cases} (p^{n-w} - 1)(p^{n-v} - 1)/2m, & m \equiv 0 \mod 2, \\ (p^{n-w} - 1)p^{n-v}/2m, & m \equiv 1 \mod 2. \end{cases}$$

**Corolario 3.74.** [4] *Para $m = 1$ y para todo $0 \le v \le n$ y $0 \le w \le n-1$, existe un $p$-subgrupo de $\mathcal{A}(P_\infty)$ tal que el género del cuerpo fijo por ese subgrupo es*

$$g = p^{n-v}(p^{n-w} - 1)/2.$$

Sea $m$ un divisor de $q^2 - 1$ tal que $m$ no divide $q-1$. Definimos

$$d = \text{mcd}(m, q+1),$$
$$s = \text{el orden de } p \text{ en } (\mathbb{Z}/m\mathbb{Z})^*,$$
$$r = \text{el orden de } p \ (\mathbb{Z}/\frac{m}{d}\mathbb{Z})^*.$$

**Teorema 3.75.** [4] *Para cada $m$ divisor de $q^2 - 1$ tal que $m$ no divide a $q-1$ y para cada $v$ y $w$ satisfaciendo*

1. $0 \le v \le n, v \mid 2n, v \nmid n$ *y $v$ es divisible por $s$,*
2. $v/2 \le w \le n$, *y $w$ es divisible por $r$,*

*existe un subgrupo $\mathcal{G}$ de $\mathcal{A}(P_\infty)$ de orden $mp^{v+w}$ tal que el género del cuerpo fijo por $H^{\mathcal{G}}$ es*

$$g = \frac{(p^{n-w} - 1)(p^{n-v} - d + 1)}{2m}.$$

Identificamos $\sigma \in \mathcal{A}(P_\infty)$ con el par $[b, c]$ si

$$\begin{cases} \sigma(x) = x + b, \\ \sigma(y) = y + b^q x + c, \end{cases}$$

donde $b \in \mathbb{F}_{q^2}$ y $c^q + c = b^{q+1}$.

Definimos $\phi : \mathcal{G} \to \mathbb{F}_{q^2}$ por

$$\phi(\sigma) = b \quad \text{si } \sigma \text{ está identificado con el par } [b, c].$$

Esta aplicación es un morfismo suryectivo en un subgrupo aditivo de $\mathbb{F}_{q^2}$. Definimos

$$V := Im(\phi) \quad \text{y} \quad W = \{c \in \mathbb{F}_{q^2} : [0, c] \in \mathcal{G}\}.$$

Tenemos que $V$ y $W$ son subgrupos aditivos de $\mathbb{F}_{q^2}$ de ordenes

$$ord(V) = 2^v, \quad ord(W) = 2^w \quad \text{y} \quad ord(\mathcal{G}) = 2^{v+w}.$$

**Teorema 3.76.** [22] *Sea $q = 2^n$ y $g \geq 1$ un entero. Entonces son equivalentes:*

1. *Existe un 2-subgrupo $\mathcal{G} \subseteq \mathcal{A}$ tal que $g = g(H^{\mathcal{G}})$;*
2. *$g = 2^{n-v-1}(2^{n-w} - 1)$ con $0 \leq v, w \leq n - 1$ y existen $\mathbb{F}_2$-espacios vectoriales $V \subseteq \mathbb{F}_{q^2}$ y $W \subseteq \mathbb{F}_q$ de ordenes $2^v$ y $2^w$ respectivamente, tales que $V^{q+1} = \{b^{q+1} : b \in V\}$ está contenido en $W$.*

**Teorema 3.77.** *Para cada $1 \leq v \leq n-1$ tal que $v = s+k$ con $s|n$, $0 \leq k \leq s$ y para cada $w$ satisfaciendo $s \leq w \leq n - 1$, las siguientes afirmaciones son equivalentes:*

1. *Existe un 2-subgrupo $\mathcal{G} \subseteq \mathcal{A}$ tal que $g = g(H^{\mathcal{G}})$;*
2. *$g = 2^{n-v-1}(2^{n-w} - 1)$.*

*Observación* 3.78. En [11] los autores consideraron subgrupos de orden $p$ y dieron ecuaciones explícitas para los subcuerpos que obtuvieron.

*3.5.2. Problema de clasificación.* Para el problema de clasificación vamos a estudiar dos tipos:

1. Existe un punto racional $P_0 \in \mathcal{C}$ tal que $m$ es una no laguna en $P_0$ con $m \cdot n = q + 1$ y $2g = (q - 1)(m - 1)$.
2. Existe un punto racional $P_0 \in \mathcal{C}$ tal que $m$ es una no laguna en $P_0$ con $m \cdot n = q$ y $2g = q(m - 1)$ con una hipótesis extra.

Estos dos tipos aparecen naturalmente ya que si $\mathcal{C}$ es una curva maximal, entonces $q$ y $q + 1$ son no lagunas en cualquier punto racional.

El primer caso, fue estudiado en [13] donde mostraron que existe una única curva maximal, a menos de isomorfismo y que está dada por

$$Y^q + Y = X^m.$$

Ahora, no podemos esperar tener unicidad en el segundo caso, como muestra el siguiente ejemplo.

**Ejemplo 3.79.** Sea $q = 2^6$ y considere las curvas $\mathcal{C}_1$ y $\mathcal{C}_2$ definidas sobre $\mathbb{F}_{q^2}$ como

$$\mathcal{C}_1 = (X^{65} + Z^{16} + Z^4 + Z = 0),$$
$$\mathcal{C}_2 = (X^{65} + W^{16} + W^8 + W^2 + W = 0).$$

Son curvas maximales, ya que son cubiertas por la Hermitiana (basta hacer $Z = Y^4 + Y$ y $W = Y^4 + Y^2 + Y$). Ambas tienen el mismo género y no son isomorfas (vea [1], [3]).

*Observación* 3.80. En [1] fue mostrado que en característica 2, para todo $n \in \mathbb{N}$ existen $2^{n-1}$ curvas maximales no isomorfas con el mismo género. En el segundo caso, Fuhrmann conjeturó que las curvas deberían ser isomorfas a

$$F(Y) = X^{q+1}, \quad \text{donde } F(Y) \text{ es un polinomio aditivo de grado } m.$$

De hecho, los ejemplos mencionados antes, son todos de este tipo.

**Teorema 3.81.** [3] *Sea $\mathcal{C}$ una curva maximal definida sobre $\mathbb{F}_{q^2}$ de género $g = (m-1)q/2$ donde $m$ es una no laguna en $P_0 \in \mathcal{C}$ tal que $nm = q$. Suponga que la extensión $\mathbb{F}_{q^2}(\mathcal{C})|\mathbb{F}_{q^2}(x)$ es una extensión de Galois donde $x \in \mathbb{F}_{q^2}(\mathcal{C})$ es una función tal que $\mathrm{div}_\infty(x) = mP_0$. Entonces la curva $\mathcal{C}$ es isomorfa a la curva dada por*

$$P(z) = A(x),$$

*donde $P(z) \in \mathbb{F}_{q^2}[z]$ es un polinomio aditivo separable de grado $m$ y $A(x) \in \mathbb{F}_{q^2}[x]$ es un polinomio de grado $q+1$.*

Para el caso $m = p$ (donde $p$ es la característica del cuerpo), sin ninguna otra hipótesis, fue mostrado que la curva $\mathcal{C}$ es $\mathbb{F}_{q^2}$-isomorfa a la curva dada por

$$\sum_{i=1}^{t} z^{p^{t-i}} = cx^{p^t+1},$$

donde $c \in \mathbb{F}_{q^2}^*$ es tal que $c^q + c = 0$ y $q = p^t$.

Este resultado generaliza algunos resultados de [5] y [6].

**3.6. Comportamiento asintótico.** El objetivo de esta sección es el de presentar la cota de Drinfeld-Vladut y dar algunos ejemplos de cuando la cota es alcanzada.

Definimos

$$N_q(g) := \text{máx}\{N(F) : F|\mathbb{F}_q \text{ es un cuerpo de funciones de género } g\}$$

y

$$A(q) := \limsup_{g\to\infty} \frac{N_q(g)}{g}.$$

La cantidad $A(q)$ fue introducida por Ihara y satisface

$$A(q) \leq \sqrt{2q + \frac{1}{4}} - \frac{1}{2}.$$

Esta cota fue mejorada por Drinfeld-Vladut [51].

**Teorema 3.82.** *(Cota de Drinfeld-Vladut)*

$$A(q) \leq \sqrt{q} - 1.$$

La prueba del teorema usa el método de Serre de fórmulas explícitas.

*Demostración.* Para todo $m \in \mathbb{N}$, considere el polinomio $\Psi_m(t)$ dado por

$$\Psi_m(t) = \sum_{r=1}^{m} c_r t^r = \sum_{r=1}^{m} \big(1 - \frac{r}{m}\big) t^r,$$

así el grado es $m-1$. Para $|t| \neq 1$, podemos escribir $\Psi_m(t)$ como

$$\Psi_m(t) = \frac{t}{(t-1)^2} \cdot \Big(\frac{t^m - 1}{m} + 1 - t\Big).$$

Definiendo $f_m(t) := 1 + \Psi_m(t) + \Psi_m(t^{-1})$, tenemos que

$$f_m(t) = \frac{2 - (t^m + t^{-m})}{m(t-1)(t^{-1}-1)}.$$

Si $\gamma \in \mathcal{C}$ es tal que $|\gamma| = 1$, entonces $(\gamma - 1) \cdot (\gamma^{-1} - 1)$ es un número real positivo y $|\gamma^m + \bar{\gamma}^{-m}| \leq 2$. Esto implica que

$$f_m(t) = \frac{2 - (\gamma^m + \bar{\gamma}^{-m})}{m|\gamma - 1|^2} \geq 0,$$

para todo $\gamma \in \mathcal{C}$ con $|\gamma| = 1$. Así, tenemos

$$\frac{N_q(g)}{g} \leq \frac{1}{\Psi_m(q^{-1/2})} + \frac{1}{g}\Big(\frac{\Psi_m(q^{1/2})}{\Psi_m(q^{-1/2})} + 1\Big).$$

Si $m \to \infty$ entonces

$$\Psi_m(q^{-1/2}) \to \frac{1}{q^{1/2} - 1}.$$

Entonces para todo $\epsilon > 0$ existe $g_0$ tal que para todo $g > g_0$ vale

$$\frac{N}{g} < q^{1/2} - 1 + \epsilon.$$

El teorema sigue de la desigualdad anterior. □

El valor exacto de $A(q)$ es desconocido para casi todos los valores de $q$. Serre mostró que $A(q) > 0$ para todo $q$. Cuando $q$ es un cuadrado, Ihara mostró que $A(q) > \sqrt{q} - 1$ usando técnicas profundas de curvas modulares de Shimura y su argumento no es constructivo.

En particular, Ihara nos dice que

$$A(q) = \sqrt{q} - 1 \quad \text{cuando } q \text{ es un cuadrado.}$$

El primer ejemplo de una construcción explícita sobre $\mathbb{F}_{q^2}$ que alcanza la cota de Drinfeld-Vladut se debe a García y Stichtenoth en [20]. Desde entonces otras construcciones explícitas fueron apareciendo.

Vamos introducir algunos conceptos nuevos (vea [21]).

**Definición 3.83.** Una *torre* $\mathcal{F}$ sobre $\mathbb{F}_q$ es una familia infinita de cuerpos

$$F_0 \subseteq F_1 \subseteq F_2 \subseteq \cdots \subseteq F_n \subseteq F_{n+1} \subseteq \cdots$$

donde

1. $F_n|\mathbb{F}_q$ es un cuerpo de funciones para todo $n$;
2. $F_{n+1}|F_n$ es una extensión finita y separable para todo $n$;
3. $g_n := g(F_n)$ va para infinito cuando $n$ va para infinito.

El *límite de la torre* $\mathcal{F}$ está definido por

$$\lambda_q(\mathcal{F}) := \lim_{n\to\infty} \frac{N(F_n)}{g(F_n)}.$$

*Observación* 3.84. El número $\lambda_q(\mathcal{F})$ existe. De hecho, si $E|F$ es una extensión separable de cuerpos de funciones definidos sobre $\mathbb{F}_q$, entonces

$$\frac{N(E)}{g(E) - 1} \leq \frac{N(F)}{g(F) - 1}.$$

Así, la sucesión $\{N(F_n)/g(F_n)\}$ es no creciente y por lo tanto es convergente.

**Definición 3.85.** Sea $\mathcal{F} = (F_0 \subseteq F_1 \subseteq F_2 \subseteq \cdots \subseteq F_n \subseteq F_{n+1} \subseteq \cdots)$ una torre sobre $\mathbb{F}_q$ y suponga que existan un polinomio $\Phi(X,Y) \in \mathbb{F}_q[X,Y]$ y elementos $x_r \in F_r$ para todo $r \geq 0$ tales que

1. $F_0 = \mathbb{F}_q(x_0)$;
2. $F_n = F_{n-1}(x_n)$ y $\Phi(x_{n-1}, x_n) = 0$ para todo $n \geq 1$;
3. $\Phi(x_{n-1}, Y) \in F_{n-1}[Y]$ es absolutamente irreducible sobre $F_{n-1}$,

entonces decimos que la torre $\mathcal{F}$ está *definida recursivamente* por $\Phi(X,Y)$.

*Observación* 3.86. Si $\mathcal{F}$ está definida recursivamente por $\Phi(X,Y)$ y $\lambda_q(\mathcal{F}) > 0$, entonces $\deg_X(\Phi) = \deg_Y(\Phi)$.

**Definición 3.87.** Una torre $\mathcal{F}$ es *moderada* si y solamente si existe $F \in \mathcal{F}$ tal que la extensión $E|F$ es moderada para toda extensión $E|F$ y $E \in \mathcal{F}$.

Para cada $F \in \mathcal{F}$, definimos el conjunto de ramificación sobre $F$ como

$V_F = \{P \in \mathbb{P}_{F \cdot \bar{\mathbb{F}}_q} : P \text{ es ramificado en alguna extensión finita } E|F \text{ y } E \in \mathcal{F}\}$.

**Definición 3.88.** La torre $\mathcal{F}$ es del *tipo de ramificación finita* si existe $F \in \mathcal{F}$ tal que $\#V_F < \infty$.

**Definición 3.89.** Una torre $\mathcal{F}$ es de *descomposición completa* si existe $F \in \mathcal{F}$ y un lugar racional $P \in \mathbb{P}_F$ tal que $P$ se descompone totalmente en toda extensión $E|F$ con $E \in \mathcal{F}$.

Para $F \in \mathcal{F}$ definimos

$\mathcal{T}_F(\mathcal{F}) = \{P \in \mathbb{P}_F : P \text{ es racional y se descompone totalmente } \forall E|F \text{ con } E \in \mathcal{F}\}$.

Así, $\mathcal{F}$ se descompone totalmente si y solamente si existe $F \in \mathcal{F}$ tal que $\mathcal{T}_F(\mathcal{F}) \neq \emptyset$.

*Observación* 3.90. Para todo $F \in \mathcal{F}$ tenemos

$$0 \leq \#\mathcal{T}_F(\mathcal{F}) \leq N(F).$$

**Proposición 3.91.** *Sea $\mathcal{F}$ una torre moderada con ramificación de tipo finito y que se descomponga completamente, definida sobre $\mathbb{F}_q$. Sea $F \in \mathcal{F}$ tal que $\#V_F < \infty$; $\#T_F(\mathcal{F}) \geq 1$, y tal que $E|F$ es moderada para todo $E \in \mathcal{F}$. Entonces*

$$\lambda_q(\mathcal{F}) \geq \frac{2\#\mathcal{T}_F(\mathcal{F})}{2g(F) - 2 + \#V_F}.$$

Ahora, volviendo a la cota de Drinfeld-Vladut, la torre de García-Stichtenoth [20] fue la primera construcción explícita que alcanzó la cota. La torre fue definida como

1. $F_0 = \mathbb{F}_{q^2}(x_0)$;
2. $F_1 = F_0(z)$ donde $z_1^q + z_1 = x_0^{q+1}$, definiendo $x_1 = z_1/x_0$;
3. $F_2 = F_1(z_2)$ donde $z_1^q + z_1 = x_1^{q+1}$, y así sucesivamente.

Cada extensión es de Artin-Schreier y usando las propiedades de este tipo de extensión, se puede controlar la ramificación en cada paso de la torre, calcular los géneros y estimar el número de puntos racionales.

### 3.7. Ejercicios.

1. Complete la prueba del Teorema 3.58: Muestre que $N-(q+1) \geq -(g\lfloor 2\sqrt{q}\rfloor)$ usando $\gamma_i := -(\alpha_i + \bar{\alpha}_i) + \lfloor 2\sqrt{q}\rfloor + 1$ en lugar de $\beta_i$ para $i = 1, \ldots, g$.
2. Considere la cuártica de Klein del Ejemplo 3.59.
   *a*) Muestre que tiene exactamente dos puntos en el infinito.
   *b*) Analice los puntos racionales de la forma $(x, 0, 1), (0, y, 1) \in \mathbb{F}_8^3$ y deduzca que pertenecen a la curva si y solamente si $x = y = 0$.
3. Muestre que la curva de Suzuki del Ejemplo 3.61 tiene $q^2+1$ puntos racionales.
4. Considere la curva de Fermat del Ejemplo 3.65.
   *a*) Muestre que la curva no tiene puntos racionales en el infinito.
   *b*) Muestre que la curva no tiene puntos racionales afines donde una de las coordenadas sea cero.
5. Muestre que $\sigma$ es un automorfismo del cuerpo de funciones Hermitiano, donde
$$\begin{cases} \sigma(x) = ax + b, \\ \sigma(y) = a^{+1}y + ab^q x + c, \end{cases}$$
$a \in \mathbb{F}_{q^2}^*, b \in \mathbb{F}_{q^2}$ y $c^q + c = b^{q+1}$.
6. Considere la curva proyectiva asociada al polinomio
$$f(X, Y) = X^3Y + Y^3 + X + X^2Y^2 + Y^2 + X^2 + X^2Y + XY^2.$$
   *a*) Muestre que la curva es no singular y tiene género 3.
   *b*) Muestre que tiene 3 puntos en el infinito.
   *c*) Muestre que $f(x, y) = 0$ para todo $x, y \in \mathbb{F}_2$.
   *d*) Deduzca que $\#\mathcal{C}(\mathbb{F}_2) = 7$.

## Referencias


[1] M. Abdón *On maximal curves in characteristic two*, Ph.D. dissertation, IMPA **F-121**, 1–50 (2000).

[2] M. Abdón, J. Bezerra y L. Quoos, *Further examples of maximal curves*, J. Pure Appl. Algebra **213(6)**, 1192–1196 (2009). DOI 10.1016/j.jpaa.2008.11.037.

[3] M. Abdón y A. Garcia, *On a characterization of certain maximal curves*, Finite Fields Appl. **10 (2)**, 133–158 (2004). DOI 10.1016/j.ffa.2003.06.002.

[4] M. Abdón y L. Quoos, *On the genera of subfields of the Hermitian function field*, Finite Fields and Appl **10 (3)**, 271–284 (2004). DOI 10.1016/j.ffa.2003.08.003.

[5] M. Abdón y F. Torres, On maximal curves in characteristic two, Manuscripta Math. **99**, 39–53 (1999). DOI 10.1007/s002290050161.

[6] M. Abdón y F. Torres, *On $\mathbb{F}_{q^2}$-maximal curves of genus $q(q-3) = 6$*, Beitr. Algebra Geom. **46 (1)**, 241–260 (2005).

[7] E. Arbarello; M. Cornalba, M.; P.A. Griffiths y J. Harris, Geometry of Algebraic Curves, Springer-Verlag, New York, 1985. DOI:10.1007/978-1-4757-5323-3

[8] D. Bartoli, M. Montanucci y F. Torres *$\mathbb{F}_p$-maximal curves with many automorphisms are Galois covered by the Hermitian curve*, arXiv:1708.03933v2 [math.AG]

[9] C. Carvalho, On $\mathcal{V}$-Weierstrass sets and gaps, Journal of Algebra **312**:2, 956–962 (2007). DOI:10.1016/j.jalgebra.2006.11.016.

[10] C. Carvalho y F. Torres, On Goppa codes and Weierstrass gaps at several points, Des. Codes Cryptogr. **35**:2, 211–225 (2005). DOI:10.1007/s10623-005-6403-4

[11] A. Cossidente, G. Korchmáros y F. Torres, *Curves of large genus covered by the Hermitian curve*, Comm. Algebra **28 (1)**, 4707–4728 (2000). DOI 10.1080/00927870008827115.

[12] A. Cossidente, G. Korchmáros y F. Torres, *On curves covered by the Hermitian curve*, Composition Math **216**, 56–76 (1999). DOI 10.1006/jabr.1998.7768.

[13] R. Fuhrmann, A. Garcia y F. Torres, *On maximal curves*, J. Number Theory **67(1)**, 29–51 (1997). DOI 10.1006/jnth.1997.2148.

[14] R. Fuhrman y F. Torres, *The genus of curves over finite fields with many rational points*, Manuscripta Mathematica **89(1)**, 103–106 (1996). DOI 10.1007/BF02567508.
[15] W. Fulton, *Algebraic Curves: An Introduction to Algebraic Geometry*, Addison-Wesley, 245 páginas (1989).
[16] A. Garcia y A. Garzón, *On Kummer covers with many rational points over finite fields*, Journal of Pure and Applied Algebra 185, 177–192 (2003). DOI 10.1016/S0022-4049(03)00110-5.
[17] A. Garcia y L. Quoos, *A construction of curves over finite fields*, Acta Arith. **98**, 181–195 (2001).
[18] A. Garcia, S.J. Kim y R.F. Lax, Consecutive Weierstrass gaps and minimum distance of Goppa codes, J. Pure Appl. Algebra **84**:2, 199–207 (1993). DOI:10.1016/0022-4049(93)90039-V
[19] A. Garcia y R.F. Lax, Goppa codes and Weierstrass gaps, Coding theory and algebraic geometry (Luminy, 1991), 33–42, Lecture Notes in Math., 1518, Springer, Berlin, 1992. DOI:10.1007/BFb0087991
[20] A. Garcia y H. Stichtenoth, *A tower of Artin-Schreier extensions of functions fields attaining the Drinfeld-Vladut¸ bound*, Inventiones Math. **121**, 211–222 (1995). DOI 10.1007/BF01884295.
[21] A. Garcia y H. Stichtenoth, *On towers over finite fields*, J. Reine Angew. **557**, 53–80 (1995).
[22] A. Garcia, H. Stichtenoth y C.P. Xing, *On subfields of the Hermitian function field* Compositio Math **120**, 137–170 (2000) DOI 10.1023/A:1001736016924.
[23] J. von zur Gathen y D. Panario, *A survey on factoring polynomials over finite fields*, Journal of Symbolic Computation **31**, 3-17 (2001).
[24] G van der Geer y M. van der Vlugt, *Kummer covers with many points*, Finite Fields Appl. **6(4)**, 327–341 (2000). DOI 10.1006/ffta.2000.0286.
[25] M. Giulietti y G. Korchmáros, *A new family of maximal curves over a finite field*, Math. Ann. **343**, 229–245 (2009). DOI 10.1007/s00208-008-0270-z.
[26] D.M. Goldschmidt, Algebraic functions and projective curves, Graduate Texts in Mathematics, 215. Springer-Verlag, New York, 2003. DOI:10.1007/b97844
[27] G.L. Matthews, Weierstrass pairs and minimum distance of Goppa codes, Des. Codes Cryptogr. **22**:2, 107–121 (2001). DOI:10.1023/A:1008311518095
[28] R. Hartshorne, Algebraic Geometry, Graduate Texts in Math. **52**, 496 páginas (1997).
[29] J.W.P. Hirschfeld, G. Korchmáros y F. Torres, Algebraic curves over a finite field. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2008 DOI:10.1515/9781400847419
[30] M. Homma, The Weierstrass semigroup of a pair of points on a curve, Arch. Math. **67**:4, 337–348 (1996). DOI:10.1007/BF01197599
[31] M. Homma y S.J. Kim, Goppa codes with Weierstrass pairs, J. Pure Appl. Algebra **162**:2-3, 273–290 (2001). DOI:S0022-4049(00)00134-1
[32] Y. Ihara, *Some remarks on the number of rational points of algebraic curves over finite fields*, J. Fac. Sci. Tokio **28**, 721–724 (1981).
[33] N. Ishii, A certain graph obtained from a set of several points on a Riemann surface, Tsukuba J. Math. **23**:1, 55–89 (1999). DOI:10.21099/tkbjm/1496163776
[34] H. Janwa, On the parameters of algebraic geometric codes, Applied algebra, algebraic algorithms and error-correcting codes (New Orleans, LA, 1991), 19–28, Lecture Notes in Comput. Sci., 539, Springer, Berlin, 1991 DOI:10.1007/3-540-54522-0_92
[35] J. Justesen, K. Larsen, H. Jensen, H. Elbrnd, A. Havemose y T. Holdt, *Constructing ans decoding of a class of algebraic geometry codes*, IEEE Trans. Infor. Theory **35(4)**, 811–821 (1989). DOI 10.1109/18.32157.
[36] S.J. Kim, On the index of the Weierstrass semigroup of a pair of points on a curve, Arch. Math. **62**:1, 73–82 (1994). DOI:10.1007/BF01200442.
[37] G. Korchmáros y F. Torres, *Embedding of a maximal curve in a Hermitian variety*, Composition Math. **128**, 95–113 (2001). DOI 10.1023/A:1017553432375.
[38] G. Korchmáros y F. Torres, *On the genus of a maximal curve*, Math. Ann. **323(3)**, 589–608 (2002). DOI 10.1007/s002080200316.
[39] G. Lachaud, *Sommes d'Eisenstein e nombre de points de certaines curves algébriques sur les corps finis*, C. R. Acd. Sci. Paris **305** Série I, 729–732 (1987).
[40] R. Lidl y H. Niederreiter, *Finite Fields*, Cambridge University Press, Cambridge, second edition, 755 páginas (1997).
[41] G. Mullen y D. Panario, *Handbook of Finite Fields*, Discrete Mathematics and Its Applications Series, CRC Press, 1068 páginas (2013).

[42] H. Niederreiter y C. P. Xing, *Rational Points on Curves over Finite Fields*, London Mathematical Society, Lecture Notes Series **288**, 256 páginas (2001).
[43] H. G Rück y H. Stichtenoth, *A characterization of Hermitian function fields over finite fields*, J. Reine Angew. Math. **457**, 185–188 (1994).
[44] J.-P. Serre, *Sur le nombre des points rationnels d'une courbe algébrique sur un corps fini.*, C. R. Acad. Sci. Paris **296**, 397–402 (1983).
[45] I. R. Shafarevich, *Basic Algebraic Geometry I*, Springer, 310 páginas (2013).
[46] J. H. Silverman, *The Arithmetic of Elliptic Curves over Finite Fields*, Undergraduate Texts in Mathematics, Springer New York, 281 páginas (1994).
[47] H. Stichtenoth, *Algebraic Function Fields and Codes*, Graduate Texts in Mathematics **254**, 360 páginas (2008).
[48] H. Stichtenoth y C. P. Xing, *The genus of maximal function fields*, Manuscripta Math. **86**, 217–224 (1995). DOI 10.1007/BF02567990.
[49] K.O. Stöhr y J. F. Voloch, *Weierstrass points and curves over finite fields* Proc. London Math. Soc. 52, 1–19 (1986). DOI 10.1112/plms/s3-52.1.1..
[50] S. Tafazolian, A. Teherán-Herrera y F. Torres, *Further examples of maximal curves which cannot be covered by the Hermitian curve*, J. Pure Appl. Algebra **220(3)**, 1122–1132 (2016). DOI 10.1016/j.jpaa.2015.08.010.
[51] S.G. Vladut y V.G. Drinfeld, *Number of points of an algebraic curve*, Funct. Anal. **17(1)**, 68–69 (1983). DOI 10.1007/BF01083182.

Universidade Federal Fluminense, Instituto de Matemática e Estatística. Rua Professor Marcos Waldemar de Freitas Reis, s/n, Blocos G e H - Campus do Gragoatá, Niterói - RJ, CEP: 24210-201 - Brasil
*Email address*: miriam_abdon@id.uff.br

Universidade Federal de Uberlândia, Faculdade de Matemática, Av. J.N. de Ávila 2121, 38400-902 Uberlândia - MG, Brasil
*Email address*: cicero@ufu.br

Carleton University, School of Mathematics and Statistics, 1125 Colonel By Dr., K1S 5B6 Ottawa - Ontario, Canada
*Email address*: daniel@math.carleton.ca

Curso

# EQUIDISTRIBUCIÓN, TEORÍA DEL POTENCIAL Y APLICACIONES ARITMÉTICAS

JOSÉ IGNACIO BURGOS GIL Y RICARDO MENARES

Publicaciones Matemáticas del Uruguay
Volumen 18, Diciembre 2023, Páginas 61–101
S 0797-1443

# EQUIDISTRIBUCIÓN, TEORÍA DEL POTENCIAL Y APLICACIONES ARITMÉTICAS

JOSÉ IGNACIO BURGOS GIL Y RICARDO MENARES

Resumen. En este curso daremos una breve introducción a la teoría del potencial en el plano complejo así como varias aplicaciones aritméticas. Entre las aplicaciones aritméticas veremos el Teorema de Fekete-Szegö y el Teorema de Equidistribución de Bilu.

## Índice



Versión final: 8 de septiembre de 2019.
2010 *Mathematics Subject Classification.* 30C85, 37P30, 11G50.
El primer autor ha sido apoyado parcialmente por los proyectos MINECO MTM2016-79400-P e ICMAT Severo Ochoa project SEV-2015-0554.
El segundo autor contó con apoyo del proyecto Fondecyt 1171329.
Estas notas corresponden al curso dictado por los autores en la escuela AGRA III, Aritmética, Grupos y Análisis, del 9 al 20 de Julio de 2018 en Córdoba, Argentina.

## 1. Introducción

En muchas situaciones en el estudio de ecuaciones diofantinas, es posible establecer la distribución asintótica de familias distinguidas de puntos algebraicos, en cuyo caso decimos que hay equidistribución de la familia.

Un ejemplo básico de este fenómeno es el caso de las raíces de la unidad. Denotamos por $\mathbb{D} = \{z \in \mathbb{C} : |z| \leq 1\}$ al disco unitario. Las soluciones complejas de la ecuación $x^n = 1$ se sitúan sobre el borde de $\mathbb{D}$ formando los vértices de un polígono regular de $n$ lados. Cuando $n$ crece el polígono aproxima cada vez mejor al círculo, de donde se deduce que en el límite, las raíces de la unidad se distribuyen siguiendo la medida uniforme del círculo.

En este curso presentaremos el Teorema de Fekete-Szegö y el Teorema de Bilu, que constituyen generalizaciones diferentes de este fenómeno. En el primer caso, buscaremos reemplazar $\mathbb{D}$ por otro conjunto, en el segundo caso buscaremos reemplazar la familia de ecuaciones $x^n = 1$.

En el contexto del Teorema de Fekete-Szegö, reemplazamos $\mathbb{D}$ por un conjunto compacto $K \subseteq \mathbb{C}$, invariante por conjugación compleja. ¿Bajo qué condiciones podemos afirmar que hay una familia infinita de enteros algebraicos, cuyos conjugados están en $K$ y que se equidistribuye con respecto a alguna medida natural? No es difícil ver que un disco centrado en el origen y de radio menor que 1 contiene solo un conjunto de conjugados de un entero algebraico, el punto cero. Esto sugiere que el conjunto $K$ debiese tener "suficiente tamaño aritmético", en algún sentido a precisar. La noción adecuada de tamaño resulta ser la *capacidad* de $K$, concepto proveniente de la teoría del potencial.

En el contexto del Teorema de Bilu, es instructivo considerar la familia de ecuaciones $(x-1)^n = 0$, que tiene una única solución. Lo que hace diferente esta familia de $x^n = 1$ es el tamaño de los coeficientes del polinomio que determina la ecuación. La noción adecuada de tamaño es la *altura* de un número algebraico. Se demuestra que una sucesión de números algebraicos de altura que tiende a cero, tiene la propiedad que sus conjugados se equidistribuyen con respecto a la medida uniforme en el círculo unitario.

En la Sección 2 presentamos un resumen de algunas propiedades básicas de los números algebraicos que serán usadas en este curso. Luego en la Sección 3 presentamos las nociones de la teoría del potencial necesarias para formular y demostrar el Teorema de Fekete-Szegö. Por último, en la Sección 4 explicamos la noción de altura y la demostración del Teorema de Bilu.

## 2. Generalidades sobe números algebraicos

### 2.1. Números algebraicos.

**Definición 2.1.** Un elemento $\alpha \in \mathbb{C}$ se dice *número algebraico* si existe un polinomio no constante, con coeficientes racionales, $f(x) \in \mathbb{Q}[x]$, tal que $f(\alpha) = 0$. Un número algebraico se denomina *entero algebraico* si el polinomio $f$ se puede elegir mónico (es decir, el coeficiente del término dominante es 1) y con coeficientes enteros $f(x) \in \mathbb{Z}[x]$.

Notar que todo número racional es algebraico. En efecto, si $\alpha = \frac{a}{b}$, con $a, b \in \mathbb{Z}$ y $b \neq 0$, entonces podemos tomar $f(x) = x - \frac{a}{b}$. Sin embargo, sólo los números enteros son, a la vez, racionales y enteros algebraicos. Más ejemplos:

**Ejemplo 2.2.**

1. $\alpha = i, \quad f(x) = x^2 + 1.$
2. $\alpha = \sqrt{2}, \quad f(x) = x^2 - 2.$
3. $\alpha = \sqrt[3]{2}, \quad f(x) = x^3 - 2.$
4. $\alpha = \zeta_n := e^{2\pi i/n}, \quad f(x) = x^n - 1, \quad n \in \mathbb{Z}_{>0}.$
5. $\alpha = \frac{1+\sqrt{5}}{2}, \quad f(x) = x^2 - x - 1.$
6. $\alpha = \sqrt{2}/2, \quad f(x) = 2x^2 - 1.$

Todos los ejemplos anteriores, salvo el último, son enteros algebraicos.

Denotamos $\overline{\mathbb{Q}}$ al conjunto de los números algebraicos. Notar que $\overline{\mathbb{Q}}$ es un conjunto numerable. En efecto, $\mathbb{Q}[x]$ es numerable y cada elemento no constante de $\mathbb{Q}[x]$ tiene a lo más un número finito de raíces. Dado que $\mathbb{C}$ no es numerable, se desprende que existen números que no son algebraicos (y de hecho son mayoría). Sin embargo, no es fácil identificar un número no algebraico. Por ejemplo, se sabe que $e$ (Hermite 1873) y $\pi$ (Lindemann 1882) no son algebraicos, pero a la redacción de estas líneas no se sabe decidir si $e + \pi$ es algebraico o no.

El polinomio $f(x)$ que figura en la Definición 2.1 no es único. Cualquier polinomio de la forma $h(x) = f(x)g(x)$, con $g(x) \in \mathbb{Q}[x]$, sirve también. Sin embargo, nos será útil contar con un polinomio asociado de manera canónica a un número algebraico.

**Proposición 2.3.** *Sea $\alpha \in \overline{\mathbb{Q}}$. Entonces existe un único polinomio no constante $f_\alpha(x)$ que cumple*

1. *$f_\alpha(x)$ tiene coeficientes racionales,*
2. *$f_\alpha(\alpha) = 0$,*
3. *$f_\alpha(x)$ es mónico,*
4. *el grado de $f_\alpha(x)$ es mínimo entre los polinomios que satisfacen* (1), (2) *y* (3).

*Demostración.* el principio del buen orden nos asegura que existe un polinomio $f(x) \in \mathbb{Q}[x]$ que satisface las cuatro propiedades del enunciado. Si $g(x)$ es otro polinomio que cumple las mismas propiedades, entonces podemos aplicar división de polinomios

$$f(x) = q(x)g(x) + r(x), \quad q(x), r(x) \in \mathbb{Q}[x], \quad \deg r(x) < \deg g(x).$$

Como $r(\alpha) = f(\alpha) - q(\alpha)g(\alpha) = 0$, de la minimalidad del grado de $f(x)$ concluimos que $r(x)$ es el polinomio nulo, es decir $f(x) = q(x)g(x)$. Como $f$ y $g$ tienen el mismo grado y son mónicos, entonces $f = g$. □

**Definición 2.4.**

- Decimos que el polinomio $f_\alpha(x)$ dado por la Proposición 2.3 es el *polinomio mínimo* de $\alpha$.
- Definimos el *grado de* $\alpha$ por $\deg \alpha := \deg f_\alpha$.

**Definición 2.5.** Un polinomio $f(x) \in \mathbb{Q}[x]$ se dice $\mathbb{Q}$-reductible si se puede escribir como producto de polinomios, con coeficientes racionales, de grado estrictamente menor. Un polinomio en $\mathbb{Q}[x]$ que no es $\mathbb{Q}$-reductible se dice $\mathbb{Q}$-irreductible.

**Observación**: la minimalidad del grado de $f_\alpha(x)$ asegura que éste es un polinomio $\mathbb{Q}$-irreductible. Recíprocamente, se tiene el siguiente resultado.

**Proposición 2.6.** *Sea $\alpha \in \overline{\mathbb{Q}}$ y sea $f(x) \in \mathbb{Q}[x]$ un polinomio $\mathbb{Q}$-irreductible y mónico tal que $f(\alpha) = 0$. Entonces $f = f_\alpha$.*

*Demostración.* aplicar el algoritmo de la división para polinomios. □

**Definición 2.7.** Dado un polinomio

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \in \mathbb{C}[x],$$

definimos el *polinomio reverso*

$$f^*(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1}x + a_n \in \mathbb{C}[x].$$

**Lema 2.8.** *Si $f(x) \in \mathbb{Q}[x]$ es $\mathbb{Q}$-irreductible, entonces $f^*(x)$ también lo es.*

*Demostración.* basta notar que $f^*(x) = x^{\deg f} f(1/x)$ y usar la Definición 2.5. □

De la Proposición 2.6 y el Lema 2.8, se deduce

**Corolario 2.9.** *Si $\alpha \in \overline{\mathbb{Q}}$ y $\alpha \neq 0$, entonces $1/\alpha \in \overline{\mathbb{Q}}$. Más aún, $f_{1/\alpha} = f_\alpha^*$.*

No siempre es fácil determinar el polinomio mínimo de un número algebraico. En los ejemplos 2.2, los polinomios indicados son todos irreductibles (luego coinciden con el polinomio mínimo) excepto en el Ejemplo 2.2, (4). En efecto,

$$x^n - 1 = (x-1)(x^{n-1} + x^{n-2} + \cdots + x + 1).$$

Dos herramientas básicas para decidir la irreductibilidad de un polinomio con coeficientes racionales son el Lema de Gauss y el Criterio de Eisenstein, que procedemos a explicar.

**Definición 2.10.** Un polinomio con coeficientes enteros $p(x) \in \mathbb{Z}[x]$ se dice $\mathbb{Z}$-reductible si existen polinomios $f(x), g(x) \in \mathbb{Z}[x]$ tales que

- $f(x), g(x) \notin \{\pm 1\}$,
- $p(x) = f(x)g(x)$.

Diremos que $p(x) \in \mathbb{Z}[x]$ es $\mathbb{Z}$-irreductible si no es $\mathbb{Z}$-reductible.

**Teorema 2.11.** *(Lema de Gauss) Un polinomio $p(x) \in \mathbb{Z}[x]$ que es $\mathbb{Z}$-irreductible es también $\mathbb{Q}$-irreductible.*

**Teorema 2.12.** *(Criterio de Eisenstein) Sea $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ un polinomio con coeficientes enteros. Suponga que existe un número primo $p$ tal que*

- $p | a_i, \quad i = 0, 1, \ldots, n-1$,
- $p \nmid a_n$,
- $p^2 \nmid a_0$.

*Entonces $f(x)$ es $\mathbb{Q}$-irreductible.*

La demostración de estos teoremas está esbozada en los Ejercicios 2.15, (4) y (5).

Denotamos por $\Phi_n(x)$ al polinomio mínimo de $\zeta_n = e^{2\pi i/n}$.

*Observación* 2.13. Se tiene que $\Phi_n(x)$ es un divisor (en $\mathbb{Q}[x]$) del polinomio $x^n - 1$ (cf. Ejercicio 2.15 (1)). Del lema de Gauss, deducimos que $\Phi_n(x) \in \mathbb{Z}[x]$.

Como aplicación de los teoremas anteriores, demostraremos el siguiente

**Lema 2.14.** *Sea $p$ un número primo. Entonces $\Phi_p(x) = x^{p-1} + x^{p-2} + \ldots + x + 1$. En particular,*

$$x^p - 1 = (x-1)\Phi_p(x).$$

*Demostración.* Sea $w(x) = x^{p-1} + x^{p-2} + \ldots + x + 1$. Veamos que $w(x)$ es un polinomio $\mathbb{Q}$-irreductible. Esto último equivale a demostrar que $h(x) := w(x+1)$ es $\mathbb{Q}$-irreductible. Usando la identidad $x^p - 1 = (x-1)w(x)$, se tiene

$$h(x) = w(x+1) = \frac{(x+1)^p - 1}{x} = x^{p-1} + \Big(\sum_{k=2}^{p-1} \binom{p}{k} x^{k-1}\Big) + p.$$

Dado que $p|\binom{p}{k}$, para todo $0 < k < p$, el criterio de Eisenstein permite concluir que $h(x)$ es $\mathbb{Q}$-irreductible. Tenemos entonces que $w(x)$ es $\mathbb{Q}$-irreductible y mónico, luego $w(x) = \Phi_p(x)$ por la Proposición 2.6. □

*Ejercicios* 2.15.

1. Sea $f(x) \in \mathbb{Q}[x]$ un polinomio $\mathbb{Q}$-irreductible y sea $\alpha \in \mathbb{C}$ tal que $f(\alpha) = 0$. Sea $g(x) \in \mathbb{Q}[x]$ un polinomio no constante tal que $g(\alpha) = 0$. Demuestre que existe un polinomio $h(x) \in \mathbb{Q}[x]$ tal que $g(x) = f(x)h(x)$.
2. Un polinomio $p(x)$ se dice que tiene raíces repetidas si se puede factorizar sobre $\mathbb{C}$ de la forma
$$p(x) = (x-z)^2 h(x), \quad z \in \mathbb{C}, \quad h(x) \in \mathbb{C}[x]$$
(de manera equivalente, $p(z) = p'(z) = 0$). Demuestre que un polinomio $\mathbb{Q}$-irreductible $p(x) \in \mathbb{Q}[x]$ no puede tener raíces repetidas.
3. Sea $p$ un primo. Denotamos $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$, el cuerpo finito de $p$ elementos. Sea
$$\nu : \mathbb{Z} \longrightarrow \mathbb{Z}/p\mathbb{Z}$$
el morfismo canónico. Para un polinomio con coeficientes enteros $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$, definimos
$$\bar{f}(x) = \nu(a_n)x^n + \nu(a_{n-1})x^{n-1} + \ldots + \nu(a_1)x + \nu(a_0) \in \mathbb{F}_p[x].$$
Muestre que la operación $f \mapsto \bar{f}$ define un morfismo de anillos $\mathbb{Z}[x] \longrightarrow \mathbb{F}_p[x]$ (es decir, verifique $\overline{f+g} = \bar{f} + \bar{g}$ y $\overline{fg} = \bar{f}\bar{g}$).
4. *Lema de Gauss.* Decimos que un polinomio $f(x) \in \mathbb{Z}[x]$ es *reducido* si el máximo común divisor de sus coeficientes es 1.
   a) Sean $f(x), g(x) \in \mathbb{Z}[x]$ dos polinomios y sea $h(x) = f(x)g(x)$. Sea $N$ el máximo común divisor de los coeficientes de $h(x)$. Suponga $N \neq 1$ y tome un primo $p$ tal que $p|N$. Usando el morfismo de anillos del ejercicio anterior, demuestre que $p$ o bien divide a todos los coeficientes de $f(x)$ o bien divide a todos los coeficientes de $g(x)$
   b) Deduzca que el producto de dos polinomios reducidos es reducido
   c) Demuestre el Teorema 2.11.
5. *Criterio de Eisenstein.* Sean $f(x) \in \mathbb{Z}[x]$ un polinomio y $p$ un primo que satisfacen las hipótesis del Teorema 2.12.
   a) Suponga que se puede factorizar $f(x) = g(x)h(x)$ con $g(x), h(x) \in \mathbb{Z}[x]$. Muestre que entonces se tiene
   $$\nu(a_n)x^n = \bar{g}(x)\bar{h}(x), \quad \text{en } \mathbb{F}_p[x].$$
   b) Justifique que $\bar{g}(x) = ux^a, \bar{h} = vx^b$, con $u, v \in \mathbb{F}_p^*$ y $a + b = n$.
   c) Demuestre el Teorema 2.12.
6. Pruebe que $\overline{\mathbb{Q}}$ es un cuerpo. Es decir,
$$\alpha, \beta \in \overline{\mathbb{Q}}, \alpha \neq 0 \Rightarrow \frac{1}{\alpha}, \alpha\beta, \alpha + \beta \in \overline{\mathbb{Q}}.$$

## 2.2. Órbita galoisiana.

**Definición 2.16.**

- Sean $\alpha \in \overline{\mathbb{Q}}$ y $f_\alpha(x)$ su polinomio mínimo. Decimos que $\beta \in \mathbb{C}$ *es un conjugado de* $\alpha$ si $f_\alpha(\beta) = 0$.
- Definimos la *órbita galoisiana* de $\alpha$ por

$$\begin{aligned}\mathrm{Gal}(\alpha) :&= \{\beta \in \mathbb{C} : \beta \text{ es un conjugado de } \alpha\}\\ &= \{\beta \in \mathbb{C} : f_\alpha(\beta) = 0\}.\end{aligned}$$

*Observación* 2.17. El número de conjugados de $\alpha$ es exactamente el grado de $f_\alpha(x)$ (ver Ejercicio 2.15, (2)).

**Ejemplo 2.18.**

- $\mathrm{Gal}(i) = \{i, -i\}$.
- $\mathrm{Gal}(\sqrt{2}) = \{\sqrt{2}, -\sqrt{2}\}$.
- $\mathrm{Gal}(\sqrt[3]{2}) = \{\sqrt[3]{2}, \zeta_3\sqrt[3]{2}, \zeta_3^2\sqrt[3]{2}\}$.
- $\mathrm{Gal}(\frac{1+\sqrt{5}}{2}) = \{\frac{1+\sqrt{5}}{2}, \frac{1-\sqrt{5}}{2}\}$.

Es de interés para nosotros calcular la órbita galoisiana de $\zeta_n$. Definimos

$$\begin{aligned}\mu_n :&= \{z \in \mathbb{C} : z^n = 1\} = \{\text{raíces de la unidad de orden } n\},\\ \tilde{\mu}_n :&= \{z \in \mu_n : z^k \neq 1, \quad \forall 1 \leq k < n\}\\ &= \{\text{raíces primitivas de la unidad de orden } n\}.\end{aligned}$$

Notar que $(\mu_n, \cdot)$ es un grupo cíclico de orden $n$. Más precisamente, si $\zeta_n = e^{2\pi i/n}$, entonces $\mu_n = \{\zeta_n^j : 0 \leq j \leq n-1\}$. Más aún, se tiene

$$\tilde{\mu}_n = \{\zeta_n^j : \mathrm{mcd}(j, n) = 1\}. \tag{2.1}$$

En particular, el número de elementos de $\tilde{\mu}_n$ es

$$\varphi(n) := |(\mathbb{Z}/n\mathbb{Z})^*| = \#\{1 \leq j \leq n-1 : \mathrm{mcd}(j, n) = 1\}$$

(función de Euler).

**Proposición 2.19.** *Sea* $\zeta_n = e^{2\pi i n}$. *Entonces*

$$\mathrm{Gal}(\zeta_n) = \tilde{\mu}_n.$$

La demostración de este hecho está esbozada en los ejercicios 2.21.

**Corolario 2.20.** *Se tiene* $\deg \Phi_n = \varphi(n)$, *para todo entero positivo* $n$.

*Ejercicios* 2.21.

1. Demuestre (2.1).
2. *a*) Demuestre que para todo $\alpha \in \overline{\mathbb{Q}}$ y todo entero positivo $k$, se tiene
$$\mathrm{Gal}(\alpha^k) = \{\beta^k : \beta \in \mathrm{Gal}(\alpha)\}.$$
Deduzca $\deg(\alpha^k) \leq \deg(\alpha)$.
*b*) Muestre que para $\alpha \in \overline{\mathbb{Q}}^*$, se tiene $\mathrm{Gal}(1/\alpha) = \{1/\beta : \beta \in \mathrm{Gal}(\alpha)\}$. Deduzca $\deg(\alpha) = \deg(1/\alpha)$.
3. Sean $n$ un entero y $p$ un número primo que no divide $n$. Sea $\alpha \in \mathbb{C}$ tal que $\Phi_n(\alpha) = 0$. El objetivo de este problema es demostrar que $\Phi_n(\alpha^p) = 0$.
*a*) Muestre que $\alpha^p \in \overline{\mathbb{Q}}$ y que su polinomio mínimo tiene coeficientes enteros.
*b*) Suponga que $\Phi_n(\alpha^p) \neq 0$. Sea $g(x) \in \mathbb{Z}[x]$ el polinomio mínimo de $\alpha^p$. Sea $h(x) = g(x^p)$. Muestre que $\Phi_n(x) | h(x)$ en $\mathbb{Z}[x]$.

*c*) Sea $j(x) \in \mathbb{F}_p[x]$ un factor irreductible de $\overline{\Phi}_n(x) \in \mathbb{F}_p[x]$ (cf. la operación

$$f \in \mathbb{Z}[x] \mapsto \bar{f} \in \mathbb{F}_p[x]$$

definida en el Ejercicio 2.15 (3)). Muestre que $j(x)|\bar{g}(x)$.

*d*) Use lo anterior para mostrar que $j(x)^2|x^n - 1$. Concluya.

4. Use el ejercicio anterior para demostrar la Proposición 2.19.

**2.3. Discriminante y Resultante.** Sea $f(x) \in \mathbb{C}[x]$ un polinomio y lo expresamos en términos de sus raíces como

$$f(x) = a(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n), \quad a, \alpha_1, \ldots, \alpha_n \in \mathbb{C}.$$

Notar que las raíces $\alpha_1, \alpha_2, \ldots, \alpha_n$ no son necesariamente distintas. Se define el *discriminante* de $f$ por

$$D(f) = a^{2n-2} \prod_{i<j} (\alpha_i - \alpha_j)^2.$$

Es inmediato de la definición que $D(f) = 0$ si y sólo si $f$ tiene raíces repetidas.

**Definición 2.22.** Sean

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \in \mathbb{C}[x], \tag{2.2}$$

$$g(x) = b_m x^m + b_{m-1}x^{m-1} + \ldots + b_1 x + b_0 \in \mathbb{C}[x]. \tag{2.3}$$

Formamos la siguiente matriz de tamaño $(n+m) \times (n+m)$:

$$M(f,g) = \begin{pmatrix} a_n & a_{n-1} & \ldots & a_0 & 0 & \ldots & 0 \\ 0 & a_n & a_{n-1} & \ldots & a_0 & \ldots & 0 \\ & \ldots & & & & \ldots & \\ 0 & 0 & \ldots & a_n & a_{n-1} & \ldots & a_0 \\ b_m & b_{m-1} & \ldots & b_0 & & & \\ 0 & b_m & b_{m-1} & \ldots & b_0 & & \\ & \ldots & & & & & \\ 0 & 0 & \ldots & b_m & b_{m-1} & \ldots & b_0 \end{pmatrix}.$$

Definimos la *resultante* $R(f,g)$ por

$$R(f,g) := \det M(f,g).$$

**Teorema 2.23.** *([11], Proposition 8.3, p. 202) Escribimos $g(x) = b_m(x-\beta_1)\cdots(x-\beta_m)$. Entonces se tiene*

$$R(f,g) = a_n^m b_m^n \prod_{i=1}^{n}\prod_{j=1}^{m}(\alpha_i - \beta_j). \tag{2.4}$$

*Observación* 2.24. De este teorema vemos que $R(f,g) = 0$ si y sólo si $f$ y $g$ tienen una raíz en común.

Central en lo que sigue es la siguiente

**Proposición 2.25.** *Si $f(x)$ tiene coeficientes enteros y $a_n$ es el coeficiente del término dominante, entonces $a_n D(f) \in \mathbb{Z}$.*

*Demostración.* Tomemos $g = f'$ en (2.4). Entonces $m = n-1$ y $b_m = na_n$. Si escribimos

$$f'(x) = na_n(x-\beta_1)\cdots(x-\beta_{n-1}),$$

entonces

$$R(f,f') = n^n a_n^{2n-1} \prod_{i=1}^n \frac{f'(\alpha_i)}{na_n} = a_n^{n-1} \prod_{i=1}^n f'(\alpha_i).$$

Pero $f'(\alpha_i) = a_n \prod_{j\neq i}(\alpha_i - \alpha_j)$, luego obtenemos

$$\begin{aligned} R(f,f') &= a_n^{2n-1} \prod_{i=1}^n \prod_{\substack{j=1 \\ j\neq i}}^n (\alpha_i - \alpha_j) = (-1)^{n(n-1)/2} a_n^{2n-1} \prod_{i<j} (\alpha_i - \alpha_j)^2 \\ &= (-1)^{n(n-1)/2} a_n D(f). \end{aligned}$$

Como $f$ y $f'$ tienen coeficientes enteros, de la definición de resultante tenemos $R(f,f') \in \mathbb{Z}$ □

## 3. Teoría del potencial y el teorema de Fekete-Szegö

En esta sección vamos a ver los rudimentos de la teoría del potencial en el plano complejo y una primera aplicación aritmética. En concreto, estudiaremos las propiedades de la capacidad logarítmica y de sus otras presentaciones, el diámetro transfinito y la constante de Chebyshev. Como aplicación de este estudio daremos una demostración del teorema de Fekete-Szegö que afirma que la capacidad logarítmica discrimina cuándo un conjunto puede contener infinitas órbitas de Galois de enteros algebraicos.

La presentación está centrada en la comparación entre el diámetro transfinito, la capacidad logarítmica y la constante de Chebyshev. La mayor parte de esta sección está basada en las excelentes notas [16]. Para más detalles y demostraciones que se han omitido, el lector puede consultar [13]. Una presentación moderna del teorema de Fekete-Szegö utilizando teoría de Arakelov se puede encontrar en las notas [1].

**3.1. Puntos de Fekete.** Empezamos con un subconjunto $E \subset \mathbb{C}$ compacto, es decir $E$ es un subconjunto de $\mathbb{C}$ cerrado y acotado, por ejemplo el disco unidad. Dado un entero $n \geq 2$, queremos buscar un conjunto de $n$ puntos de $E$ que estén lo más alejados posible entre sí. Para dar una idea precisa de qué entendemos por estar lo más alejado posible, definimos la cantidad

$$\delta_n(E) = \max_{z_1,\ldots,z_n \in E} \left( \prod_{1\leq i<j\leq n} |z_i - z_j| \right)^{2/n(n-1)}$$

y buscamos colecciones de puntos donde se alcance esta cantidad. Cualquier colección de puntos $\mathcal{F}_n = \{z_1^{(n)}, \ldots, z_n^{(n)}\}$ con

$$\left( \prod_{1\leq i<j\leq n} |z_i^{(n)} - z_j^{(n)}| \right)^{2/n(n-1)} = \delta_n(E)$$

se llama un conjunto de Fekete (de $n$ puntos) y sus elementos puntos de Fekete. Por definición, estos puntos están lo más alejados entre sí dentro de $E$.

La primera observación que podemos hacer es que los puntos de Fekete se sitúan siempre en la frontera exterior del subconjunto $E$ (la frontera exterior es la frontera de la componente conexa no acotada del complementario de $E$). Por ejemplo, si $E$ es el anillo

$$A = \{z \in \mathbb{C} : r \leq z \leq R\},$$

los puntos de Fekete de $A$ se concentran en la circunferencia de radio $R$. Este resultado se puede demostrar usando el principio del máximo de las funciones analíticas.

**Ejemplo 3.1.** Como veremos en el Ejercicio 3.7 (2), si $E$ es el disco unidad, el conjunto de las raíces de la unidad $n$-ésimas es un conjunto de Fekete.

**Ejemplo 3.2.** En el conjunto $E = [-1, 1]$, para cada $n$, hay un único conjunto de Fekete de $n$ puntos. Y está dado por los ceros del polinomio $(x^2 - 1)P'_{n-1}(x)$, donde $P_{n-1}$ es el polinomio de Legendre de grado $n - 1$ y $P'_{n-1}$ su derivada.



FIGURA 1. Conjunto de Fekete de 11 puntos en el intervalo $[-1, 1]$.

Pero, ¿para qué sirven los puntos de Fekete?

Los puntos de Fekete son muy útiles en la interpolación de Lagrange. Sea $f$ una función continua definida en un intervalo compacto $E \subset \mathbb{R}$ y sean $z_0, \ldots, z_n$ un conjunto de $n+1$ puntos en $E$. El método de aproximación de Lagrange consiste en buscar un polinomio de grado $n$ que coincida con $f$ en los puntos elegidos. Siendo $f$ continua, es de esperar que el polinomio obtenido se parezca a la función $f$.

Para construir el polinomio de interpolación se definen los polinomios fundamentales de Lagrange

$$L_k(x) = \frac{\prod_{i \neq k}(x - z_i)}{\prod_{i \neq k}(z_k - z_i)}.$$

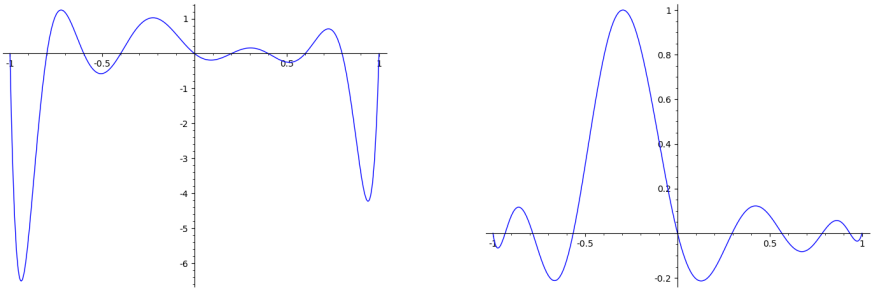
Estos polinomios cumplen $L_k(z_k) = 1$ y $L_k(z_i) = 0$ si $i \neq k$. En otras palabras $L_k(z_i) = \delta_{ik}$.

Una vez construidos los polinomios fundamentales de Lagrange, el polinomio de interpolación se escribe fácilmente

$$P(x) = \sum_{k=0}^{n} f(z_k) L_k(x).$$

Claramente, en los puntos de interpolación tenemos $P(z_k) = f(z_k)$, $k = 1, \ldots, n$. Veamos cómo funciona este método en la práctica. Vamos a interpolar funciones continuas en el intervalo $[-1, 1]$ usando puntos equiespaciados (este método se usa

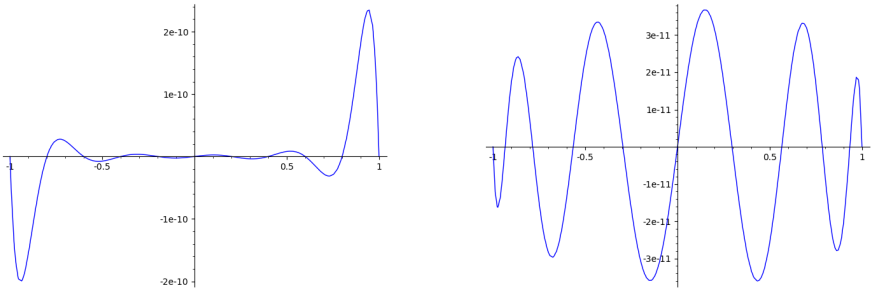
mucho por ser más cómodo de cálculo) y usando puntos de Fekete. En ambos casos usaremos 11 puntos.




(A) Puntos equiespaciados (B) Puntos de Fekete

FIGURA 2. Polinomios de Lagrange

En la Figura 2 podemos ver un polinomio fundamental de Lagrange para el caso de puntos equiespaciados y para el caso de puntos de Fekete. Obsérvese la diferencia de escala en ambos gráficos. En ambos casos hay un punto donde el polinomio toma el valor 1 y otros puntos donde el polinomio toma el valor 0. Sin embargo el polinomio de Lagrange obtenido con puntos equiespaciados oscila mucho al acercarnos al borde del intervalo, alcanzando un valor cercano a $-6$. Por el contrario, en el polinomio obtenido usando los puntos de Fekete, las oscilaciones están mucho más controladas y en ningún momento el valor absoluto del polinomio es mayor a 1. Parece claro que la mayor oscilación en el caso de puntos equiespaciados puede ser un problema. Vamos a verlo en acción. En primer lugar aproximaremos una función suave: $\exp(x)$. Esta función en el intervalo $[-1, 1]$ se aproxima muy bien, así que representamos en la Figura 3 el error cometido aproximando la función exponencial por un polinomio.




(A) Puntos equiespaciados (B) Puntos de Fekete

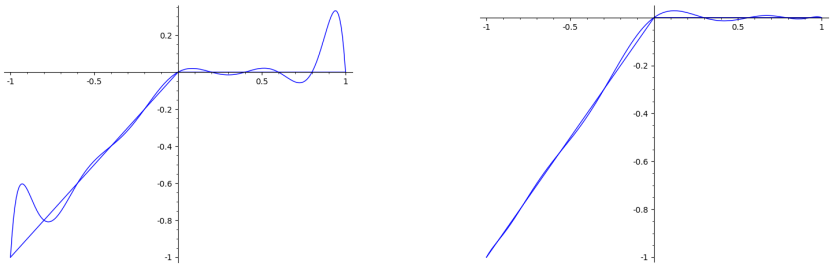
FIGURA 3. Error en la aproximación de la función exponencial.

Atención, la escala es diferente en ambos gráficos. Si bien el método de puntos equiespaciados proporciona mejores resultados en la zona central, el error máximo es casi 10 veces mayor en el caso de puntos equiespaciados respecto al caso de puntos de Fekete.

Podemos ponerle las cosas más difíciles al método de aproximación e intentar aproximar una función continua $f$ que no es diferenciable en el punto 0. Por ejemplo:

$$f(x) = \text{mín}(0, x).$$

En este caso el error se observa a simple vista y los resultados están en la Figura 4. De nuevo vemos que la presencia de oscilaciones incontroladas es un problema al



(A) Puntos equiespaciados (B) Puntos de Fekete

FIGURA 4. Aproximación de la función $f$.

usar puntos equiespaciados en el método de interpolación.

El Ejercicio 3.7 (3) explica por qué los puntos de Fekete son útiles en el método de aproximación. Para una función continua concreta puede que no dé el resultado óptimo. Pero da un resultado suficientemente bueno para todas las funciones continuas.

En muchos casos, calcular los puntos de Fekete de forma exacta es muy difícil y nos tenemos que conformar con unos puntos de Fekete *aproximados*, que hagan un trabajo razonable. Para saber si un conjunto de puntos se parece a un conjunto de Fekete es importante responder a las siguiente preguntas:

**Pregunta 3.3.** ¿Cuál es el comportamiento asintótico de $\delta_n(E)$ cuando $n$ tiende a infinito?

**Pregunta 3.4.** ¿Cuál es la distribución de los puntos de Fekete cuando $n$ tiende a infinito?

**Lema 3.5.** *La sucesión $\delta_n(E)$, $n \geq 2$, es decreciente.*

*Demostración.* Veamos que la sucesión $\log \delta_n(E)$, $n \geq 2$ es decreciente. Sea $\mathcal{F}_n = \{z_1, \ldots, z_n\} \subset E$ un conjunto de Fekete. Para cada $k$ entre 1 y $n$,

$$\begin{aligned}
\frac{n(n-1)}{2} \log \delta_n(E) &= \sum_{1 \leq i < j \leq n} \log |z_i - z_j| \\
&= \sum_{i \neq k} \log |z_i - z_k| + \sum_{\substack{1 \leq i < j \leq n \\ i \neq k \ j \neq k}} \log |z_i - z_j| \\
&\leq \sum_{i \neq k} \log |z_i - z_k| + \frac{(n-1)(n-2)}{2} \log \delta_{n-1}(E).
\end{aligned}$$

Sumando las $n$ desigualdades obtenidas variando $k$ deducimos

$$\frac{n^2(n-1)}{2} \log \delta_n(E) \leq n(n-1) \log \delta_n(E) + \frac{n(n-1)(n-2)}{2} \log \delta_{n-1}(E).$$

Por tanto $\log \delta_n(E) \leq \log \delta_{n-1}(E)$. □

Como consecuencia del Lema 3.5 la sucesión $\delta_n(E)$, $n \geq 2$ converge a un número real $\geq 0$.

**Definición 3.6.** El *diámetro transfinito* del conjunto $E$ es

$$\tau(E) = \lim_{n\to\infty} \delta_n(E).$$

Una primera respuesta a la Pregunta 3.3 es el cálculo del diámetro transfinito. Para responder a la Pregunta 3.4 usaremos el lenguaje de medidas e investigaremos el análogo continuo del problema de buscar puntos de Fekete.

*Ejercicios* 3.7.

1. La matriz de Vandermonde de tamaño $n \times n$ está dada por $(z_i^{j-1})$, para $1 \leq i \leq n$ y $1 \leq j \leq n$. Demostrar que el determinante de la matriz de Vandermonde es $\prod_{1\leq i<j\leq n}(z_j - z_i)$. Por tanto, un conjunto de puntos de Fekete maximiza el valor absoluto del determinante de la correspondiente matriz de Vandermonde.
2. Utilizando el ejercicio anterior y la desigualdad de Hadamard para determinantes, demostrar que las raíces de la unidad forman un conjunto de Fekete en el círculo unidad.
3. Sea $E$ un subconjunto compacto de $\mathbb{C}$ y $\mathcal{F}_{n+1} = \{z_1, \ldots, z_{n+1}\}$ un conjunto de Fekete de $n+1$ puntos para $E$.
   *a*) Demostrar que, si $L_k$ es el polinomio fundamental de Lagrange $k$-ésimo asociado a $\mathcal{F}_{n+1}$, entonces
   $$|L_k(z)| \leq 1, \text{ para todo } z \in E.$$
   *b*) Si, para un subconjunto $S \subset \mathbb{C}$ y una función $f$ denotamos
   $$\|f\|_S = \sup_{z\in S}\{|f(z)|\},$$
   demostrar que todo polinomio $P$ de grado menor o igual que $n$ satisface
   $$\|P\|_E \leq (n+1)\|P\|_{n+1}.$$
   *c*) Sea $f\colon E \to \mathbb{C}$ una función continua, sea $P_{n,f}$ el polinomio de grado menor o igual a $n$ que aproxima mejor a $f$ en la norma $\|\cdot\|_E$. Sea $P_{\mathcal{F}_{n+1}}$ el polinomio obtenido mediante interpolación de Lagrange en los puntos $\mathcal{F}_{n+1}$. Demostrar que
   $$\|f - P_{\mathcal{F}_{n+1}}\|_E \leq (n+2)\|f - P_{n,f}\|_E.$$

**3.2. Medidas y convergencia débil.** Para describir de forma cuantitativa la forma a la que tienden los conjuntos de Fekete o las órbitas de Galois de enteros algebraicos, es necesario introducir el lenguaje de medidas. En estas notas no podemos dar un curso completo de teoría de la medida y usaremos muchos resultados como una caja negra. El lector interesado puede consultar, por ejemplo [18] o [4].

Sea $D \subset \mathbb{R}^m$ un subconjunto con la topología inducida y sea $C_c^0(D, \mathbb{R})$ el espacio de las funciones continuas con soporte compacto en $D$. El espacio $C_c^0(D, \mathbb{R})$ tiene una norma, la norma del supremo, definida, para $f \in C_c^0(D, \mathbb{R})$, por

$$\|f\|_D = \sup_{z\in D} |f(z)|.$$

Una *medida de Radon* en $D$ es un funcional lineal continuo en $C_c^0(D,\mathbb{R})$. Las medidas de Radon son compatibles con la topología de $D$ y la mayor parte de las medidas que aparecen en análisis son de este tipo. Si $\mu$ es una medida en $D$ y $f \in C_c^0(D,\mathbb{R})$ es una función continua con soporte compacto, usamos tanto la notación funcional $\mu(f)$ como la notación integral

$$\mu(f) = \int_D f\, d\mu = \int_D f(x)\, d\mu(x).$$

Una medida se denomina *positiva* si, para toda $f \in C_c^0(D,\mathbb{R})$ con $f(z) \geq 0$, $\forall z \in D$, se tiene $\mu(f) \geq 0$.

Si $D$ es compacto y $\mu$ es una medida, se denomina la masa de $\mu$ a la cantidad

$$\mu(D) \coloneqq \int_D 1\, d\mu(z).$$

Si $D$ no es compacto, escribimos $D = \bigcup E_n$ con $E_1 \subset E_2 \subset \ldots$ compactos, y la medida de $D$ se define como el límite

$$\mu(D) = \lim \mu(E_n)$$

en caso de que este exista. Si la medida es positiva, el límite siempre existe aunque puede ser infinito. Una medida positiva se llama *finita* si la masa total es finita.

Una *medida de probabilidad* es una medida de Radon positiva de masa total 1. El espacio de medidas de probabilidad de $D$ se denota $\mathcal{M}(D)$.

**Ejemplo 3.8.** Si $z \in D$ la asignación $f \mapsto f(z)$ es una medida de probabilidad en $D$ que se denomina delta de Dirac en $z$ y se denota $\delta_z$:

$$f(z) = \int_X f(x)\, d\delta_z(x).$$

Más generalmente, si $F = \{z_1, \ldots, z_n\}$ es un conjunto finito, la *medida de contaje* de $F$ es la medida de probabilidad

$$\nu(F) \coloneqq \frac{1}{n} \sum_{i=1}^{n} \delta_{z_i}. \tag{3.1}$$

En otras palabras

$$\nu(F)(f) = \int_X f\, d\nu(F) = \frac{1}{n} \sum_{i=1}^{n} f(z_i),$$

es la media de los valores de $f$ en el conjunto $F$.

La teoría de medidas proporciona un lenguaje conveniente para estudiar la distribución asintótica de conjuntos de puntos mediante el concepto de convergencia débil-$*$.

Decimos que una sucesión de medidas $(\mu_n)_{n\geq 1}$ converge en la topología débil-$*$ a una medida $\mu$ si para toda función $f \in C_c^0(D,\overline{\mathbb{R}})$ se cumple

$$\int_D f\, d\mu = \lim_{n\to\infty} \int_D f\, d\mu_n.$$

El hecho de que la sucesión $(\mu_n)_{n\geq 1}$ converja en la topología débil-$*$ a $\mu$ se denota $\mu_n \xrightarrow{*} \mu$.

El siguiente resultado es consecuencia del Teorema de Banach-Alaoglu (ver [14] 3.15 y 3.16)

**Teorema 3.9.** *Sea $E \subset \mathbb{C}$ un subconjunto compacto. Entonces el espacio $\mathcal{M}(E)$ con la topología de la convergencia débil-$*$ es un espacio metrizable compacto. En particular es secuencialmente compacto.*

**Ejemplo 3.10.** Sea $F_n = \{e^{\frac{2\pi i k}{n}} \mid 0 \leq k \leq n-1\}$ el conjunto de raíces $n$-ésimas de 1. Para toda función $f \in C_c^0(\mathbb{C}, \mathbb{R})$ se tiene

$$\lim_{n\to\infty} \int_{\mathbb{C}} f \, d\nu(F_n) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{\theta i}) \, d\theta.$$

Por tanto la sucesión $\nu(F_n)$, $n \geq 1$ converge en la topología débil-$*$ a la medida de Haar normalizada en el círculo unidad $\frac{d\theta}{2\pi}$.

Dada una medida $\mu$ en $D$ y $U \subset D$ un subconjunto, decimos que el *soporte* de $\mu$ está contenido en $U$ si para toda función $f \in C_c^0(D, \mathbb{R})$, la condición $f|_U = 0$ implica $\mu(f) = 0$.

**Lema 3.11.** *Sea $E \subset C$ un subconjunto compacto y $B_n$, $n \geq 0$ subconjuntos compactos con $B_n \supset B_{n+1}$ y $E = \bigcap B_n$. Si $(\mu_n)_{n\geq 0}$ es una sucesión de medidas de probabilidad con $\mu_n \xrightarrow{*} \mu$ y* soporte$(\mu_n) \subset B_n$*, entonces* soporte$(\mu) \subset E$.

Necesitaremos un resultado técnico para extender determinadas medidas a funciones no necesariamente continuas.

**Definición 3.12.** Sea $D \subset \mathbb{R}^m$ un subconjunto con la topología inducida. Una función $f\colon D \to \mathbb{R} \cup \{\infty\}$ se denomina semi-continua inferior si para todo $z_0 \in D$ se cumple

$$f(z_0) \leq \liminf_{z\to z_0} f(z).$$

**Lema 3.13.** *Una función es semi-continua inferior si y sólo si, para todo compacto $K \subset D$ existe una sucesión creciente de funciones continuas que converge puntualmente a $f$.*

Si $f$ es una función semi-continua inferior y $\mu$ es una medida positiva con soporte compacto $K$, se define

$$\int_D f(x) \, d\mu(x) = \lim_{n\to\infty} f_n(x) \, d\mu(x),$$

para cualquier sucesión creciente de funciones continuas que converge puntualmente a $f$ en $K$.

**Lema 3.14.** *Sea $f$ una función semi-continua inferior y $\mu_n \xrightarrow{*} \mu$ una sucesión de medidas positivas, con soporte contenido en un compacto $K$, que converge, en la topología débil-$*$ a una medida $\mu$. Entonces*

$$\int_K f \, d\mu \leq \liminf_{n\to\infty} \int_K f \, d\mu_n.$$

*Ejercicios* 3.15.

1. Demostrar la afirmación del Ejemplo 3.10.
2. Demostrar el Lema 3.11.
3. Buscar una demostración del Lema 3.13.

**3.3. La capacidad logarítmica.** Discutimos ahora el análogo continuo del problema de encontrar los puntos de Fekete. Para ello imaginamos que los $n$-puntos representan cargas eléctricas del mismo signo que se repelen entre sí. De esta forma la configuración de puntos de Fekete es una posición de equilibrio para el problema electrostático.

Sea $E \subset \mathbb{C}$ de nuevo un subconjunto compacto. Podemos imaginar que $E$ es una lámina metálica a la que añadimos una carga eléctrica. Al ser una lámina metálica la carga se puede mover libremente y, dado que cargas de igual signo se repelen, tenderá a extenderse lo más posible.

Hay que tener cuidado con el símil electromagnético pues estamos acostumbrados al caso tridimensional, donde la fuerza de atracción/repulsión entre dos cargas eléctricas es proporcional al inverso de la distancia al cuadrado, mientras que la energía de dichas cargas es proporcional al inverso de la distancia. Si estamos confinados en un espacio bidimensional (como es el caso que nos ocupa) la fuerza eléctrica debe ser proporcional al inverso de la distancia, mientras que la energía debe ser proporcional al logaritmo de la distancia.

La densidad de carga se puede representar mediante una medida de probabilidad $\mu$ en $\mathbb{C}$ cuyo soporte está contenido en $E$. Por ejemplo, si toda la carga está concentrada en un punto $z$, la representamos mediante la medida delta de Dirac $\delta_z$.

**Definición 3.16.** Sea $\mu$ una densidad de carga en $E$, el *potencial electrostático* asociado a $\mu$ es

$$U^{\mu}(z) = \int_E \log \frac{1}{|z-t|} d\mu(t),$$

mientras que la *energía electrostática* de esta carga en $E$ es

$$I(\mu) = \int U^{\mu}(z) d\mu(z) = \int \int \log \frac{1}{|z-t|} d\mu(z) d\mu(t). \tag{3.2}$$

Hay que tener cuidado con esta definición pues, en caso de haber cargas puntuales, la energía es infinita. Para tener una cantidad finita con la que trabajar, introducimos la siguiente variante de la energía. Si $\mu = \nu(F)$ es la medida de contaje de un conjunto finito $F = \{z_1, \ldots, z_n\}$, ponemos

$$\begin{aligned} I'(F) = I'(\mu) &= \frac{1}{n(n-1)} \sum_{1 \le i \ne j \le n} \log \frac{1}{|z_i - z_j|} \\ &= -\log \left( \prod_{i<j} |z_i - z_j| \right)^{2/n(n-1)}. \end{aligned} \tag{3.3}$$

El problema de encontrar un conjunto de Fekete de $n$ puntos consiste en hallar

$$\inf_{\#F=n} I'(F).$$

**Definición 3.17.** La *constante de Robin* de un conjunto compacto $E \subset \mathbb{C}$ es

$$V_E = \inf_{\mu \in \mathcal{M}(E)} \{I(\mu)\}.$$

Por tanto, el análogo continuo de buscar las cantidades $\delta_n(E)$ y los puntos de Fekete es determinar el valor de la constante de Robin y encontrar una medida cuya energía coincida con este mínimo.

**Teorema 3.18.** *El valor $V_E$ se alcanza en una medida $\mu_E$. Es más, si $V_E < \infty$ entonces esta medida es única.*

La demostración de este resultado se basa en el análogo continuo del Lema 3.5.

**Lema 3.19.** *Sea $\mu_n$ una sucesión de medidas en $\mathcal{M}(E)$ que converge en la topología débil-$*$ a una medida $\mu$. Entonces, para todo $z \in \mathbb{C}$*

$$U^\mu(z) \leq \liminf_{n\to\infty} U^{\mu_n}(z). \tag{3.4}$$

*Además*

$$I(\mu) \leq \liminf_{n\to\infty} I(\mu_n). \tag{3.5}$$

*Demostración.* Este resultado es consecuencia de que $\log(1/|z-t|)$ es semi-continua inferior respecto de $t$. □

*Argumento de la demostración del Teorema 3.18.* Sean $\mu_n$ una sucesión de medidas en $\mathcal{M}(E)$ tales que $\lim_{n\to\infty} I(\mu_n) = V_E$. Por el Teorema 3.9 el espacio $\mathcal{M}(E)$ con la topología débil-$*$ es secuencialmente compacto. Así existe una subsucesión de $\mu_n$ que converge en la topología débil-$*$ a una medida $\mu$. Por el Lema 3.19 esta medida cumple $V_E = I(\mu)$.

Si $V_E = \infty$, cualquier medida $\mu$ con soporte en $E$ cumple $I(\mu) = V_E$ y minimiza la energía, por lo que no hay unicidad. Por el contrario si $V_E < \infty$, usando que la función $I$ en $\mathcal{M}(E)$ es estrictamente convexa (un punto técnico que hay que precisar pues estamos en un espacio de dimensión infinita) , se deduce la unicidad. □

**Definición 3.20.** La cantidad

$$\operatorname{Cap}(E) = e^{-V_E}$$

se denomina la *capacidad logarítmica* de $E$. Una medida $\mu_E$ que cumpla $V(E) = I(\mu_E)$ se denomina una *medida de equilibrio* para $E$ y el potencial $U^{\mu_E}$ asociado a la medida de equilibrio se denomina el *potencial de equilibrio* de $E$.

Enunciamos en el siguiente resultado unas propiedades básicas de la capacidad logarítmica.

**Teorema 3.21.**

1. *Si $E_1 \subset E_2$ entonces $\operatorname{Cap}(E_1) \leq \operatorname{Cap}(E_2)$.*
2. *Si $E \subset \mathbb{C}$ y $\alpha, \beta \in \mathbb{C}$, entonces $\operatorname{Cap}(\alpha E + \beta) = |\alpha| \operatorname{Cap}(E)$.*
3. *Si $K \subset \mathbb{C}$ es compacto, entonces $\operatorname{Cap}(K) = \operatorname{Cap}(\partial K)$.*
4. *Si $K_1 \supset K_2 \supset \dots$ son compactos y $K = \bigcap K_i$, entonces*
$$\operatorname{Cap}(K) = \lim_{i\to\infty} \operatorname{Cap}(K_i).$$
5. *Si $\operatorname{Cap}(K) = 0$ entonces $K$ tiene medida de Lebesgue cero. Más aún, si $\nu$ es una medida que satisface $I(\nu) < \infty$, entonces $\nu(K) = 0$.*

El problema discreto de encontrar los puntos de Fekete y el problema continuo de encontrar la medida de equilibrio están muy relacionados y la teoría electrostática del plano complejo nos permite dar respuesta a las preguntas 3.3 y 3.4.

**Teorema 3.22.** *Sea $E \subset \mathbb{C}$ un conjunto compacto. Entonces*

$$\operatorname{Cap}(E) = \tau(E). \tag{3.6}$$

*Además, si* $\mathrm{Cap}(E) > 0$ *y* $\mathcal{F}_n$, $n \geq 2$ *son colecciones de puntos de Fekete para* $E$ *entonces*

$$\nu(\mathcal{F}_n) \xrightarrow{*} \mu_E. \tag{3.7}$$

*Esta última propiedad se expresa diciendo que los punto de Fekete se* equidistribuyen *según la medida de equilibrio de* $E$.

*Demostración.* En el argumento de esta demostración vamos a necesitar la medida producto de dos medidas. Más concretamente, si $\mu$ y $\nu$ son medidas en $E$ entonces, la medida producto es una medida $\mu \times \nu$ en $E \times E$. Para la definición y las propiedades básicas de la medida producto el lector puede consultar la Sección 1.7 de [18]. La propiedad fundamental de la medida producto es el Teorema de Fubini-Tonelli [18, Corollary 1.7.23].

Empezamos demostrando la desigualdad $V_E \geq \log(1/\tau(E))$. Escribimos

$$F(z_1, \ldots, z_n) \coloneqq I'(\{z_1, \ldots, z_n\}) = \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \log \frac{1}{|z_i - z_j|}.$$

El mínimo de la función $F$ es $\log(1/\delta_n(E))$. El valor esperado de $F$ respecto a cualquier medida de probabilidad en el producto $\mathbb{C}^n$ debe ser mayor o igual a este mínimo. En particular, si consideramos la medida producto $d\mu_E(z_1) \ldots d\mu_E(z_n)$, donde $\mu_E$ es la medida de equilibrio de $E$, deducimos

$$V_E = \int_{E \times \cdots \times E} F(z_1, \ldots, z_n)\, d\mu_E(z_1) \ldots d\mu_E(z_n) \geq \log \frac{1}{\delta_n(E)}.$$

Consideramos la sucesión de medidas $\nu_n \coloneqq \nu(\mathcal{F}_n)$, $n \geq 2$. Por el Teorema 3.9 existe una medida $\hat{\mu}$ que es el límite en la topología débil-$*$ de una subsucesión $(\nu_{n_k})_{k \in \mathbb{N}}$. Una consecuencia del Teorema de Fubini-Tonelli para medidas es que $(\nu_{n_k} \times \nu_{n_k})_{k \in \mathbb{N}} \xrightarrow{*} \hat{\mu} \times \hat{\mu}$.

Para poder aplicar la convergencia débil de medidas necesitamos aplicarlas a funciones continuas. La función $\log|x|$ no es continua cuando $x$ tiende a cero.

Para cada $M \in \mathbb{R}$, $M > 0$, escribimos $\log_M(|x|) = \min(\log(|x|), M)$. Las funciones $\log_M(1/|x|)$ son continuas para todo $x \in \mathbb{R}$. Y cuando $M$ tiende a infinito, la colección de funciones $(\log_M(1/|x|))_M$ converge monótonamente a la función $\log(1/|x|)$. Calculamos

$$\begin{aligned}
I(\hat{\mu}) &= \int\int \log \frac{1}{|z-t|}\, d\hat{\mu}(z) d\hat{\mu}(t) \\
&= \lim_{M \to \infty} \int\int \log_M \frac{1}{|z-t|}\, d\hat{\mu}(z) d\hat{\mu}(t) \\
&= \lim_{M \to \infty} \lim_{k \in \mathbb{N}} \int\int \log_M \frac{1}{|z-t|}\, d\hat{\nu}_{n_k}(z) d\hat{\nu}_{n_k}(t) \\
&\leq \lim_{M \to \infty} \lim_{k \in \mathbb{N}} \left( \frac{n_k - 1}{n_k} \log \frac{1}{\delta_{n_k}(E)} + \frac{M}{n_k} \right) \\
&= \lim_{M \to \infty} \lim_{k \in \mathbb{N}} \log \frac{1}{\delta_{n_k}(E)} = \log \frac{1}{\tau(E)}.
\end{aligned}$$

La igualdad en la segunda línea es consecuencia del teorema de convergencia dominada, mientras que la igualdad en la tercera línea se sigue de la convergencia

débil de medidas junto con el hecho de que $\log_M(1/|x|)$ es continua en cualquier compacto de $\mathbb{C}$. Por la minimalidad de $V_E$ deducimos

$$V_E \leq I(\hat{\mu}) \leq \log \frac{1}{\tau(E)} \leq V_E$$

lo que demuestra (3.6). Si $\mathrm{Cap}(E) > 0$, la unicidad de la medida de equilibrio implica $\hat{\mu} = \mu_E$ de donde se deduce (3.7). □

**Ejemplo 3.23.** En el Ejercicio 3.26 1 se busca la capacidad y la medida de equilibrio en el caso del disco unidad.

**Ejemplo 3.24.** Si $E \subset \mathbb{R} \subset \mathbb{C}$ es el segmento $E = [a, b]$, entonces $\mathrm{Cap}(E) = \frac{b-a}{4}$ y la medida de equilibrio es

$$\mu_E = \frac{1}{\pi} \frac{dx}{\sqrt{(x-a)(b-x)}}. \tag{3.8}$$

En el Ejercicio 3.26 (2) vemos que el potencial de equilibrio es constante en el disco unidad. Esto es consistente con la intuición física, pues si hubiera diferencia de potencial, las cargas se desplazarían para disminuir la energía total. En general el resultado matemático preciso es

**Teorema 3.25** (Teorema de Frostman)**.** *Sea $E \subset \mathbb{C}$ un subconjunto compacto con* $\mathrm{Cap}(E) > 0$. *Entonces*

1. *$U^{\mu_E}(z) \leq V_E$ para todo $z \in \mathbb{C}$,*
2. *$U^{\mu_E}(z) = V_E$ para todo $z \in E \setminus S$, donde $S$ es un subconjunto con* $\mathrm{Cap}(S) = 0$.
3. *Si $\Omega$ es la componente conexa no acotada de $\mathbb{C} \setminus E$, entonces para todo $z \in \Omega$ se tiene $U^{\mu_E}(z) < V_E$.*

La demostración de este resultado usa el principio del máximo para potenciales 3.35. Esta demostración la daremos en la Sección 3.6.

*Ejercicios* 3.26.

1. Usando que las raíces de la unidad de orden $n$ forman un conjunto de Fekete de $n$ puntos en el disco unidad, demostrar que la capacidad del disco unidad es 1 y que la medida de equilibrio del disco unidad es la medida de Haar $\frac{d\theta}{2\pi}$ en el círculo unidad.
2. Demostrar que el potencial de equilibrio del disco unidad es

$$U^{\mu_E}(z) = \begin{cases} 0, & \text{if } |z| \leq 1 \\ \log(1/|z|), & \text{if } |z| > 1. \end{cases}$$

**3.4. La constante de Chebyshev.** En esta sección, $E \subset \mathbb{C}$ continúa designando un subconjunto compacto. El siguiente problema es determinar el mínimo valor de la norma $\|p\|_E$ donde $p$ es un polinomio mónico de grado $n$ y encontrar un polinomio que alcance este mínimo. Es decir, buscamos el valor de

$$t_n(E) \coloneqq \inf_{p \in \mathcal{P}_n} \|p\|_E, \tag{3.9}$$

donde $\mathcal{P}_n$ es el conjunto de los polinomios mónicos de grado $n$. Si $E$ contiene infinitos puntos existe un único polinomio mónico $T_n \in \mathcal{P}_n$ con $t_n(E) = \|T_n\|_E$, que se llama el polinomio de Chebyshev $n$-ésimo para $E$. Ver el ejercicio 1.

Debido a que

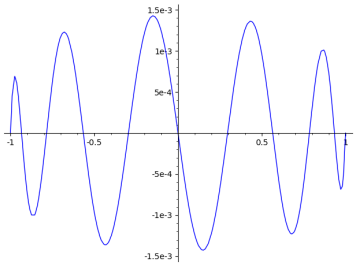
$$t_{m+n}(E) = \|T_{m+n}\|_E \leq \|T_m\|_E \|T_n\|_E = t_m(E) t_n(E),$$

la sucesión $\log(t_n(E))$ es subaditiva y por tanto, el límite

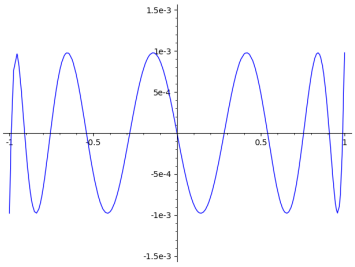
$$\mathrm{Cheb}(E) = \lim_{n\to\infty} t_n(E)^{1/n} = \inf_{n\geq 1} t_n(E)^{1/n}$$

existe y se denomina la constante de Chebyshev de $E$.

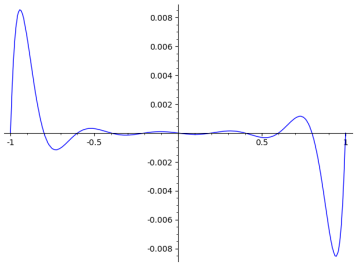
Muy relacionado con el polinomio de Chebyshev, el polinomio de Fekete $F_n$ es un polinomio mónico de grado $n$ cuyas raíces forman un conjunto de Fekete de $n$ puntos.




(A) Polinomio de Fekete de grado 11.




(B) Polinomio de Chebyshev de grado 11.




(C) Polinomio de grado 11 con raíces equiespaciadas.

FIGURA 5. Polinomios de Fekete y Chebyshev.

En la Figura 5 se compara el polinomio de Chebyshev de grado 11 con el polinomio de Fekete de grado 11 en el intervalo $[-1, 1]$. Ambos son razonablemente similares. Se incluye también el polinomio mónico de grado 11 con las raíces equiespaciadas. Observad la diferencia de escala. La norma del supremo de este último polinomio es 8 veces mayor que la del correspondiente polinomio de Chebyshev.

El resultado fundamental de la teoría del potencial es el siguiente.

**Teorema 3.27.** *Sea $E \subset \mathbb{C}$ un subconjunto compacto.*

1. $\tau(E) = \mathrm{Cap}(E) = \mathrm{Cheb}(E)$.
2. *Los polinomios de Fekete son asintóticamente óptimos para el problema de Chebyshev:*

$$\lim_{n\to\infty} \|F_n\|_E^{1/n} = \mathrm{Cheb}(E).$$

*Si además,* $\mathrm{Cap}(E) > 0$ *y por tanto $\mu_E$ es única, entonces*

3. *Los puntos de Fekete tienen distribución asintótica $\mu_E$. Es decir $\nu(\mathcal{F}_n) \xrightarrow{*} \mu_E$ cuando $n \to \infty$.*

*4. La convergencia*

$$\lim_{n\to\infty} |F_n(z)|^{1/n} = \exp(-U^{\mu_E}(z))$$

*es uniforme en cada subconjunto compacto de $\Omega$, la componente conexa no acotada de $\mathbb{C}\setminus E$.*

*Demostración.* La afirmación (3) ya ha sido establecida en el Teorema 3.22. La afirmación (4) se deduce de ella. De hecho, de la definición de potencial se deduce que

$$U^{\nu(\mathcal{F}_n)}(z) = \frac{1}{n}\log\frac{1}{|F_n(z)|}$$

y todos los ceros de $F_n$ están contenidos en $E$. Por tanto, si $K\subset\Omega$ es un subconjunto compacto y $z\in K$, dado que $\nu(\mathcal{F}_n)\xrightarrow{*}\mu_E$ tenemos que

$$\frac{1}{n}\log\frac{1}{|F_n(z)|} = \int \log\frac{1}{|t-z|}\,d\nu(\mathcal{F}_n)(t) \xrightarrow{n\to\infty} \int \log\frac{1}{|t-z|}\,d\mu_E(t) = U^{\mu_E}(z),$$

debido a que, para $z\notin E$, la función $t\mapsto -\log|t-z|$ es continua en $E$. Como $K$ es compacto y para $t\in E$, la función $\frac{1}{|t-z|}$ es continua en $K$, esta está uniformemente acotada en $K$. Por tanto la convergencia es uniforme en $K$.

Nos queda por demostrar (1) y (2), de los que ya hemos demostrado $\mathrm{Cap}(E) = \tau(E)$. Sea $\mathcal{F}_n = \{z_1^{(n)},\dots,z_n^{(n)}\}$ un conjunto de Fekete de $n$ puntos. Entonces

$$\delta_{n+1}(E)^{n(n+1)/2} = \max_{\{z_i\}\subset E}\prod_{1\le i<j\le n+1}|z_i-z_j| \ge \left(\prod_{1\le i\le n}|z-z_i^{(n)}|\right)\delta_n(E)^{(n-1)n/2}$$

para todo $z\in E$. Por tanto, poniendo $\delta_n = \delta_n(E)$,

$$\delta_{n+1}^{n(n+1)/2}/\delta_n^{(n-1)n/2} \ge |F_n(z)|, \quad \forall z\in E.$$

Tomando las raíces $n$-ésimas

$$\delta_{n+1}^{(n+1)/2}/\delta_n^{(n-1)/2} \ge \|F_n\|_E^{1/n}.$$

Como la sucesión $\delta_n$ es decreciente,

$$\delta_{n+1}^{(n+1)/2}/\delta_n^{(n-1)/2} \le (\delta_{n+1}\delta_n)^{1/2}.$$

Usando la definición de $\mathrm{Cheb}(E)$ deducimos

$$(\delta_{n+1}\delta_n)^{1/2} \ge \|F_n\|_E^{1/n} \ge \mathrm{Cheb}(E).$$

Haciendo tender $n$ a $\infty$ obtenemos

$$\tau(E) \ge \limsup_{n\to\infty}\|F_n\|_E^{1/n} \ge \liminf_{n\to\infty}\|F_n\|_E^{1/n} \ge \mathrm{Cheb}(E).$$

Por lo que para deducir (1) y (2), basta demostrar que $\mathrm{Cheb}(E)\ge\tau(E)$. Si $\tau(E)=0$ esta desigualdad es obvia, por lo que podemos suponer $\mathrm{Cap}(E)>0$. Sea $T_n(z)$ un polinomio de Chebyshev de grado $n$ en $E$ y denotamos mediante $\nu(T_n)$ la medida de contaje en los ceros de $T_n(z)$. Entonces

$$\frac{1}{n}\log\frac{1}{t_n(E)} = \inf_{z\in E}\frac{1}{n}\log\frac{1}{|T_n(z)|} = \inf_{z\in E} U^{\nu(T_n)}(z).$$

Ahora bien,

$$\inf_{z\in E} U^{\nu(T_n)}(z) \le \int U^{\nu(T_n)}\,d\mu_E = \int U^{\mu_E}\,d\nu(T_n),$$

donde la última igualdad es consecuencia del teorema de Fubini-Tonelli. Por el Teorema de Frostman 3.25 tenemos que

$$\int U^{\mu_E}\, d\nu(T_n) \leq V_E.$$

Juntando las desigualdades obtenidas, deducimos que

$$\frac{1}{n} \log \frac{1}{t_n(E)} \leq V_E,$$

y tomando el límite cuando $n$ tiende a infinito,

$$\mathrm{Cheb}(E) \geq \mathrm{Cap}(E) = \tau(E).$$

□

*Ejercicios* 3.28.

1. Sea $E \subset \mathbb{C}$ u subconjunto compacto que contiene al menos $n+1$ puntos.
   *a*) Demostrar que el ínfimo en la ecuación (3.9) se alcanza en algún polinomio $T_n \in \mathcal{P}_n$.
   *b*) Demostrar que hay al menos $n+1$ puntos de $E$ donde se cumple la igualdad $t_n(E) = |T_n(z)|$. (Indicación: si no fuera cierto, podríamos encontrar un polinomio $q$ de grado menor o igual a $n-1$ que coincide con $T_n$ en todo punto $z$ con $|T_n(z)| = t_n(E)$. Demostrar que, para $\epsilon > 0$ suficientemente pequeño $\|T_n - \epsilon q\| < t_n(E)$ contradiciendo la definición de $t_n(E)$.)
   *c*) Demostrar que $T_n$ es único. (Indicación: si hubiera dos candidatos, aplicar 1*b* a la media aritmética de ambos).
   *d*) Demostrar que los ceros de $T_n$ están en la envolvente convexa de $E$.
   *e*) Sea $E = [-1,1] \times \{-i, i\}$. Demostrar que, para $n$ impar, $T_n$ tiene un cero en el eje real y por tanto fuera de $E$.
2. Demostrar que los polinomios de Chebyshev en el intervalo $[-1,1]$ son los polinomios
$$T_n(x) = 2^{1-n} \cos(n \arc\cos x).$$
Concluir que $\mathrm{Cap}([-1,1]) = \mathrm{Cheb}([-1,1]) = 1/2$.
3. Sean $0 \leq b < a$ números reales y sea $E = [-a,-b] \cup [b,a]$.
   *a*) Demostrar que, si $p(x)$ es un polinomio de Chebyshev para $E$ de grado par, entonces $(p(x)+p(-x))/2$ también lo es. Por tanto podemos buscar polinomios de Chebyshev de grado $2k$ de la forma $q(x^2)$ con $q$ polinomio mónico de grado $k$.
   *b*) Demostrar que los polinomios
$$P_{2k}(x) = 2^{1-2k}(a^2-b^2)^k \cos\left(n \arc\cos\left(2\frac{x^2-b^2}{a^2-b^2} - 1\right)\right)$$
son polinomios de Chebyshev para $E$.
   *c*) Demostrar que
$$\mathrm{Cap}(E) = \frac{\sqrt{a^2-b^2}}{2}.$$
4. Sea $p(x)$ un polinomio mónico de grado $n$. Consideramos la *lemniscata*
$$L = \{z \in \mathbb{C} \mid |p(z)| \leq R^n\}.$$

Determinar los polinomios de Chebyshev de grado $nk$ para $E$ y mostrar que $\mathrm{Cap}(L) = R$.

**3.5. Potenciales y funciones armónicas.** En esta sección definiremos funciones armónicas y superarmónicas, daremos algunas propiedades y las usaremos para estudiar los potenciales $U^\mu$. Una referencia para el material de este capítulo es el libro [13].

**Definición 3.29.** Una función $f\colon D \to \mathbb{R}$ definida en un abierto $D \subset \mathbb{C}$ es armónica si cumple la *propiedad del valor medio*: Para todo $z \in D$, si el disco $|\zeta - z| \leq r$ está contenido en $D$, entonces

$$f(z) = \frac{1}{2\pi} \int_0^{2\pi} f(z + re^{i\theta})\, d\theta. \tag{3.10}$$

Una función $f$ es armónica si y sólo si $f$ es continua, sus primeras y segundas derivadas parciales existen y son continuas y para todo $z \in D$,

$$\Delta f(z) \coloneqq f_{xx}(z) + f_{yy}(z) = 0,$$

donde $z = x + iy$ y $f_{xx}$, $f_{yy}$ denotan las segundas derivadas parciales de $f$ respecto de $x$ e $y$. De hecho, esta propiedad se puede usar como definición equivalente de función armónica, por ejemplo en [13] Definición 1.1.1 y Teorema 1.2.7.

Una consecuencia directa de (3.10) es el *principio max-min* para funciones armónicas.

**Teorema 3.30.** *Sea $D \subset \mathbb{C}$ un abierto conexo y $f\colon D \to \mathbb{R}$ una función armónica. Si $f$ alcanza su mínimo o su máximo en $D$ entonces es constante. En consecuencia, si $f$ es una función continua en un compacto y es armónica en el interior del compacto entonces alcanza el mínimo y el máximo en la frontera del compacto.*

Las funciones superarmónicas son el análogo complejo de las funciones cóncavas.

**Definición 3.31.** Sea $D \subset \mathbb{C}$ un abierto. Una función $f\colon D \to \mathbb{R} \cup \{\infty\}$ es *superarmónica* si no es constante igual a $\infty$ en ninguna componente conexa de $D$ y

1. es semicontinua inferior en $D$;
2. el valor en cada punto es mayor o igual al valor medio en un círculo centrado en el punto. Esto es, si $z \in D$ y $r > 0$ cumplen que el disco $|\zeta - z| \leq r$ está contenido en $D$, entonces

$$f(z) \geq \frac{1}{2\pi} \int_0^{2\pi} f(z + re^{i\theta})\, d\theta. \tag{3.11}$$

Una función $f$ se llama subarmónica si $-f$ es superarmónica.

Las funciones superarmónicas verifican el *principio del mínimo*, y por tanto las subarmónicas verifican el principio del máximo [13, Theorem 2.3.1].

**Teorema 3.32.** *Sea $D \subset \mathbb{C}$ un conjunto abierto, conexo y acotado y $g$ una función superarmónica no constante en $D$ tal que*

$$\liminf_{z \to \zeta} g(z) \geq m, \qquad \forall \zeta \in \partial D.$$

*Entonces $g(z) > m$ para todo $z \in D$.*

La conexión entre funciones superarmónicas y potenciales viene dado por el siguiente resultado.

**Proposición 3.33.** *Sea $\mu$ una medida positiva cuyo soporte está contenido en un compacto $E$ y sea $\Omega$ la componente conexa no acotada de $\mathbb{C}\setminus E$, entonces el potencial*

$$U^{\mu}(z)=\int \log\frac{1}{|z-t|}\,d\mu(t)$$

*es una función superarmónica y su restricción a $\Omega$ es armónica.*

*Demostración.* Dado que la función $\log(1/|z-t|)$ es semicontinua inferior en $z$ para todo $t$ se sigue que $U^{\mu}(z)$ es semicontinua inferior. Es fácil comprobar que, para $t$ fija, la función $f_t(z)=\log(1/|z-t|)$ es superarmónica. De hecho, para $z\neq t$, $f_t(z)$ es armónica, mientras que, para $z=t$, $f_t(t)=\infty$ que es mayor o igual que la media en el círculo $|z-t|=r$ para $r$ suficientemente pequeño.

Usando el teorema de Fubini-Tonelli obtenemos

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi}U^{\mu}(z+re^{i\theta})\,d\theta&=\int\frac{1}{2\pi}\int_0^{2\pi}\log\frac{1}{|z+re^{i\theta}-t|}\,d\theta\,d\mu(t)\\&\leq\int\log\frac{1}{|z-t|}\,d\mu(t)=U^{\mu}(z),\end{aligned}$$

probando que $U^{\mu}$ es superarmónica. Que la restricción de $U^{\mu}$ a $\Omega$ es armónica se demuestra de la misma forma, usando que $\log(1/|z-t|)$ con $t$ fijo y $z\neq t$ es una función armónica. □

De hecho, toda función superarmónica se parece a un potencial con respecto de una medida positiva. Este es el contenido del Teorema de Riesz [13, Theorem A.3.2].

**Teorema 3.34** (Teorema de Riesz)**.** *Si $f$ es superarmónica en un abierto conexo $D\subset\mathbb{C}$, entonces existe una medida positiva $\lambda$ con soporte contenido en $D$ y en cada subconjunto abierto conexo $D^*$ con $\overline{D^*}\subset D$ se tiene*

$$f(z)=h(z)+\int_D\log\frac{1}{|z-t|}\,d\lambda(t),\quad z\in D^*$$

*donde $h$ es armónica en $D^*$.*

A partir del principio del máximo para funciones superarmónicas se obtiene un principio del máximo para potenciales.

**Teorema 3.35.** *Sea $\mu$ una medida positiva finita con soporte compacto. Si $U^{\mu}(z)\leq M$ para todo $z$ en el soporte de $\mu$, entonces $U^{\mu}(z)\leq M$ para todo $z\in\mathbb{C}$.*

El principio del máximo para potenciales es un caso particular del teorema de dominación (Ver [17] II Teorema 3.2.)

**Teorema 3.36.** *Sean $\mu$ y $\nu$ dos medidas positivas finitas con soporte compacto, masa total cumpliendo $\nu(\mathbb{C})\leq\mu(\mathbb{C})$ y tal que $\mu$ tiene energía logarítmica finita ($I(\mu)<\infty$). Si para una constante $c$, la desigualdad*

$$U^{\mu}(z)\leq U^{\nu}(z)+c$$

*se cumple salvo en un conjunto de $\mu$-medida cero, entonces la desigualdad se cumple en todo punto complejo.*

Muy relacionada con el potencial $U^{\mu_E}$ tenemos la función de Green para $E$ con polo en el infinito. Para las propiedades de la función de Green y más detalles sobre los próximos resultados se puede consultar la sección II.4 de [17]. La función de Green está caracterizada por el siguiente resultado.

**Teorema 3.37.** *Sea $E \subset \mathbb{C}$ un subconjunto compacto y $\Omega$ la componente conexa no acotada de $\mathbb{C} \setminus E$. Entonces existe una única función $g_E(\cdot, \infty)\colon \Omega \to \mathbb{R}$ caracterizada por las propiedades*

1. *$g_E(\cdot, \infty)$ es armónica en $\Omega$ y está acotada superior e inferiormente fuera de todo entorno de $\infty$;*
2. *la diferencia $g_E(z, \infty) - \log|z|$ está acotada en un entorno de $\infty$;*
3. $$\lim_{z \to \zeta} g(z, \infty) = 0$$
   *para todo $\zeta \in \partial\Omega$ excepto en un conjunto de capacidad cero.*

La relación precisa entre la función de Green, la constante de Robin y el potencial viene dada por el siguiente resultado.

**Teorema 3.38.** *Sean $E$ y $\Omega$ como en el teorema anterior y $g_E(\cdot, \infty)$ la función de Green para $E$ con polo en el infinito. Entonces*

1. *$\lim_{z\to\infty} g(z, \infty) - \log|z| = V_E$,*
2. *el potencial de equilibrio de $E$ es*
$$U^{\mu_E}(z) = V_E - g(z, \infty).$$

En los ejercicios 3.39 del 7 al 11 usamos las propiedades que caracterizan la función de Green para estudiar el efecto de las aplicaciones holomorfas en la capacidad.

*Ejercicios* 3.39.

1. Demostrar el Teorema 3.30.
2. Demostrar que si una función $f$ es armónica en un conjunto abierto $D \subset \mathbb{C}$ que contiene el disco $|z - z_0| \leq r$ entonces
$$f(z_0) = \frac{1}{\pi r^2} \int\int f(x + iy)\, dxdy.$$
3. Demostrar que si $(f_n)_{n\geq 1}$ es una sucesión de funciones armónicas que converge localmente uniformemente a una función $f$, entonces $f$ es armónica.
4. Demostrar que una función $f \in C^2(D)$ en un abierto conexo $D$ es superarmónica si y sólo si
$$\Delta f \coloneqq f_{xx} + f_{yy} \leq 0$$
en todos los puntos de $D$.
5. Sea $F\colon D \to \mathbb{C}$ una función analítica en un conjunto abierto conexo $D$ que no se anula idénticamente. Demostrar que, para $p > 0$, $|F|^p$ es subarmónica y que $\log(1/|F|)$ es superarmónica.
6. Buscar una demostración del Teorema 3.21.
7. Sea $f\colon D \to \mathbb{C}$ una función analítica en un abierto conexo $D$. Sea $D' = f(D)$ y $u\colon D' \to \mathbb{R}$ una función armónica (resp. superarmónica), demostrar que $u \circ f$ es armónica (resp. superarmónica).
8. Sean $E_1$ y $E_2$ conjuntos compactos conexos tales que $\Omega_1 = \mathbb{P}^1(\mathbb{C}) \setminus E_1$ y $\Omega_2 = \mathbb{P}^1(\mathbb{C}) \setminus E_2$ sean conexos y simplemente conexos. Sea $f\colon \Omega_1 \to \Omega_2$ una función analítica que envía $\partial\Omega_1$ continuamente a $\partial\Omega_2$ y que cumple $f(\infty) = \infty$ y $f'(\infty) > 0$. Si $g_{E_2}(\cdot, \infty)$ es la función de Green de $E_2$ entonces $g_{E_2}(f(z), \infty)$ es la función de Green para $E_1$. Utilizar este resultado para comparar la capacidad de $E_1$ y $E_2$ y sus potenciales de equilibrio.

9. Demostrar que el potencial de equilibrio del disco unidad $\overline{B} = \{|z| \leq 1\}$ es la función
$$-\log^+ |z| = -\max(0, \log |z|).$$
10. Utilizar los problemas 9 y 8 y la transformación $z \mapsto z + 1/z$ para calcular la capacidad logarítmica, la medida de equilibrio y el potencial de equilibrio del segmento $[-2, 2]$.
11. Mostrar que, si $E$ es un conjunto compacto de capacidad positiva y $p$ es un polinomio mónico de grado $n$ entonces
$$\mathrm{Cap}(p^{-1}(E)) = \mathrm{Cap}(E)^{1/n}.$$

**3.6. Demostración del Teorema de Frostman.** En esta sección damos la demostración del teorema de Frostman. Recordemos el resultado. Sea $E$ un conjunto compacto con $\mathrm{Cap}(E) > 0$, $\mu_E$ la medida de equilibrio de $E$ y $U^{\mu_E}$ el potencial de equilibrio asociado. Tenemos que ver

1. $U^{\mu_E}(z) \leq V_E$ para todo $z \in \mathbb{C}$.
2. $U^{\mu_E}(z) = V_E$ para todo $z \in E \setminus S$, donde $S$ tiene capacidad 0.
3. $U^{\mu_E}(z) < V_E$ para todo $z \in \Omega$, donde $\Omega$ es la componente conexa no acotada de $\mathbb{C} \setminus E$.

La demostración se realiza en varios pasos. En primer lugar sea
$$E_n = \{z \in E \,|\, U^{\mu_E} \leq V_E - 1/n\}, \quad n > 0.$$
Demostraremos por contradicción que
$$\mathrm{Cap}(E_n) = 0, \quad \forall n > 0. \tag{3.12}$$
Supongamos que $\mathrm{Cap}(E_n) > 0$ para algún $n > 0$. Como $V_E = I(\mu_E) = \int U^{\mu_E}\, d\mu_E$, existe un $z_0$ en el soporte de $\mu_E$ con $U^{\mu_E}(z_0) \geq V_E$. Por la semicontinuidad inferior de $U^{\mu_E}$ existe un $r > 0$ de tal forma que $U^{\mu_E}(z) > V_E - 1/2n$ en el disco $\overline{B}(z_0, r)$. En particular $\overline{B}(z_0, r) \cap E_n = \emptyset$. Como $z_0$ está en el soporte de $\mu_E$, el número
$$a \coloneqq \mu_E(\overline{B}(z_0, r))$$
es estrictamente positivo.

Como suponemos que $\mathrm{Cap}(E_n) > 0$, existe una medida $\mu \in \mathcal{M}(E_n)$ con $I(\mu) < \infty$. Definimos una medida con signo
$$\sigma = \begin{cases} \mu, & \text{en } E_n, \\ -\mu_E/a, & \text{en } \overline{B}(z_0, r), \\ 0, & \text{en el resto.} \end{cases}$$
Para cada $t \in (0, a)$ la medida $\mu_t = \mu_E + t\sigma$ es positiva y de masa total 1. Por tanto $\mu_t \in \mathcal{M}(E)$. Un cálculo directo muestra que
$$I(\mu_E) - I(\mu_t) \geq 2t\,(V_E - 1/2n - V_E + 1/n) + O(t^2).$$
por tanto, para $t$ suficientemente pequeño, $I(\mu_E) \geq I(\mu_t)$ lo que contradice la minimalidad de $I(\mu_E)$.

En el segundo paso, sea
$$L_n = \{z \in \mathrm{soporte}(\mu_E) \mid U^{\mu_E}(z) > V_E + 1/n\}, \quad n > 1.$$
Demostraremos por contradicción que
$$L_n = \emptyset. \tag{3.13}$$

Supongamos que algún $L_n$ es no vacío y sea $z_1 \in L_n$, por la semicontinuidad inferior de $U^{\mu_E}$ existe un $s > 0$ tal que $U^{\mu_E}(z) > V_E + 1/n$ en $\overline{B}(z_1, s)$. Como $z_1$ está en el soporte de $\mu_E$, el número

$$b \coloneqq \mu_E(\overline{B}(z_1, s))$$

es estrictamente positivo.

Por (3.12) y el Teorema 3.21 (5) sabemos que $\mu_E(E_n) = 0$ y por tanto $U^{\mu_E}(z) \geq V_E$ salvo en un conjunto de $\mu_E$ medida cero. Así,

$$\begin{aligned} V_E = I(\mu_e) &= \int_E U^{\mu_E}\, d\mu_E \\ &= \int_{\overline{B}(z_1,s)} U^{\mu_E}\, d\mu_E + \int_{E\setminus B(z_1,s)} U^{\mu_E}\, d\mu_E \\ &\geq \left(V_E + \frac{1}{n}\right) b + V_E(1-b) \\ &> V_E, \end{aligned}$$

obteniendo una contradicción.

Para concluir la demostración del teorema, observamos que (3.13) implica que $U^{\mu_E} \leq V_E$ en el soporte de $\mu_E$. Así, aplicando el principio del máximo para potenciales (Teorema 3.35) deducimos la afirmación (1). Usando la condición (3.12) se puede ver que el conjunto $S = \bigcup_n E_n$ tiene capacidad cero, por lo que obtenemos la afirmación (2). La afirmación (3) se sigue del hecho que $U^{\mu_E}$ es armónica en $\Omega$ junto con el Teorema 3.30.

### 3.7. El teorema de la lemniscata de Hilbert.

**Definición 3.40.** Sea $p \in \mathbb{C}[x]$ un polinomio mónico de grado $d$ y $0 < \rho \in \mathbb{R}$ una constante real. La *lemniscata de polinomio $p$ y constante $\rho$* es el conjunto

$$L = L_{p,\rho} = \{z \in \mathbb{C} \mid |p(z)| \leq \rho^d\}.$$

En esta sección veremos que todo conjunto de capacidad mayor que cero se puede aproximar por una lemniscata.

**Teorema 3.41** (El teorema de la lemniscata de Hilbert)**.** *Sea $E$ un conjunto compacto con $\mathrm{Cap}(E) > 0$ y sea $U \supset E$ un entorno abierto de $E$ con $\mathbb{C} \setminus U$ conexo. Entonces existe un polinomio mónico de grado $d > 0$, $p[x] = x^d + \cdots \in \mathbb{C}[x]$ y una constante $\rho > \mathrm{Cap}(E)$ tal que*

$$E \subset L_{p,\rho} \subset U.$$

*Demostración.* Como $E$ es compacto, está acotado. Restringiéndonos a un $U$ más pequeño si hace falta, podemos suponer que $U$ también está acotado. Sea $R > 0$ tal que

$$U \subset B(0, R) = \{z \in \mathbb{C} \mid |z| < R\}.$$

El conjunto $K \coloneqq \overline{B}(0, R) \setminus U$ es compacto y, por el Teorema de Frostman 3.25, $U^{\mu_E} < V_E$ en $K$. Por tanto existe $\epsilon > 0$ tal que $U^{\mu_E}(z) \leq V_E - \epsilon$ para todo $z \in K$.

Por el Teorema 3.27 4 existe un $n_0$ y para todo $n \geq n_0$ y todo $z \in K$

$$\left| \log \frac{1}{|F_n(z)|^{1/n}} - U^{\mu_E}(z) \right| < \epsilon/2,$$

lo que implica

$$\log \frac{1}{|F_n(z)|^{1/n}} < V_E - \epsilon/2. \tag{3.14}$$

Como $\log(1/|F_n(z)|)$ es armónica en el complementario de $\overline{B}(0,R)$, por el Teorema 3.30, la desigualdad (3.14) en $K$ implica que esta desigualdad es cierta en todo $\mathbb{C} \setminus U$.

En consecuencia, para todo $n \geq n_0$ y $z \notin U$,

$$|F_n(z)|^{1/n} > e^{-V_E} e^{\epsilon/2} = \mathrm{Cap}(E) e^{\epsilon/2}. \tag{3.15}$$

Utilizando ahora el Teorema 3.27 apartados (1) y (2), sabemos que existe un $n_1$ y para todo $n \geq n_1$ se tiene

$$\|F_n\|_E^{1/n} < \mathrm{Cap}(E) e^{\epsilon/2}.$$

Eligiendo $d > \max(n_0, n_1)$ y poniendo $p = F_d$, $\rho = \mathrm{Cap}(E) e^{\epsilon/2}$ obtenemos el resultado. □

**3.8. Las órbitas de Galois y el teorema de Fekete.** Un entero algebraico es un número complejo que es solución de un polinomio mónico con coeficientes enteros. El conjunto de enteros algebraicos se denota como $\overline{\mathbb{Z}}$. Por tanto $\zeta \in \overline{\mathbb{Z}}$ si y sólo si existe un polinomio $p(z) = z^n + \cdots \in \mathbb{Z}[z]$ tal que

$$p(\zeta) = 0. \tag{3.16}$$

Entre todos los polinomios mónicos con coeficientes enteros que cumplen la ecuación (3.16) hay uno de grado mínimo denominado el *polinomio mínimo* de $\zeta$. Este polinomio es irreducible sobre los enteros (en caso contrario no sería minimal) y, por tanto, tiene todas sus raíces complejas simples. La órbita de Galois de $\zeta$ es el conjunto de raíces del polinomio mínimo de $\zeta$. La órbita de Galois de $\zeta$ la denotaremos como $\mathrm{Gal}(\zeta)$. El grado del polinomio mínimo de $\zeta$ se denomina el grado de $\zeta$. El primer resultado [9] utiliza la identificación entre la capacidad logarítmica y el diámetro transfinito.

**Teorema 3.42** (Teorema de Fekete)**.** *Sea $E \subset \mathbb{C}$ un conjunto compacto. Si $\mathrm{Cap}(E) < 1$, entonces existe un abierto $U$ que contiene a $E$ y tal que el conjunto de enteros $\zeta \in \overline{\mathbb{Z}}$ con $\mathrm{Gal}(\zeta) \subset U$ es finito. En particular $E$ contiene únicamente un número finito de órbitas de Galois.*

*Demostración.* Para cada $\varepsilon > 0$ escribimos

$$B(E,\varepsilon) = \{z \in \mathbb{C} \mid d(z,E) < \varepsilon\}$$
$$\overline{B}(E,\varepsilon) = \{z \in \mathbb{C} \mid d(z,E) \leq \varepsilon\},$$

donde $d(z,E) = \min_{x \in E} |z - x|$. Como $E$ es compacto, los conjuntos $\overline{B}(E,1/n)$ son compactos y

$$E = \bigcap_{n \geq 1} \overline{B}(E,1/n).$$

Por tanto

$$\mathrm{Cap}(E) = \lim_{n \to \infty} \mathrm{Cap}(\overline{B}(E,1/n)).$$

Como $\mathrm{Cap}(E) < 1$, existe $n_0 > 0$ con $\mathrm{Cap}(\overline{B}(E,1/n_0)) < 1$. Si demostramos que $\overline{B}(E,1/n_0)$ solo contiene un número finito de órbitas de Galois, entonces $U \coloneqq B(E,1/n_0)$ es un abierto que contiene a $E$ y que solo contiene un número

finito de órbitas de Galois. Reemplazando $E$ por $\overline{B}(E, 1/n_0)$ basta demostrar que $E$ solo contiene un número finito de órbitas de Galois.

Supongamos que $E$ contiene un número infinito de órbitas de Galois de enteros algebraicos. Elijamos una sucesión numerable de tales órbitas, $\mathrm{Gal}(\zeta_n)$, $n \geq 1$ y sean $p_n$ los polinomios mínimos de cada $\zeta_n$. Supongamos que hubiera una cota uniforme $d$ al grado de estos polinomios. Como $E$ es compacto, está acotado. Dado que los coeficientes de cada $p_n$ son las funciones simétricas elementales en las raíces de $p_n$ y, por hipótesis, estas están en $E$, deducimos que los coeficientes de los polinomios $p_n$ están uniformemente acotados. Como el grado es finito solo puede haber un número finito de ellos, contradiciendo la suposición de que el número de órbitas de Galois contenidas en $E$ es infinito. Así podemos suponer que el grado de los polinomios mínimos no está acotado.

Como estamos asumiendo que el grado de los polinomios $p_n$ no está acotado, después de restringirse a una subsucesión, podemos suponer que la sucesión de grados

$$d_n \coloneqq \deg(p_n)$$

es estrictamente creciente.

Por definición

$$\delta_{d_n}(E) \geq \left| \prod_{\xi \neq \eta \in \mathrm{Gal}(\zeta_n)} (\xi - \eta) \right|^{1/d_n(d_n-1)} = |\mathrm{disc}(p_n)|^{1/d_n(d_n-1)}, \tag{3.17}$$

donde $\mathrm{disc}(p_n)$ es el discriminante del polinomio $p_n$ (ver sección 2.3). Dado que $p_n$ es el polinomio mínimo de un entero algebraico, es un polinomio mónico con coeficientes enteros. Por la Proposición 2.25, sabemos que $\mathrm{disc}(p_n)$ es un número entero y por ser $p_n$ irreducible, $\mathrm{disc}(p_n)$ es un entero distinto de cero. La ecuación (3.17) implica que $\delta_{d_n}(E) \geq 1$. Por tanto

$$\mathrm{Cap}(E) = \tau(E) \geq 1$$

lo que contradice la hipótesis de que $\mathrm{Cap}(E) < 1$. □

**3.9. El teorema de Fekete-Szegö.** El teorema de Fekete-Szegö [10] es casi un recíproco del teorema de Fekete y apareció más de 30 años más tarde que el teorema de Fekete. Este resultado asegura que, si la capacidad logarítmica de un conjunto simétrico respecto a la conjugación compleja, es mayor que uno, entonces todo entorno del conjunto contiene infinitas órbitas de Galois de enteros algebraicos.

**Teorema 3.43** (Fekete-Szegö)**.** *Sea $E \subset \mathbb{C}$ un subconjunto compacto, simétrico respecto a la conjugación compleja y tal que* $\mathrm{Cap}(E) \geq 1$. *Si $U \supset E$ es un abierto, entonces $U$ contiene infinitas órbitas de Galois de enteros algebraicos.*

*Demostración.* Por simplicidad, demostraremos solo el caso en el que $\mathbb{C} \setminus U$ es conexo. Veamos primero que una lemniscata respecto a un polinomio mónico con coeficientes enteros y con constante $\rho = 1$ contiene infinitas órbitas de Galois de enteros algebraicos. Sea $p(z) = z^n + \cdots \in \mathbb{Z}[z]$ un polinomio mónico con coeficientes enteros y sea

$$L = \{z \in \mathbb{C} \mid |p(z)| \leq 1\}.$$

Para cada $n \geq 1$ consideramos el polinomio

$$f_n(z) = p^n(z) - 1.$$

Claramente $f_n(z)$ es un polinomio mónico con coeficientes enteros y todas las raíces de $f_n$ están contenidas en $L$. Además el conjunto

$$S = \bigcup_{n\geq 1} \{x \in \mathbb{C} \mid f_n(x) = 0\}$$

es infinito, pues la aplicación $S \to \{\text{raíces de } 1\}$ dada por $x \mapsto p(x)$ es exhaustivo. Por lo que $S$, y por tanto $L$, contiene un número infinito de órbitas de Galois.

En consecuencia, para obtener el resultado, basta demostrar que el conjunto $U$ contiene la lemniscata de un polinomio mónico con coeficientes enteros y constante 1.

Ya hemos visto, en el Teorema de la lemniscata de Hilbert 3.41, que existe un polinomio $p(z) = z^d + \cdots \in \mathbb{C}[z]$ y una constante $\rho > 1$ tal que la lemniscata de polinomio $p$ y constante $\rho$ está contenida en $U$. Ya tenemos un polinomio $p$, pero tiene coeficientes complejos. Debemos hacer una serie de reducciones para conseguir un polinomio con coeficientes enteros.

Dado que $E$ es simétrico respecto a conjugación compleja, podemos reemplazar $U$ por $U \cap \overline{U}$ y suponer que $U$ también es invariante por conjugación compleja. Si ahora $p$ cumple $|p(z)| > \rho^d > 1$ para todo $z \in \mathbb{C} \setminus U$, entonces $|p(z)\overline{p}(z)| > \rho^{2d} > 1$ para todo $z \in \mathbb{C} \setminus U$. En consecuencia, la lemniscata de polinomio $p\overline{p}$ y constante $\rho > 1$ está contenido en $U$. Pero el polinomio $p_1 = p\overline{p}$ tiene coeficientes reales. Esta ha sido la primera reducción.

Usando que los números racionales son densos dentro de $\mathbb{R}$, es fácil ver que podemos encontrar un polinomio $p_2$ con coeficientes racionales y una constante $1 < \rho' < \rho$ tal que la lemniscata de polinomio $p_2$ y constante $\rho'$ está contenida en $U$. Nos hemos reducido a un polinomio con coeficientes racionales.

El último paso es ver que podemos construir una lemniscata de polinomio con coeficientes enteros y constante 1. Este proceso se denomina corrección (en ingles "patching") y está explicado en el siguiente lema.

**Lema 3.44.** *Sea $p(x) = x^d + \cdots \in \mathbb{Q}[x]$ un polinomio mónico con coeficientes racionales y $1 < \rho \in \mathbb{R}$ un número mayor que uno. Sea $L$ la lemniscata de polinomio $p$ y constante $\rho$:*

$$L = \{z \in \mathbb{C} \mid |p(z)| \leq \rho^d\}.$$

*Entonces existe un polinomio mónico $\Gamma(x) \in \mathbb{Z}[x]$ con coeficientes enteros, y la lemniscata de polinomio $\Gamma$ y constante 1 está contenida en $L$.*

*Demostración del Lema 3.44.* Sea $0 < n \in \mathbb{Z}$ un entero tal que

$$p(x) = x^d + \frac{1}{n}\gamma(x), \quad \gamma(x) \in \mathbb{Z}[x].$$

Si $n = 1$, $p$ ya tiene coeficientes enteros, así que supondremos que $n \geq 2$. Sea $\mu \geq 1$ un número entero que fijaremos más tarde. Escribimos $\sigma = \mu d$ y $\nu = \sigma! n^\sigma$. Entonces el polinomio $p^\nu$ tiene los coeficientes de mayor grado enteros:

$$p^\nu(x) = x^{\nu d} + \frac{\nu}{d}x^{d(\nu-1)}\gamma + \cdots + \frac{\nu(\nu-1)\ldots(\nu-\sigma-1)}{\sigma! n^\sigma}x^{d(\nu-\sigma)}\gamma^\sigma + \binom{\nu}{\sigma+1}\frac{1}{n^{\sigma+1}}x^{d(\nu-\sigma-1)}\gamma^{\sigma+1} + \ldots,$$

dada la elección de $\nu$ todos los coeficientes de la primera fila son enteros. Los coeficientes restantes pueden ser racionales. Observemos que el grado en el que

aparece el primer coeficiente que puede ser no entero es

$$d(\nu-\sigma-1)+(d-1)(\sigma+1)=d\nu-\sigma-1=d(\nu-\mu)-1.$$

Así, como mucho, tenemos que corregir $d(\nu-\mu)$ coeficientes racionales. Ahora corregimos estos coeficientes racionales intentando controlar el tamaño de la corrección. Se pueden encontrar polinomios $q_\ell$, $\ell=0,\dots,\nu-\mu-1$, de grado menor o igual a $d-1$, cuyos coeficientes sean racionales en el intervalo $[0,1)$ y tales que

$$\Gamma(x)\coloneqq p^\nu(x)+\sum_{\ell=0}^{\nu-\mu-1}p^\ell(x)q_\ell(x)\in\mathbb{Z}[x]$$

tenga coeficientes enteros.

Obsérvese que el polinomio $\Gamma$ depende de la elección de $\mu$. Queda por ver que, mediante una elección adecuada de $\mu$, el polinomio $\Gamma$ cumple que su lemniscata con coeficiente 1 está contenido en $L$. La frontera de $L$ es

$$\partial L=\{z\in\mathbb{C}\mid |p(z)|=\rho^d\}.$$

Escribimos $\Delta=\Gamma-p^\nu$. Como los coeficientes de los polinomios $q_\ell$ están en el intervalo $[0,1)$, para todo $z\in\partial L$ tenemos la cota

$$\frac{|\Delta(z)|}{|p^\nu(z)|}\le M\cdot\left|\frac{1}{\rho^{d(\mu+1)}}+\cdots+\frac{1}{\rho^{d\nu}}\right|\le\frac{M}{\rho^{d\mu}(\rho^d-1)},$$

donde

$$M=\sup_{z\in\partial L}(1+|z|+\cdots+|z|^{d-1}).$$

Como $\rho>1$, existe un $\mu_0$ tal que, para todo $\mu\ge\mu_0$ y $z\in\partial L$,

$$\frac{|\Delta(z)|}{|p^\nu(z)|}\le\frac{1}{2}. \tag{3.18}$$

Recordemos el teorema de Rouché de variable compleja. Dadas dos funciones analíticas $f,g$ en un dominio $K$, si $|f(z)|<|g(z)|$ para todo $z\in\partial K$, entonces el número de raíces (contadas con su multiplicidad) de $g$ y de $f+g$ contenidas en $K$ es igual. Este teorema, junto con la ecuación (3.18), implica que todas las raíces de $\Gamma$ están contenidas en $L$.

Por otro lado, para todo $z\in\partial L$, tenemos la estimación

$$|\Gamma(z)|\ge|p^\nu(z)|-|\Delta(z)|=\rho^{d\nu}\left(1-\frac{|\Delta(z)|}{|p^\nu(z)|}\right)>\frac{\rho^{d\nu}}{2}>\frac{\rho^{d\mu}}{2}.$$

De nuevo, debido a que $\rho>1$ podemos encontrar $\mu_1$, tal que, para todo $\mu\ge\mu_1$ y todo $z\in\partial L$, se cumple

$$|\Gamma(z)|>1. \tag{3.19}$$

Utilizando el principio del mínimo para funciones holomorfas, del hecho de que todas las raíces de $\Gamma$ estén contenidas en $L$ y la ecuación (3.19) se deduce que, si $z\notin L$, entonces $|\Gamma(z)|>1$. Por tanto la lemniscata de polinomio $\Gamma$ y constante 1 esta contenida en $L$ concluyendo la demostración del Lema. □

El Teorema 3.43 es consecuencia del Lema 3.44 y de la discusión precedente. □

**3.10. Equidistribución de órbitas de Galois en el caso de capacidad uno.** Como siempre $E \subset \mathbb{C}$ denota un conjunto compacto. Vamos a ver que, en el caso límite, cuando $\mathrm{Cap}(E) = 1$, si bien tenemos infinitas órbitas de Galois en un entorno de $E$, estas no tienen espacio para moverse y se equidistribuyen según la medida de equilibrio de $E$.

**Teorema 3.45.** *Sea $E \subset \mathbb{C}$ un conjunto compacto con $\mathrm{Cap}(E) = 1$ y sea $(\alpha_n)_{n>0}$ una sucesión de enteros algebraicos de grado $d_n$ cumpliendo*

1. $\lim_{n\to\infty} d_n = \infty$,
2. *Para todo $n > 0$, $\mathrm{Gal}(\alpha_n) \subset B(E, 1/n)$.*

*Entonces*

$$\nu(\mathrm{Gal}(\alpha_n)) \xrightarrow{*} \mu^E.$$

*Demostración.* Por el Teorema 3.9, la sucesión $(\nu(\mathrm{Gal}(\alpha_n)))_{n>0}$ tiene una subsucesión convergente en la topología débil-$*$. Si vemos que toda subsucesión convergente de $\nu(\mathrm{Gal}(\alpha_n))$ converge a $\mu_E$, entonces, necesariamente toda la sucesión $\nu(\mathrm{Gal}(\alpha_n))$ converge a $\mu^E$.

Así basta considerar el caso en que la sucesión $\nu(\mathrm{Gal}(\alpha_n))$ converge en la topología débil-$*$ a una medida $\hat{\mu}$ y demostrar que $\hat{\mu} = \mu^E$. Para cada $n$ denotamos por $P_n$ el polinomio entero mónico minimal de $\alpha_n$. De esta forma, $\mathrm{Gal}(\alpha_n)$ es el conjunto de raíces de $P_n$.

Por el Lema 3.11, tenemos que $\mathrm{soporte}(\hat{\mu}) \subset E$ y, por tanto,

$$I(\hat{\mu}) \geq V_E = 0.$$

Por otro lado,

$$\begin{aligned}
I(\hat{\mu}) &= \int\int \log \frac{1}{|z-t|} d\hat{\mu}(z) d\hat{\mu}(t) \\
&= \lim_{M\to\infty} \int\int \log_M \frac{1}{|z-t|} d\hat{\mu}(z) d\hat{\mu}(t) \\
&= \lim_{M\to\infty} \lim_{n\to\infty} \int\int \log_M \frac{1}{|z-t|} d\nu(\mathrm{Gal}(\alpha_n))(z) d\nu(\mathrm{Gal}(\alpha_n)(t) \\
&\leq \lim_{M\to\infty} \lim_{n\to\infty} \left( \frac{1}{d_n^2} \sum_{\beta\neq\beta'\in\mathrm{Gal}(\alpha_n)} \log \frac{1}{|\beta-\beta'|} + \frac{M}{d_n} \right) \\
&\leq \lim_{M\to\infty} \lim_{n\to\infty} \left( \frac{-\log(|\operatorname{disc} P_n|)}{d_n^2} + \frac{M}{d_n} \right) \leq 0.
\end{aligned}$$

En la última desigualdad hemos usado de nuevo la Proposición 2.25 que implica que $|\operatorname{disc} P_n|$ es un entero positivo no nulo y, por tanto $-\log(|\operatorname{disc} P_n|) \leq 0$. En consecuencia $I(\hat{\mu}) = 0 = V_E$. Por la unicidad de la medida de equilibrio deducimos que $\hat{\mu} = \mu^E$ concluyendo la demostración. □

## 4. Alturas y teorema de Bilu

**4.1. Medida de Mahler de un polinomio en una variable.** Sea $f(x) \in \mathbb{C}[x]$ un polinomio no nulo con coeficientes complejos. Definimos su *medida de Mahler* por

$$M(f) = \exp\Big(\frac{1}{2\pi} \int_0^{2\pi} \log|f(e^{i\theta})| d\theta\Big).$$

No es difícil ver que la integral es absolutamente convergente, incluso si el polinomio tiene raíces sobre el círculo unitario.

**Lema 4.1.** *Para todo polinomio no nulo $f \in \mathbb{C}[x]$, se tiene $M(f) = M(f^*)$.*

*Demostración.* si $n = \deg f$, tenemos $f^*(x) = x^n f(1/x)$, luego

$$M(f^*) = \exp\Big(\frac{1}{2\pi}\int_0^{2\pi} \log|f(e^{-i\theta})|d\theta\Big).$$

El enunciado resulta entonces de un cambio de variable apropiado. □

Nos será útil en lo que sigue usar la función $\log^+$, dada por

$$\log^+ z = \begin{cases} \log z & \text{si } z \geq 1 \\ 0 & \text{si } 0 < z < 1. \end{cases}$$

La medida de Mahler satisface las propiedades siguientes:

**Teorema 4.2.**

1. *$M(fg) = M(f)M(g)$, para todo $f, g \in \mathbb{C}[x], fg \neq 0$.*
2. *$M(f) > 0$, para todo $f \in \mathbb{C}[x]$, no nulo.*
3. *(Fórmula de Jensen) Sea $f(x) \in \mathbb{C}[x]$, un polinomio no nulo de grado $n$, que escribimos de la forma*

$$f(x) = a(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n), \quad a, \alpha_1, \ldots, \alpha_n \in \mathbb{C}. \tag{4.1}$$

*Entonces*

$$\log M(f) = \log|a| + \sum_{i=1}^{n} \log^+|\alpha_i|. \tag{4.2}$$

*Demostración.* Las primeras dos propiedades se deducen directamente de la definición. Usando (1), nos reducimos a mostrar que para todo $\alpha \in \mathbb{C}$, se tiene

$$\frac{1}{2\pi}\int_0^{2\pi} \log|e^{i\theta} - \alpha|d\theta = \log^+|\alpha|. \tag{4.3}$$

Para demostrar esta identidad, usaremos que el promedio de una función armónica $h$ sobre el borde del círculo unitario es $h(0)$.

Si $|\alpha| > 1$, entonces la función $h(x) = \log|x - \alpha|$ es armónica en el disco unitario, de manera que el término izquierdo de (4.3) es $h(0) = \log|\alpha|$.

Si $|\alpha| < 1$, entonces la función $w(x) = \log|1 - \alpha\bar{x}|$ es armónica en el disco unitario y coincide con $h(x)$ cuando $|x| = 1$. Por lo tanto el término izquierdo de (4.3) es $w(0) = \log|1| = 0$.

El caso $|\alpha| = 1$ se deduce por la continuidad de la función $\alpha \mapsto \int_0^{2\pi} \log|e^{i\theta} - \alpha|d$. □

*Ejercicio* 4.3. Demostrar

$$|D(f)| \leq n^n M(f)^{2n-2},$$

donde $D(f)$ denota el discriminante de $f$ (Sección 2.3)
Indicación: exprese el discriminante como un determinante de Vandermonde.

**4.2. Altura de un número algebraico.** Sea $\alpha \in \overline{\mathbb{Q}}^*$ y sea $f_\alpha(x) \in \mathbb{Q}[x]$ su polinomio mínimo. Existe un único entero positivo $m$ tal que el polinomio

$$g_\alpha(x) := m f_\alpha(x)$$

satisface

- $g_\alpha(x)$ tiene coeficientes enteros
- el máximo común divisor de los coeficientes de $g_\alpha(x)$ es 1.

Decimos que $g_\alpha(x)$ es el *polinomio mínimo sobre* $\mathbb{Z}$ *de* $\alpha$.

El claro que la órbita galoisiana de $\alpha$ se calcula por

$$\mathrm{Gal}(\alpha) = \{\beta \in \mathbb{C} : g_\alpha(\beta) = 0\}$$

(cf. Definición 2.16).

**Definición 4.4.** Sea $\alpha \in \overline{\mathbb{Q}}^*$. Definimos su *altura* (de Weil) por

$$h(\alpha) = \frac{1}{\deg \alpha} \log M(g_\alpha).$$

De los lemas 2.8 y 4.1 deducimos

**Lema 4.5.** *Para todo* $\alpha \in \overline{\mathbb{Q}}^*$, *se tiene* $h(\alpha) = h(1/\alpha)$.

Escribiendo

$$g_\alpha(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0,$$

de la fórmula de Jensen (4.2) se tiene

$$h(\alpha) = \frac{1}{\deg \alpha} \Big( \log |a_n| + \sum_{\beta \in \mathrm{Gal}(\alpha)} \log^+ |\beta| \Big). \tag{4.4}$$

**Proposición 4.6.** *Se tiene que* $h(\alpha) \geq 0$, *para todo* $\alpha \in \overline{\mathbb{Q}}^*$. *Más aún,* $h(\alpha) = 0$ *si y sólo si* $\alpha$ *es una raíz de la unidad.*

*Demostración.* En la expresión (4.4), $|a_n|$ es un entero positivo, de donde es claro que $h(\alpha) \geq 0$. Si $\alpha$ es una raíz de la unidad, la Proposición 2.19 asegura que $|\beta| = 1$, para todo $\beta \in \mathrm{Gal}(\alpha)$. Además, de la Observación 2.13 tenemos que su polinomio mínimo sobre $\mathbb{Z}$ es mónico. Usando nuevamente (4.4), tenemos $h(\alpha) = 0$.

Supongamos ahora $h(\alpha) = 0$. Sea $n = \deg(\alpha)$. Escribamos

$$g_\alpha(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0, \quad a_i \in \mathbb{Z}. \tag{4.5}$$

Entonces la expresión (4.4) garantiza que $a_n = 1$ y

$$|\beta| \leq 1, \tag{4.6}$$

para todo $\beta \in \mathrm{Gal}(\alpha)$.

La desigualdad (4.6) también vale para productos de elementos en $\mathrm{Gal}(\alpha)$. Entonces, de la fórmula $g_\alpha(x) = \prod_{\beta \in \mathrm{Gal}(\alpha)} (x - \beta)$, vemos que los coeficientes de $g_\alpha(x)$ satisfacen

$$|a_j| \leq n, \quad \forall j = 0, 1, 2, \ldots, n. \tag{4.7}$$

Sea $S_n$ el conjunto formado por todos los polinomios con coeficientes enteros, de grado a lo más $n$, de la forma (4.5) que además satisfacen (4.7). Tenemos que $S_n$ es un conjunto finito.

Ahora bien, si $k$ es un entero positivo, usando el Ejercicio 2,21, (2) vemos que el razonamiento anterior también se aplica a $\alpha^k$, es decir, $\{g_{\alpha^k}(x) : k \in \mathbb{Z}_{>0}\} \subseteq S_n$.

Como $S_n$ es finito, el conjunto de raíces de polinomios en $S_n$ también lo es, luego el conjunto

$$\{\alpha^k : k \in \mathbb{Z}_{>0}\}$$

es finito. Por lo tanto, existen enteros $k_1 \neq k_2$ tales que $\alpha^{k_1} = \alpha^{k_2}$ (equivalentemente, $\alpha^{k_1-k_2} = 1$), lo que prueba que $\alpha$ es una raíz de la unidad □

Terminamos esta sección con un problema abierto.

**Pregunta 4.7.** (Lehmer [12]). Decidir si la siguiente afirmación es cierta: existe una constante $c > 0$ tal que

$$h(\alpha) \geq \frac{c}{\deg \alpha}, \quad \forall \alpha \in \overline{\mathbb{Q}}^* \text{ que no es una raíz de la unidad.}$$

*Ejercicio* 4.8. Adapte la demostración de la Proposición 4.6 para demostrar el teorema de Northcott: Sean $A, B > 0$. El conjunto

$$\{\alpha \in \overline{\mathbb{Q}}^* : \deg(\alpha) \leq A, \quad h(\alpha) \leq B\}$$

es finito.

### 4.3. Enunciado del Teorema de Bilu en dimensión 1.

**Definición 4.9.** Sean $C \subset \mathbb{C}$ y $\nu \in \mathcal{M}(\mathbb{C})$. Decimos que $\nu$ *está soportada en* $C$ si para toda $f \in C_0(\mathbb{C})$ cuyo soporte es disjunto a $C$, se tiene

$$\int_{\mathbb{C}} f\mu = 0.$$

Si $\nu$ está soportada en un conjunto compacto, decimos que $\nu$ tiene soporte compacto.

Denotamos por $\mathcal{M}(C) \subseteq \mathcal{M}(\mathbb{C})$ al conjunto de medidas soportadas en $C$.

**Ejemplo 4.10.** Denotamos por $\nu_S$ la medida caracterizada por

$$\int_{\mathbb{C}} f\nu_S = \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\theta}) d\theta, \quad \forall f \in C_0(\mathbb{C}). \tag{4.8}$$

Es claro que $\nu_S$ está soportada en el círculo unitario y una medida de contaje de la forma $\nu(E)$ está soportada en $E$. Todas estas medidas tienen soporte compacto.

Recordemos el Teorema 3.9, cuando $C$ es compacto, el espacio $\mathcal{M}(C)$ es secuencialmente compacto. Es decir, toda sucesión $(\nu_n) \subset \mathcal{M}(C)$ admite una subsucesión convergente.

**Definición 4.11.** Consideremos una sucesión $E_n \subset \mathbb{C}$, donde cada $E_n$ es un conjunto finito. Decimos que la familia $(E_n)$ *se equidistribuye con respecto a* $\nu \in M$ si

$$\lim_{n\to\infty} \nu(E_n) = \nu.$$

Aquí, $\nu(E_n)$ es la medida de contaje definida en (3.1).

De manera equivalente, se pide que para toda $f \in C_0(\mathbb{C})$,

$$\lim_{n\to\infty} \frac{1}{\#E_n} \sum_{z\in E_n} f(z) = \int_{\mathbb{C}} f\nu.$$

**Definición 4.12.** Decimos que una sucesión $(x_n) \subset \mathbb{C}$ es *genérica* si para todo $m \in \mathbb{N}$, el conjunto

$$\{k \in \mathbb{N} : x_k = x_m\}$$

es finito.

**Teorema 4.13.** *(Bilu,* [5]*) Sea $\nu_S$ la medida de probabilidad uniforme sobre el círculo, dada por* (4.8)*. Sea $(\alpha_n) \subset \overline{\mathbb{Q}}^*$ una sucesión genérica de números algebraicos tal que*

$$\lim_{n\to\infty} h(\alpha_n) = 0.$$

*Entonces la familia de conjuntos* $\mathrm{Gal}(\alpha_n)$ *se equidistribuye con respecto a $\nu_S$.*

En particular, se tiene el siguiente resultado notable:

**Corolario 4.14.** *La sucesión de conjuntos $(\tilde{\mu}_n)$ se equidistribuye con respecto a la medida $\nu_S$.*

*Demostración.* tomando en cuenta la Proposición 2.19 y la Proposición 4.6, el enunciado se deduce directamente del Teorema de Bilu □

*Ejercicios* 4.15. El propósito de los ejercicios de esta sección es de indicar una demostración del Corolario 4.14 que no utiliza el Teorema de Bilu.

1. Sea $X$ un espacio métrico compacto. Los espacios $\mathcal{M}(X), C_0(X)$ y la noción de convergencia de medidas se definen como en la Sección 3.2. El espacio $C_0(X) = C(X)$ se encuentra dotado de la topología dada por la norma supremo:
$$\|f\| = \sup_{z\in X} |f(z)|.$$
Sea $H \subset C(X)$ una subálgebra (es decir, $f, g \in H \to fg, f + g \in H$). Sea $V \subset H$ un conjunto generador (es decir, todo elemento de $H$ es una combinación $\mathbb{C}$-lineal finita de elementos de $V$). Suponga que $H$ es conjunto denso. Muestre que la sucesión de medidas $(\nu_n) \in \mathcal{M}(X)$ converge a $\nu \in \mathcal{M}(X)$ si y sólo si
$$\int_X f\nu_n = \int_X f\nu, \quad \forall f \in V.$$
2. *a*) En la notación del ejercicio anterior, tome $X = \{z \in \mathbb{C} : |z| = 1\}$. Para cada $k \in \mathbb{Z}$, definimos $h_k : X \to \mathbb{C}$ por $h_k(x) = x^k$. Sea $H \subset C(X)$ la subálgebra generada por $\{h_k : k \in \mathbb{Z}\}$. Muestre que $H$ es densa en $C(X)$. Indicación: use el teorema de Stone-Weierstrass.
   *b*) (Criterio de Weyl) Sea $E_n \subset X$ una sucesión de conjuntos finitos. Deduzca que la familia $(E_n)$ se equidistribuye con respecto a $\nu_S$ si y sólo si para todo $k \in \mathbb{Z}$, con $k \neq 0$, se tiene
$$\lim_{n\to\infty} \frac{1}{\#E_n} \sum_{z\in E_n} z^k = 0.$$
   *c*) Utilizando el punto anterior, demuestre que la sucesión de conjuntos $\mu_n$ se equidistribuye con respecto a $\nu_S$.
3. El propósito de este ejercicio es demostrar el Corolario 4.14, usando los ejercicios anteriores.
   *a*) Sea $\mu : \mathbb{N} \to \{-1, 0, 1\}$ la función de Möbius, dada por
$$\mu(n) = \begin{cases} 0 & \text{si existe un entero } m \geq 2 \text{ tal que } m^2|n \\ (-1)^r & \text{si } n \text{ es el producto de } r \text{ primos distintos.} \end{cases}$$
   Sea $\delta_1 : \mathbb{N} \to \{0, 1\}$ dada por
$$\delta_1(n) = \begin{cases} 1 & \text{si } n = 1 \\ 0 & \text{si } n > 1. \end{cases}$$

Demuestre

$$\sum_{d|n} \mu(d) = \delta_1(n).$$

*b*) *Sumas de Ramanujan.* Denotamos por $(a, b)$ el máximo común divisor del par $a, b$. Para $n \in \mathbb{N}$ y $k \in \mathbb{Z} - \{0\}$, definimos

$$S(n,k) = \sum_{\substack{0 \le j \le n-1 \\ (n,j)=1}} \zeta_n^{jk}, \quad \zeta_n = e^{2\pi i/n}.$$

Muestre que

$$S(n,k) = \sum_{d|(n,k)} d\mu\Big(\frac{n}{d}\Big).$$

Indicación: escriba $S(n,k) = \sum_{j=0}^{n-1} \delta_1\big((n,k)\big)\zeta_n^{jk}$ y aplique el punto anterior.

*c*) Demuestre el Corolario 4.14.

**4.4. Energía.** Consideremos la medida $\nu_S$ definida en (4.8). Usando (4.3) y la definición de energía de una medida en (3.2), vemos que $I(\nu_S) = 0$. Además, para todo conjunto finito $G \subset \mathbb{C}$, se tiene $I\big(\nu(G)\big) = \infty$.

En vista del último ejemplo, es útil considera la cantidad $I'(G)$ definida en (3.3). El nexo entre la teoría de alturas y la teoría del potencial está dado por el siguiente

**Lema 4.16.** *(Comparación Energía-Altura) Sea $\alpha \in \overline{\mathbb{Q}}^*$ con $n := \deg(\alpha) \ge 2$. Para $a \in \mathbb{C}$ y $0 < \varepsilon < 1$ denotamos $D(a,\varepsilon) = \{\beta \in \mathrm{Gal}(\alpha) : |\beta - a| \le \varepsilon/2\}$ y $n_\varepsilon = \#D(a,\varepsilon)$. Entonces se tiene*

$$|\log \varepsilon| \cdot \frac{n_\varepsilon(n_\varepsilon - 1)}{n(n-1)} - \frac{2n}{n-1}\tilde{h}(2\alpha) \le I'\big(\mathrm{Gal}(\alpha)\big) \le 3h(\alpha).$$

*Aquí,* $\tilde{h}(2\alpha) := \frac{1}{n}\sum_{\beta \in \mathrm{Gal}(\alpha)} \log^+(|2\beta|)$.

*Demostración.* Primero demostraremos la cota superior. Sea $a_n$ el coeficiente dominante de $g_\alpha$ y escribimos $\mathrm{Gal}(\alpha) = \{\alpha_1, \dots, \alpha_n\}$. Como $g_\alpha$ es irreductible, tenemos $D(g_\alpha) \ne 0$. Además, $a_n D(g_\alpha) \in \mathbb{Z}$ (Proposición 2.25), de donde

$$\log(a_n |D(g_\alpha)|) \ge \log 1 = 0. \tag{4.9}$$

Se tiene

$$\begin{aligned} 0 \le \log a_n + \log|D(g_\alpha)| &= (2n-1)\log a_n + 2\sum_{i<j} \log|\alpha_i - \alpha_j| \\ &= (2n-1)\log a_n + \sum_{i \ne j} \log|\alpha_i - \alpha_j| \\ &= (2n-1)\log a_n - n(n-1)I'\left(\mathrm{Gal}(\alpha)\right). \end{aligned}$$

Por otro lado, de (4.4) tenemos

$$\log a_n = nh(\alpha) - \sum_i \log^+|\alpha_i| \le nh(\alpha).$$

Combinando ambas desigualdades se obtiene la cota superior.

Ahora demostraremos la cota inferior. Escribimos $I'(\alpha) = A + B + C$, con

$$A = \frac{1}{n(n-1)} \sum_{\substack{|\alpha_i - \alpha_j| \le \varepsilon \\ i \neq j}} \log \frac{1}{|\alpha_i - \alpha_j|},$$

$$B = \frac{1}{n(n-1)} \sum_{\varepsilon < |\alpha_i - \alpha_j| \le 1} \log \frac{1}{|\alpha_i - \alpha_j|},$$

$$C = \frac{1}{n(n-1)} \sum_{|\alpha_i - \alpha_j| > 1} \log \frac{1}{|\alpha_i - \alpha_j|}.$$

Claramente $B \ge 0$. Además, como $\alpha_i, \alpha_j \in D(a, \varepsilon)$ implica $|\alpha_i - \alpha_j| \le \varepsilon$, deducimos

$$A \ge \log(1/\varepsilon) \frac{n_\varepsilon(n_\varepsilon - 1)}{n(n-1)}.$$

Para acotar inferiormente $C$, supondremos $|\alpha_1| \le |\alpha_2| \le \cdots \le |\alpha_n|$. Entonces

$$\begin{aligned}
\sum_{|\alpha_i - \alpha_j| > 1} \log |\alpha_i - \alpha_j| &= 2 \sum_{i<j} \log^+ |\alpha_i - \alpha_j| \\
&= 2 \sum_{i=1}^{n} \sum_{j=i+1}^{n} \log^+ |\alpha_i - \alpha_j| \\
&\le 2 \sum_{i=1}^{n} \sum_{j=i+1}^{n} \log^+ |2\alpha_j| \\
&\le 2 \sum_{i=1}^{n} n \tilde{h}(2\alpha) \\
&= 2n^2 \tilde{h}(2\alpha),
\end{aligned}$$

de donde se obtiene $C \ge -\frac{2n}{n-1} \tilde{h}(2\alpha)$. Reuniendo las cotas inferiores para $A, B$ y $C$, se obtiene el resultado. □

Enunciamos aquí un resultado que caracteriza la medida de Lebesgue sobre el círculo en términos del funcional de energía.

**Proposición 4.17.** *Sea $\mathcal{M}(S)$ el conjunto de todas las medidas en $\mathcal{M}$ soportadas en $\{z \in \mathbb{C} : |z| = 1\}$. Entonces*

$$I(\nu) \ge 0, \quad \forall \nu \in \mathcal{M}(S).$$

*Más aún,*

$$\nu \in \mathcal{M}(S), \quad I(\nu) = 0 \Leftrightarrow \nu = \nu_S.$$

*Demostración.* Usar el Ejercicio 3.26 (1) de la Sección 3. Alternativamente, ver [16], Ex. 1.10. □

**4.5. Demostración del Teorema de Bilu.** En toda esta sección asumimos que la sucesión $(\alpha_n) \subset \overline{\mathbb{Q}}^*$ es genérica y cumple

$$h(\alpha_n) \to 0, \quad n \to \infty.$$

Unidas al Ejercicio 4.8, estas hipótesis aseguran que $\deg(\alpha_n)$ tiende a infinito con $n$.

**Lema 4.18.** *(Equidistribución en radio). Para cada $n \in \mathbb{N}$, se puede escoger un conjunto $E_n \subseteq \mathrm{Gal}(\alpha_n)$ de manera que*

1. *se tiene*
$$\lim_{n\to\infty} \frac{\#E_n}{\#\operatorname{Gal}(\alpha_n)} = 1.$$
2. *Para todo $\varepsilon > 0$, existe $n_0$ tal que para todo $n \geq n_0$ se tiene*
$$z \in E_n \Rightarrow 1 - \varepsilon \leq |z| \leq 1 + \varepsilon.$$

*Demostración.* Notar que por el Lema 4.5, también tenemos
$$h(1/\alpha_n) \to 0, \quad n \to \infty.$$
Procederemos por contradicción. Pasando a una subsucesión si fuese necesario, podemos suponer que existen subconjuntos $V_n \subseteq \operatorname{Gal}(\alpha_n)$ tales que

- $\lim_{n\to\infty} \frac{\#V_n}{\#\operatorname{Gal}(\alpha_n)} = c > 0$.
- existe $a > 1$ tal que para todo $n$, se tiene
$$z \in V_n \Rightarrow |z| > a \text{ ó } \left|\frac{1}{z}\right| > a.$$

Descomponemos $V_n = A_n \cup B_n$, donde
$$z \in A_n \Leftrightarrow |z| > a, \qquad z \in B_n \Leftrightarrow \left|\frac{1}{z}\right| > a.$$
Usando la fórmula de Jensen tenemos
$$h(\alpha_n) \geq \frac{1}{\deg(\alpha_n)}\Big(\sum_{\beta\in A_n} \log|\beta|\Big) \geq \frac{\#A_n}{\deg(\alpha_n)} \log a.$$
Por otro lado, usando el Ejercicio 2.21 (2) también se tiene
$$h(1/\alpha_n) \geq \frac{1}{\deg(\alpha_n)}\Big(\sum_{\beta\in B_n} \log|1/\beta|\Big) \geq \frac{\#B_n}{\deg(\alpha_n)} \log a.$$
Luego
$$h(\alpha_n) + h(1/\alpha_n) \geq \frac{\#V_n}{\deg \alpha_n} \log a.$$
Tomando el límite cuando $n \to \infty$, obtenemos
$$0 \geq c \log a > 0,$$
lo que es absurdo. □

**Fin de la demostración del Teorema de Bilu**: la condición (2) en el Lema 4.18, asegura que existe un compacto $K \subset \mathbb{C}$ que contiene al círculo unitario tal que $\nu(E_n) \in \mathcal{M}(K)$, para todo $n$. Como $\mathcal{M}(K)$ es secuencialmente compacto (cf. Teorema 3.9), tenemos que la sucesión $\nu(E_n)$ admite una subsucesión $\nu(E_{n_j})$ convergente a una medida $\nu \in \mathcal{M}(K)$. Notar que la condición (1) en el Lema 4.18 asegura que $\nu\big(\operatorname{Gal}(\alpha_{n_j})\big)$ también converge a $\nu$. Ahora estableceremos que la diagonal tiene medida nula con respecto a la medida producto $\nu \times \nu$ (para ver generalidades y referencias sobre medidas producto revisar la demostración del Teorema 3.22).

**Afirmación.** *Sea $D = \{(x, x) : x \in \mathbb{C}\} \subseteq \mathbb{C}^2$. Entonces $\nu \times \nu(D) = 0$.*

*Demostración.* Para demostrar la afirmación, razonaremos por contradicción. Por Fubini-Tonelli, tenemos

$$\nu \times \nu(D) = \int_{\mathbb{C}} \nu(\{x\}) d\nu(x).$$

Luego si $\nu \times \nu(D) > 0$, entonces existe $a \in \mathbb{C}$ tal que $\nu(\{a\}) > 0$. En particular, existe $\kappa > 0$ y un conjunto infinito $J \subset \mathbb{N}$ tal que para todo $n \in J$ y todo $\varepsilon > 0$ suficientemente pequeño, se cumple

$$\frac{\#\{\beta \in \mathrm{Gal}(\alpha_n) : |\beta - a| \leq \varepsilon/2\}}{\deg \alpha_n} \geq \kappa.$$

Usando el Lema 4.16, deducimos que para $n \in J$ se tiene

$$|\log \varepsilon| \cdot \kappa^2 - \frac{2\deg(\alpha_n)}{\deg(\alpha_n) - 1} \tilde{h}(2\alpha_n) \leq 3h(\alpha_n). \tag{4.10}$$

Es sencillo ver que $\tilde{h}(2\alpha_n) \leq h(\alpha_n) + \log 2$. Como $\lim_{n\to\infty} h(\alpha_n) = 0$, vemos que $\tilde{h}(2\alpha_n)$ es acotado. De la desigualdad (4.10) deducimos que existe $A \in \mathbb{R}$ tal que para todo $\varepsilon$ suficientemente pequeño, se tiene $|\log \varepsilon| \cdot \kappa^2 \leq A$. Esto es absurdo, luego la afirmación es verdadera. ■

Ahora demostraremos que $\nu = \nu_S$. Primero notemos que usando la condición (2) en el Lema 4.18, podemos concluir que $\nu$ está soportada en el círculo unitario. De la Proposición 4.17, deducimos que $I(\nu) \geq 0$. Sea $\phi : \mathbb{R}_{\geq 0} \to \mathbb{R}$ una función continua no decreciente tal que

- $0 \leq \phi \leq 1$,
- $\phi(t) = 0$ si $t \leq \frac{1}{2}$,
- $\phi(t) = 1$ si $t \geq 1$.

Para $\varepsilon > 0$, definimos $\phi_\varepsilon(t) = \phi(t/\varepsilon)$. Entonces, para todo $(z, t) \in \mathbb{C}^2 - D$, se tiene

$$\lim_{\varepsilon\to 0} \phi_\varepsilon(|z - t|) \log \frac{1}{|z - t|} = \log \frac{1}{|z - t|}.$$

Como $D$ es un conjunto de medida nula según $\nu \times \nu$, y $\phi$ es no decreciente, el Teorema de convergencia monótona asegura que

$$\begin{aligned} I(\nu) &= \int\int \log \frac{1}{|z - t|} d\nu(z) d\nu(t) \\ &= \lim_{\varepsilon\to 0} \int\int \phi_\varepsilon(|z - t|) \log \frac{1}{|z - t|} d\nu(z) d\nu(t). \end{aligned}$$

Sea

$$S_{j,\varepsilon} = \frac{1}{|E_{n_j}|^2} \sum_{\beta,\beta' \in E_{n_j}} \phi_\varepsilon(|\beta - \beta'|) \log \frac{1}{|\beta - \beta'|}.$$

Por convergencia débil tenemos

$$\int\int \phi_\varepsilon(|z - t|) \log \frac{1}{|z - t|} d\nu(z) d\nu(t) = \lim_{j\to\infty} S_{j,\varepsilon}.$$

Si $\varepsilon$ es suficientemente pequeño, tenemos que $|\beta - \beta'| \leq \varepsilon \Rightarrow \log \frac{1}{|\beta - \beta'|} \geq 0$. Además, $|\beta - \beta'| \geq \varepsilon \Rightarrow \phi_\varepsilon(|\beta - \beta'|) = 1$. Como $0 \leq \phi_\varepsilon \leq 1$ en todos los casos,

concluimos que

$$S_{j,\varepsilon} \leq \frac{1}{|E_{n_j}|^2} \sum_{\substack{\beta,\beta' \in E_{n_j} \\ \beta \neq \beta'}} \log \frac{1}{|\beta - \beta'|}.$$

Combinando las estimaciones anteriores, deducimos que

$$\begin{aligned} I(\nu) &\leq \liminf \frac{1}{|E_{n_j}|^2} \sum_{\substack{\beta,\beta' \in E_{n_j} \\ \beta \neq \beta'}} \log \frac{1}{|\beta - \beta'|} \\ &= \liminf I'(E_{n,j}) \\ &= \liminf I'\big(\operatorname{Gal}(\alpha_{n_j})\big) \qquad \text{condición (1) en el Lema 4.18} \\ &\leq 3 \liminf h(\alpha_{n_j}) \qquad \text{Lema 4.16} \\ &= 0. \end{aligned}$$

Por lo tanto, $I(\nu) = 0$. Aplicando nuevamente la Proposición 4.17, deducimos $\nu = \nu_S$.

Hemos establecido que toda subsucesión convergente de $\nu(E_n)$ tiene como límite a la medida $\nu_S$. Esto muestra que $\nu(E_n)$ converge a $\nu_S$. Luego $\nu\big(\operatorname{Gal}(\alpha_n)\big)$ converge a $\mu_S$, terminando la demostración del Teorema de Bilu.

## 5. Generalizaciones y aplicaciones diofantinas

En esta sección haremos algunos comentarios sobre extensiones de los teoremas vistos en el curso, así como aplicaciones.

*Varias variables.* El teorema que Bilu demuestra en su artículo [5] es más general que lo que hemos formulado aquí, pues trata el caso de un número finito arbitrario de variables. La demostración en el caso de dimensión general se reduce al caso de dimensión 1 por un argumento estándar de análisis de Fourier.

*Versiones p-ádicas.* Los números complejos tienen un análogo $p$-ádico: si $p$ es un número primo y $\mathbb{Q}_p$ el cuerpo de los números $p$-ádicos, entonces norma $p$-ádica admite una única extensión a la cerradura algebraica $\overline{\mathbb{Q}_p}$. Así, podemos considerar su completación, denotada $\widehat{\overline{\mathbb{Q}_p}}$. El cuerpo resultante no es algebraicamente cerrado. Denotamos por $\mathbb{C}_p$ a la cerradura algebraica de $\widehat{\overline{\mathbb{Q}_p}}$. Este último cuerpo es a la vez completo y algebraicamente cerrado. El lector podrá consultar en [15] el desarrollo de los conceptos y relaciones que dan lugar a una teoría del potencial sobre $\mathbb{C}_p$, análoga a la teoría sobre $\mathbb{C}$ que hemos revisado. También encontrará una formulación del Teorema de Fekete-Szegö en tal contexto. El Teorema de Bilu también tiene un análogo $p$-ádico, ver [2], [7] o [8].

*Relación con problemas diofantinos* Una de las principales motivaciones para estudiar teoremas de equidistribución aritméticos es la posibilidad de aplicarlos en cuestiones diofantinas. Un ejemplo básico es el siguiente: decimos que $(\alpha, \beta) \in \mathbb{C}^2$ es un punto de torsión si $\alpha$ y $\beta$ son raíces de la unidad. Una subvariedad de torsión de $\mathbb{C}^2$ es un conjunto de la forma

$$\{x, y : x^n y^m = \zeta\},$$

donde $m, n$ son números enteros, no ambos cero y $\zeta \in \mathbb{C}$ es una raíz de la unidad.

**Teorema 5.1.** *Sea $X \subseteq \mathbb{C}^2$ un subconjunto Zariski cerrado propio. Entonces $X$ contiene infinitos puntos de torsión si y sólo si $X$ contiene una subvariedad de torsión.*

Este resultado puede deducirse del Teorema de Bilu (ver [5], Theorem 5.1) y es un caso particular de la conjetura de Manin-Mumford. Para mayor información sobre el nexo entre teoremas de equidistribución y problemas de geometría diofantina se recomienda ver [19], [20].

## Referencias


[1] P. Autissier, *Autour du théorème de Fekete-Szegö* , École d'été de géométrie d'Arakelov, Grenoble, 2017. `https://www.math.u-bordeaux.fr/~pautissi/`

[2] M. Baker, R. Rumely, *Equidistribution of small points, rational dynamics, and potential theory*, Ann. Inst. Fourier (Grenoble) **56** (2006), no. 3, 625–688

[3] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[4] P. Billingsley. *Probability and Measure*. John Wiley & Sons, Inc., New York, second edition, 1986.

[5] Y. Bilu. Limit distribution of small points on algebraic tori. *Duke Math. J.*, 89(3):465–476, 1997.

[6] E. Bombieri and W. Gubler. *Heights in Diophantine geometry*, volume 4 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2006.

[7] A. Chambert-Loir, *Mesures et équidistribution sur les espaces de Berkovich*, J. Reine Angew. Math. **595** (2006), p. 215–235

[8] Ch. Favre, J. Rivera-Letelier, *Équidistribution quantitative des points de petite hauteur sur la droite projective*, Math. Ann. **335** (2006), no. 2, p. 311–361

[9] M. Fekete, *Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten*, Math. Z. **17** (1923), p. 228–249.

[10] M. Fekete, G. Szegö, *On algebraic equations with integral coeffients whose roots belong to a given point set*, Math. Z. **63** (1955), 158–172.

[11] S. Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2002.

[12] D. H. Lehmer. Factorization of certain cyclotomic functions. *Ann. of Math. (2)*, 34(3):461–479, 1933.

[13] T. Ransford, *Potential theory in the complex plane*, Cambridge Univ. Press, 1995.

[14] W. Rudin, *Functional analysis* McGraw Hill, 1973.

[15] R. Rumely. *Capacity theory on algebraic curves*, volume 1378 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1989.

[16] E. B. Saff, *Logarithmic potential theory with applications to approximation theory*, Surv. Approx. Theory **5** (2010), 165–200.

[17] E. B. Saff, V. Totik, *Logarithmic potential with external fields*, Grundlehren der mathematischen Wissenschaften **316**, Springer Verlag 1997.

[18] T. Tao, *An introduction to measure theory*, Graduate Studies in Mathematics **126**, AMS 2011.

[19] E. Ullmo, *Manin-Mumford, André-Oort, the equidistribution point of view*, Equidistribution in number theory, an introduction, 103–138, NATO Sci. Ser. II Math. Phys. Chem., 237, Springer, Dordrecht, 2007

[20] E. Ullmo, *Théorie ergodique et géométrie arithmétique*, Proceedings of the International Congress of Mathematicians, Vol. II (Beijing, 2002), 197–206.



ICMAT (CSIC UAM UCM UC3). Nicolás Cabrera 13, 28049 Madrid, España.
*Email address*: `burgos@icmat.es`

Pontificia Universidad Católica de Chile. Vicuña Mackenna 4860, Macul, Santiago, Chile
*Email address*: `rmenares@mat.uc.cl`

Curso

# REPRESENTACIONES DE GALOIS

LUIS DIEULEFAIT, ARIEL PACETTI, Y FERNANDO RODRIGUEZ VILLEGAS

Publicaciones Matemáticas del Uruguay
Volumen 18, Diciembre 2023, Páginas 105–161
S 0797-1443

# REPRESENTACIONES DE GALOIS

LUIS DIEULEFAIT, ARIEL PACETTI, Y FERNANDO RODRIGUEZ VILLEGAS

## Índice



Versión final: 27 de diciembre de 2019.

Estas notas corresponden al curso dictado por los autores en la escuela AGRA III, Aritmética, Grupos y Análisis, del 9 al 20 de Julio de 2018 en Córdoba, Argentina.

## 1. Teoría de Cuerpos

**1.1. Generalidades.** Recordemos la siguiente definición.

**Definición 1.1.** Un cuerpo es una terna $(K, +, *)$ dada por un conjunto $K$, y dos operaciones binarias

$$+, * : K \times K \mapsto K,$$

que satisfacen las siguientes propiedades:

- $(K, +)$ es un grupo abeliano (o sea la operación es asociativa, conmutativa, posee un neutro denotado 0 y todo elemento tiene inverso).
- $(K \setminus \{0\}, *)$ es un grupo abeliano.
- Vale la propiedad distributiva del producto sobre la suma, o sea

$$a * (b + c) = a * b + a * c,$$

para toda terna de elementos $a, b, c$ en $K$.

**Ejemplos 1.2.**

1. El conjunto de números racionales con sus operaciones naturales es un cuerpo $(\mathbb{Q}, +, *)$. Lo mismo sucede con el conjunto de números reales $(\mathbb{R}, +, *)$ y de números complejos $(\mathbb{C}, +, *)$.
2. Otros cuerpos de gran utilidad son aquellos para los cuales el conjunto $K$ es finito. Por ejemplo, si $p$ es un número primo, y denotamos por $\mathbb{F}_p = \mathbb{Z}/p$ el conjunto de clases de equivalencia de números enteros módulo $p$ (donde identificamos dos números enteros si su diferencia es divisible por $p$), entonces el conjunto $(\mathbb{F}_p, +, *)$ es un cuerpo con $p$ elementos.

Un *morfismo de cuerpos* es simplemente un morfismo de anillos, o sea si $K$ y $L$ son cuerpos, un morfismo de cuerpos $\varphi : K \to L$ es una función que satisface:

- $\varphi(x + y) = \varphi(x) + \varphi(y)$ para todo par de elementos $x, y \in K$.
- $\varphi(x * y) = \varphi(x) * \varphi(y)$ para todo par de elementos $x, y \in K$.
- $\varphi(1) = 1$.

De manera usual, un *isomorfismo* de cuerpos, es un morfismo de cuerpos biyectivo (es fácil ver que todo morfismo de cuerpos es inyectivo).

Recordar que si $K$ es un cuerpo, entonces el anillo de polinomios $K[x]$ posee un algoritmo de división; esto es dados dos polinomios $f(x), g(x) \in K[x]$, con $g(x)$ no nulo, existen únicos polinomios $q(x), r(x) \in K[x]$ con $r(x) = 0$ o de grado menor que el grado de $g(x)$ tales que

$$f(x) = g(x)q(x) + r(x).$$

Dicho algoritmo permite definir el máximo común divisor de polinomios de manera análoga a lo hecho para números enteros (imponiendo la condición de que el máximo común divisor es un polinomio mónico, para obtener unicidad), y demostrar que dados dos polinomios $f(x), g(x) \in K[x]$, alguno no nulo, el máximo común divisor entre ellos se escribe como combinación lineal de ambos, o sea: existen $r(x), s(x) \in K[x]$ tales que

$$\gcd(f(x), g(x)) = r(x)f(x) + s(x)g(x).$$

**Proposición 1.3.** *Sea $K$ un cuerpo, y sea $p(x) \in K[x]$ un polinomio irreducible. Entonces el anillo $L = K[x]/(p(x))$ es un cuerpo.*

*Demostración.* Como estamos cocientando por un ideal (el generado por $p(x)$), el cociente tiene automáticamente una estructura de anillo. Lo que precisamos demostrar para obtener un cuerpo es que todo elemento no nulo tiene inverso multiplicativo. Sea así $\overline{f(x)}$ la clase de representantes de un elemento no nulo del cociente, y sea $f(x) \in K[x]$ algún polinomio en dicha clase. Como $\overline{f(x)}$ es no nulo, $p(x) \nmid f(x)$. Como $p(x)$ es irreducible, $\gcd(f(x), p(x)) = 1$, con lo cual existen $r(x), s(x) \in K[x]$ tales que

$$1 = r(x)f(x) + s(x)p(x). \tag{1.1}$$

Entonces en el cociente, la clase de $r(x)$ es un inverso multiplicativo de la clase de $f(x)$. □

Así obtenemos muchos ejemplos nuevos de cuerpos. Por ejemplo:

- Tomando $p(x) = x^2 + 1 \in \mathbb{Q}[x]$ obtenemos el cuerpo $\mathbb{Q}[x]/(x^2+1)$ que es isomorfo (como cuerpo) al cuerpo $\mathbb{Q}[i] = \{a + bi \ : \ a, b \in \mathbb{Q}\}$ (donde $i^2 = -1$).
- Tomando $p(x) = x^2 + 1$ en $\mathbb{F}_3[x]$ (verificar que es irreducible), obtenemos el cuerpo $\mathbb{F}_3[x]/(x^2+1)$, que denotamos $\mathbb{F}_9$. ¿Cuántos elementos tiene dicho cuerpo?
- Más generalmente, si $p$ es un número primo, y si $q(x) \in \mathbb{F}_p[x]$ es un polinomio irreducible de grado $d$, ¿ cuántos elementos tiene el cuerpo $\mathbb{F}_p[x]/(q(x))$?

Notar que si $K$ y $L$ son cuerpos y $K \subset L$, podemos pensar a $L$ como un $K$-espacio vectorial. En particular tiene sentido mirar la dimensión de $L$ como $K$-espacio vectorial y preguntarse si es finita o no.

**Definición 1.4.** Si $K$ y $L$ son cuerpos y $K \subset L$, decimos que $L$ es una extensión de cuerpos de $K$. Definimos el *grado* de la extensión, y lo denotamos $[L : K]$, como la dimensión de $L$ como $K$-espacio vectorial.

*Ejercicio* 1.5. Sea $p(x) \in K[x]$ un polinomio irreducible de grado $d$, y denotemos por $L = K[x]/(p(x))$. Calcular $[L : K]$.

**Definición 1.6.** Un cuerpo de números es un cuerpo $K$ tal que $\mathbb{Q} \subset K$, y el grado de la extensión es finito.

El siguiente resultado vale en contextos muy generales, pero lo enunciamos solamente para el caso de cuerpos de números, que es en el que lo vamos a usar.

**Teorema 1.7** (Teorema de elementos primitivos)**.** *Si $K$ es un cuerpo de números, y $[K : \mathbb{Q}] = d$, entonces existe un polinomio $p(x) \in \mathbb{Q}[x]$ irreducible de grado $d$ tal que $K$ es isomorfo a $\mathbb{Q}[x]/(p(x))$.*

El elemento de $K$ que se corresponde con la variable $x$ bajo el isomorfismo se suele llamar un *elemento primitivo* de $K$.

**1.2. Los números $p$-ádicos.** Una referencia clásica sobre los números $p$-ádicos es el libro ([6]). Existen otros cuerpos además de los mencionados anteriormente que juegan un rol preponderante en la teoría de números. Dichos cuerpos son los que se obtienen a partir de completar el cuerpo de números racionales con un valor absoluto (o sea son construcciones similares a la de los números reales). La diferencia, es que en lugar de utilizar el valor absoluto usual (llamado arquimediano), es preciso utilizar otros valores absolutos (los no-arquimedianos).

Fijemos un primo $p$. Definimos la valuación $p$-ádica como la función $v_p : \mathbb{Z}\setminus\{0\} \mapsto \mathbb{Z}$ dada por

$$v_p(a) = \text{máx}\{n \in \mathbb{N} \cup \{0\} \ : \ p^n \mid a\} \tag{1.2}$$

Así, $v_3(12) = 1$ y $v_3(10) = 0$. Extendemos la valuación a los racionales no nulos, por

$$v_p\left(\frac{a}{b}\right) = v_p(a) - v_p(b).$$

**Definición 1.8.** El valor absoluto $p$-ádico esta dado por:

$$\left|\frac{a}{b}\right|_p := \begin{cases} 0 & \text{si } \frac{a}{b} = 0, \\ p^{-v_p(\frac{a}{b})} & \text{en otro caso.} \end{cases}$$

Luego, por ejemplo $\left|\frac{2}{3}\right|_3 = 3$, y $\left|\frac{9}{7}\right|_3 = \frac{1}{9}$. Es fácil verificar las siguientes propiedades.

**Proposición 1.9.** *El valor absoluto $p$-ádico satisface las siguientes propiedades:*

1. $|a|_p = 0$ *si y sólo si* $a = 0$.
2. $|a \cdot b|_p = |a|_p \cdot |b|_p$.
3. $|a + b|_p \leq \text{máx}\{|a|_p, |b|_p\}$. *Más aún, si* $|a|_p$ *y* $|b|_p$ *son distintos, entonces vale la igualdad.*

La última propiedad de la Proposición se conoce como la propiedad *ultramétrica*. Dicha propiedad implica la desigualdad triangular (dejamos esto como ejercicio), pero es mucho más fuerte. En particular, el valor absoluto $p$-ádico nos da una manera de medir qué tan cerca están dos números racionales. Notar que esta forma de medir dista bastante de la forma usual, por ejemplo, el número $p^n$ tienen valor absoluto $p$-ádico muy pequeño cuando $n$ es grande ($|p^n|_p = \frac{1}{p^n}$). En particular, dos números enteros están cerca con el valor absoluto $p$-ádico si su diferencia es divisible por potencias grandes de $p$.

*Observación* 1.10. El valor absoluto tiene una propiedad muy importante: el conjunto de valores que toma tiene como único punto de acumulación el cero. De la definición se puede ver fácilmente que

$$|\mathbb{Q}|_p = \{0\} \cup \{p^{\mathbb{Z}}\}.$$

Así por ejemplo, si $p = 2$ y un número racional tiene valor absoluto 2-ádico mayor que 1, automáticamente dicho valor es mayor o igual que 2 (dado que en el intervalo $(1, 2)$ no hay ningún valor posible).

**Definición 1.11.** Definimos el *cuerpo de números $p$-ádicos* y lo denotamos $\mathbb{Q}_p$ al conjunto obtenido al completar el conjunto de números racionales $\mathbb{Q}$ con respecto al valor absoluto $p$-ádico. Dicho conjunto tiene una suma y un producto que lo hacen un cuerpo (ver Ejercicio 1.12).

Recordemos el proceso para completar un espacio métrico $\mathcal{B}$ con respecto a su valor absoluto $|\,.\,|$: consideramos sucesiones de Cauchy en $\mathcal{B}$ (con respecto al valor absoluto $|\,.\,|$), y definimos una relación de equivalencia, donde identificamos dos sucesiones $\{a_n\}$ y $\{b_n\}$ (y notamos $\{a_n\} \sim \{b_n\}$) si su diferencia $\{a_n - b_n\}$ tiende a cero. Luego la completación $\overline{\mathcal{B}}$ se define como el cociente del conjunto de sucesiones de Cauchy por la relación $\sim$.

*Ejercicio* 1.12. Probar que si $(\mathcal{B}, +, *)$ es un cuerpo con un valor absoluto, existe una forma natural de extender la suma y producto a la completación $\overline{\mathcal{B}}$, que hacen de $(\overline{\mathcal{B}}, +, *)$ un cuerpo.

A la vez, existe una manera natural de extender el valor absoluto $|\,.\,|$ a la completación $\overline{\mathcal{B}}$. En particular, $\mathbb{Q}_p$ también posee un valor absoluto $p$-ádico.

*Ejercicio* 1.13.
- Probar que si $(x_n) \subset \mathbb{Q}$ es una sucesión de Cauchy entonces la sucesión de valores absolutos $|x_n|_p$ es de Cauchy. En particular, existe $\lim_n |x_n|_p$.
- Probar que si $(x_n), (y_n) \subset \mathbb{Q}$ son dos sucesiones de Cauchy equivalentes, $\lim_n |x_n|_p = \lim_n |y_n|_p$.
- Deducir que si $x \in \mathbb{Q}_p$, podemos definir $|x|_p := \lim_n |x_n|_p$, donde $(x_n)$ es cualquier sucesión de Cauchy que converge a él.
- Probar que el valor absoluto definido en un elemento $x \in \mathbb{Q}$ coincide con el valor absoluto $p$-ádico.

*Ejercicio* 1.14. Demostrar las siguientes propiedades del valor absoluto:

1. El conjunto de valores que alcanza el valor absoluto de $\mathbb{Q}_p$ es el mismo que el de $\mathbb{Q}$, o sea $|\mathbb{Q}_p|_p = |\mathbb{Q}|_p = \{0\} \cup \{p^{\mathbb{Z}}\}$.
2. Probar que las bolas abiertas en $\mathbb{Q}_p$ son cerradas, y las bolas cerradas son abiertas; esto es: si $x_0 \in \mathbb{Q}_p$ y $r > 0$, la bola de centro $x_0$ y radio $r$ es
$$B_r(x_0) = \{y \in \mathbb{Q}_p \ : \ |x_0 - y|_p < r\}.$$
Probar que dado $r > 0$, existe un número $r' > 0$ tal que $B_r(x_0) = \overline{B_{r'}(x_0)}$. A la vez, dado $r' > 0$, existe un $r > 0$ tal que $\overline{B_{r'}(x_0)} = B_r(x_0)$. Esto en particular implica que como espacio métrico, $\mathbb{Q}_p$ es totalmente disconexo.

**Definición 1.15.** El conjunto de los *enteros p-ádicos* se define como la bola cerrada de radio 1 centrada en cero, esto es $\mathbb{Z}_p := \{x \in \mathbb{Q}_p \ : \ |x|_p \leq 1\}$.

*Ejercicio* 1.16. Probar que el conjunto $\mathbb{Z}_p$ es un anillo (sugerencia: usar la propiedad de que el valor absoluto $p$-ádico es una ultramétrica, ver Proposición 1.9). A la vez, probar que $\mathbb{Z} \subset \mathbb{Z}_p$, y $\mathbb{Z}_p$ es la clausura topológica de $\mathbb{Z}$ en $\mathbb{Q}_p$.

Existe una manera de entender los enteros $p$-ádicos, mediante la expansión en base $p$. Recordar que todo número natural $N \in \mathbb{N}$ se puede expresar de forma única como

$$N = a_0 + a_1 p + \ldots + a_r p^r, \tag{1.3}$$

donde $0 \leq a_i < p$, para $i = 0, 1, \ldots r$ (esta es la expresión en base $p$ de $N$). A la vez, si $N_1, N_2$ son dos números naturales tales que $v_p(N_1 - N_2) \geq t$, entonces los primeros $t+1$ términos de la expresión en base $p$ de $N_1$ y $N_2$ coinciden.

*Ejercicio* 1.17. Probar que si $\{b_n\}$ es una sucesión de Cauchy para el valor absoluto $p$-ádico (donde cada $b_n$ es un número natural), existen únicos $(a_n)_{n \geq 0}$ con $0 \leq a_n < p$ tales que la sucesión $\{b_n\}$ converge a la serie

$$\sum_{n \geq 0} a_n p^n \in \mathbb{Q}_p. \tag{1.4}$$

Esto sugiere que considerar clases de equivalencia de sucesiones de Cauchy para el valor absoluto $p$-ádico de números naturales es equivalente a considerar series como en (1.4). Es fácil ver que dada una tal serie, define una sucesión de Cauchy de

números enteros (por ejemplo tomando la sucesión cuyo término $n$-ésimo consiste en la suma de los primeros $n$-términos). Luego resta entender que sucede al considerar sucesiones de números racionales.

*Ejercicio* 1.18. Probar que si $x \in \mathbb{Q}_p$, entonces podemos representar el número $x$ de manera única como una serie de la forma

$$\sum_{i \geq N_0} a_i p^i,$$

donde $0 \leq a_i < p$ y $N_0 \in \mathbb{Z}$. A la vez, probar que $\mathbb{Z}_p$ se puede caracterizar como aquellas series que no poseen términos no nulos con índices negativos, o sea si $i < 0$ entonces $a_i = 0$.

*Observación* 1.19. Esta manera de describir los números $p$-ádicos es muy útil para entender propiedades de los mismos, pero poco eficiente para operar. No entraremos en detalles de los problemas computacionales que aparecen naturalmente al trabajar con los números $p$-ádicos (ver por ejemplo [6]).

**1.3. Teoría de Galois en extensiones finitas.** El propósito del presente curso (y sus notas) no es dar un exposición exhaustiva sobre teoría de Galois, pero precisamos repasar algunas propiedades importantes de la misma (para mas detalles, ver por ejemplo los libros [14], [23] o [1]).

Si $K$ es un cuerpo, posee un neutro para el producto que denotamos por 1. Tenemos un único morfismo natural de anillos $\psi : \mathbb{Z} \mapsto K$, determinado por $1 \to 1$. El núcleo de $\psi$ es un ideal $\mathfrak{a}$ de $\mathbb{Z}$ y obtenemos una inclusión

$$\psi : \mathbb{Z}/\mathfrak{a} \hookrightarrow K,$$

con lo cual $\mathbb{Z}/\mathfrak{a}$ debe ser un dominio íntegro, y por lo tanto $\mathfrak{a}$ es un ideal primo. Existen dos tipos de ideales primos en $\mathbb{Z}$:

- $\mathfrak{a} = \{0\}$,
- $\mathfrak{a} = p\mathbb{Z}$, donde $p$ es un número primo.

Decimos que el cuerpo $K$ tiene *característica cero* si $\psi$ es inyectiva, y decimos que $K$ tiene *característica* $p$ si el núcleo de $\psi$ es el ideal $p\mathbb{Z}$.

**Ejemplos 1.20.**
1. El cuerpo de números racionales $\mathbb{Q}$ tiene característica 0. Lo mismo sucede con el cuerpo $\mathbb{R}$ de números reales.
2. El cuerpo finito $\mathbb{F}_p$ tiene característica $p$.

En este curso solamente consideraremos cuerpos $K$ que tengan característica 0 o que sean finitos, con lo cual de ahora en más nos vamos a restringir a dichos cuerpos.

**Definición 1.21.** Sean $K \subset L$ cuerpos. Decimos que $\alpha \in L$ es *algebraico* sobre $K$ si existe un polinomio mónico $p(x) \in K[x]$ tal que $p(\alpha) = 0$. Decimos que la extensión $L/K$ es *algebraica* si todos los elementos de $L$ son algebraicos sobre $K$.

**Ejemplos 1.22.** Consideremos la extensión $\mathbb{Q} \subset \mathbb{R}$.
1. El elemento $\sqrt{2} \in \mathbb{R}$ es algebraico sobre $\mathbb{Q}$, dado que es raíz del polinomio $x^2 - 2 \in \mathbb{Q}[x]$.
2. El número $\pi$ es trascendente, o sea no es raíz de ningún polinomio $p(x) \in \mathbb{Q}[x]$. Luego $\mathbb{R}/\mathbb{Q}$ no es una extensión algebraica (buscar información en Wikipedia).

*Ejercicio* 1.23. Probar que si $[L : K] < \infty$, entonces la extensión $L/K$ es algebraica.

Consideremos en $\mathbb{Q}[x]$ el polinomio $p(x) = x^2 - 2$. Claramente dicho polinomio es irreducible (dado que si no lo fuera se podría escribir como producto de dos polinomios de grado 1, y en particular tendría una raíz racional, que claramente no tiene). Nos podemos preguntar cuál es el menor cuerpo donde dicho polinomio se vuelve reducible. La Proposición 1.3 nos dice que el cociente $L := \mathbb{Q}[t]/(t^2 - 2)$ es un cuerpo. Notar que el polinomio $x^2 - 2$ tiene una raíz en dicho cuerpo, dado que $t^2 = 2$, o sea

$$x^2 - 2 = (x - t)(x + t).$$

Además, $[L : \mathbb{Q}] = 2$, con lo cual cualquier cuerpo que tenga una raíz de $p(x)$ "contiene" a $L$ (en el sentido de que si $K$ es un cuerpo donde $p(x)$ posee una raíz, entonces existe un morfismo inyectivo de cuerpos de $L$ en $K$).

**Proposición 1.24.** *Si $p(x) \in K[x]$, existe un único cuerpo $L$ (salvo isomorfismos) que satisface las siguientes dos propiedades:*

- *el polinomio $p(x)$ tiene todas sus raíces en $L$,*
- *si $M$ es un cuerpo tal que $p(x)$ tiene todas sus raíces en $M$, entonces existe un morfismo de cuerpos de $L$ en $M$.*

*Al cuerpo $L$ se lo llama el* cuerpo de descomposición de $p(x)$.

*Demostración.* La demostración es constructiva, y se hace mediante un proceso inductivo agregando de a una raíz siguiendo el procedimiento del ejemplo anterior. Ver por ejemplo el Teorema 48 de [14]. □

**Definición 1.25.** Sea $L/K$ una extensión algebraica, y $K$ de característica 0 o finito. La extensión $L/K$ se dice *Galois* si todo polinomio irreducible en $K[x]$ que posee una raíz en $L$ tienen todas sus raíces en $L$.

Si $L/K$ es una extensión algebraica, definimos el conjunto de automorfismos de $L$ sobre $K$ como

$$\mathrm{Aut}_K(L) = \{\psi : L \to L \text{ morfismo de cuerpos tal que } \psi(k) = k \,\forall k \in K\}.$$

O sea son los morfismos del cuerpo $L$ en sí mismo que al restringirlos a $K$ dan la identidad.

Hay una caracterización alternativa de extensiones Galois en términos del grupo $\mathrm{Aut}_K(L)$: una extensión $L/K$ finita es Galois si y sólo si $[L : K] = \#\,\mathrm{Aut}_K(L)$. Con esta caracterización, no es difícil ver que el cuerpo de descomposición de un polinomio $p(x) \in K[x]$ siempre es una extensión Galois de $K$ (ver por ejemplo la demostración del Teorema 56 en [14]).

**Ejemplo 1.26.** Toda extensión $L/K$ cuadrática (o sea $[L : K] = 2$) es claramente Galois.

Existe otra construcción muy importante que es la de agregar no las raíces de un polinomio irreducible de $K$, sino agregar todas las raíces de todos los polinomios irreducibles de $K$.

**Definición 1.27.** Una clausura algebraica de $K$ es una extensión algebraica $L/K$ que satisface la propiedad de que todo polinomio $p(x) \in K[x]$ posee todas sus raíces en $L$, o sea $p(x)$ se factoriza como producto de polinomios de grado 1 en $L[x]$.

**Teorema 1.28.** *Todo cuerpo $K$ posee una clausura algebraica que denotamos por $\overline{K}$. Dicha clausura algebraica es única salvo isomorfismos.*

Así por ejemplo, la clausura algebraica de $\mathbb{C}$ es $\mathbb{C}$, mientras que la clausura algebraica de $\mathbb{R}$ es $\mathbb{C}$. La clausura algebraica de $\mathbb{Q}$ la podemos pensar como los números algebraicos de $\mathbb{C}$, o sea $\overline{\mathbb{Q}} = \{\alpha \in \mathbb{C} \ : \ \alpha \text{ es algebraico sobre } \mathbb{Q}\}$. Notar que $\overline{\mathbb{Q}}$ es mucho mas chico que $\mathbb{C}$. Es fácil ver que $\overline{\mathbb{Q}}$ es numerable, mientras que $\mathbb{C}$ claramente no lo es.

Si la extensión $L/K$ es Galois, denotamos por $\mathrm{Gal}(L/K)$ al grupo $\mathrm{Aut}_K(L)$. El principal resultado de la teoría de Galois es el siguiente.

**Teorema 1.29** (Galois)**.** *Si $L/K$ es una extensión Galois y finita, entonces:*

1. $\#\,\mathrm{Gal}(L/K) = [L : K]$.
2. *Existe una biyección entre los siguientes conjuntos:*
   - $\{$*Subextensiones $N$ con $K \subset N \subset L$*$\}$,
   - *El conjunto de subgrupos de* $\mathrm{Gal}(L/K)$*, o sea* $\{H : H < \mathrm{Gal}(L/K)\}$.

   *La biyección esta dada por:*

$$N \quad \to \quad \mathrm{Aut}_N(L) \subset \mathrm{Gal}(L/K) \tag{1.5}$$

$$\{\alpha \in L \ : \ \sigma(\alpha) = \alpha \, , \, \forall \sigma \in H\} \quad \leftarrow \quad H \tag{1.6}$$

   *Al conjunto $\{\alpha \in L \ : \ \sigma(\alpha) = \alpha \, , \, \forall \sigma \in H\}$ se lo denota $L^H$.*
3. *Si $K \subset N \subset L$, entonces la extensión $N/K$ es Galois si y sólo si el grupo* $\mathrm{Gal}(L/N)$ *es un subgrupo normal de* $\mathrm{Gal}(L/K)$*. En tal caso,* $\mathrm{Gal}(N/K) \simeq \mathrm{Gal}(L/K)/\mathrm{Gal}(L/N)$.

**Ejemplos 1.30.**

1. Consideremos el cuerpo $K := \mathbb{Q}(\sqrt{2})$, que corresponde al cuerpo de descomposición del polinomio irreducible $x^2 - 2$ en $\mathbb{Q}[x]$. La extensión $K/\mathbb{Q}$ es Galois (por ser de grado 2, o por ser $K$ un cuerpo de descomposición), y el grupo $\mathrm{Gal}(\mathbb{Q}(\sqrt{2})/\mathbb{Q})$ tiene orden 2. El elemento no trivial corresponde a la involución $\sigma(a + b\sqrt{2}) = a - b\sqrt{2}$, donde $a, b \in \mathbb{Q}$.
2. Consideremos $K$ el cuerpo de descomposición del polinomio $x^3 - 2$. En términos de radicales, $K = \mathbb{Q}(\sqrt[3]{2}, \sqrt{-3})$ (recordar que las raíces cúbicas de la unidad son $\xi_3 = \frac{-1+\sqrt{-3}}{2}$ y su conjugado). En particular $[K : \mathbb{Q}] = 6$ (¿por qué?). El grupo $\mathrm{Gal}(K/\mathbb{Q})$ tiene orden 6, y está generado por los morfismos:
   - $\tau(\sqrt[3]{2}) = \sqrt[3]{2}$, $\tau(\sqrt{-3}) = -\sqrt{-3}$,
   - $\sigma(\sqrt[3]{2}) = \xi_3\sqrt[3]{2}$, $\sigma(\sqrt{-3}) = \sqrt{-3}$.

   El morfismo $\tau$ tiene orden 2 mientras que $\sigma$ tiene orden 3. Es fácil verificar que $\sigma\tau \neq \tau\sigma$, con lo cual $\mathrm{Gal}(K/\mathbb{Q}) \simeq S_3$.

Por último, queremos mencionar la estructura que tienen los grupos de Galois de cuerpos finitos. Si $K$ es un cuerpo de característica $p$, entonces $\mathbb{F}_p \subset K$. En particular, existe un morfismo llamado de *Frobenius*, y denotado $\sigma_p$ dado por

$$\sigma_p(x) = x^p.$$

*Ejercicio* 1.31. Probar que si $K$ tiene característica $p$ (en particular cualquier múltiplo de $p$ es cero en $K$), entonces el morfismo de Frobenius es un morfismo de cuerpos, o sea $(a + b)^p = a^p + b^p$ y $(ab)^p = a^p b^p$.

**Teorema 1.32.** *Si $K$ es un cuerpo finito de $p^d$ elementos, entonces* $\mathrm{Gal}(K/\mathbb{F}_p)$ *es cíclico de orden $d$. Más aún,* $\mathrm{Gal}(K/\mathbb{F}_p) = \langle \mathrm{Frob}_p \rangle$.

## 2. Representaciones de Artin

Sea $K/\mathbb{Q}$ una extensión finita de Galois. Vamos a estudiar las representaciones

$$\rho : \mathrm{Gal}(K/\mathbb{Q}) \to \mathrm{GL}(V)$$

donde $V$ es un espacio vectorial de dimensión finita sobre $\mathbb{C}$. Veremos cómo Artin le asocia a $\rho$ una serie de Dirichlet $L(\rho, s)$ que extiende simultáneamente la definición de la función zeta de Riemann $\zeta(s)$ y de las series de Dirichlet $L(\chi, s)$ donde $\chi$ es un carácter de Dirichlet. Estas *L-series de Artin* $L(\rho, s)$ codifican mucha de la aritmética del cuerpo de números $K$.

**2.1. Representaciones lineales de grupos finitos.** Empezamos considerando sólo un grupo finito $G$ (que luego será $\mathrm{Gal}(K/\mathbb{Q})$) dejando de lado la extensión $K/\mathbb{Q}$. Una *representación* de $G$ es un homomorfismo

$$\rho : G \to \mathrm{GL}(V)$$

donde $V$ es un espacio vectorial de dimensión finita sobre $\mathbb{C}$. En otras palabras, $G$ actúa en $V$ por medio de transformación lineales. También hablaremos de la representación $V$, usando la notación $g \cdot v$ (o simplemente $gv$) para $g \in G$ y $v \in V$ en vez de $\rho(g)(v)$, cuando la representación $\rho$ en cuestión es clara del contexto.

Sea $n := \dim V$. Si elegimos una base de $V$ tenemos que $\mathrm{GL}(V)$ es isomorfo al grupo lineal $\mathrm{GL}_n(\mathbb{C})$ de matrices $n \times n$ complejas inversibles. Nos interesan las representaciones sólo módulo isomorfismo. Por definición, la clase de isomorfía de $\rho$ es la clase módulo conjugación del homomorfismo resultante

$$\rho : G \to \mathrm{GL}_n(\mathbb{C}),$$

que abusando la notación seguimos llamando $\rho$ si no lleva a confusión, independientemente de la base elegida. Pasaremos de una versión de $\rho$ a otra según convenga.

En lo que sigue damos un resumen breve de las propiedades principales de las representaciones de grupos finitos que necesitamos. Las demostraciones se pueden encontrar en cualquier libro introductorio (como [4]).

**Definición 2.1.** i) Una subrepresentación de $V$ es un subespacio $U \subseteq V$ que es estable por multiplicación por $G$ vía $\rho$. Es decir, para todo $u \in U$ y para todo $g \in G$ vale que

$$\rho(g)u \in U.$$

ii) La representación $\rho$ es irreducible si y sólo si sus únicas subrepresentaciones son $\{0\}$ y $V$.

**Teorema 2.2.** *Toda representación de $G$ es suma directa de representaciones irreducibles.*

*Demostración.* La idea de la demostración es probar que si $U \subseteq V$ es una subrepresentación entonces existe otra subrepresentación $U'$ tal que

$$V = U \oplus U'.$$

Sabemos que existe siempre un tal subespacio vectorial $U'$ (un complemento de $U$). Lo que necesitamos probar es que se puede elegir $U'$ tal que también sea estable por multiplicación por $G$ vía $\rho$.

Una forma útil de mostrar esto es probando que existe un producto Hermitiano $(\cdot, \cdot)$ no degenerado en $V$ para el cual $\rho(g)$ es una isometría. En efecto, dado

este producto basta tomar $U' = U^{\perp}$. Para conseguir un tal producto Hermitiano empezamos con un producto Hermitiano positivo cualquiera $(\cdot,\cdot)_0$ y definimos

$$(u,v) := \sum_{g\in G} (gu, gv)_0$$

como el promedio de $(\cdot,\cdot)_0$ sobre $G$. No es difícil verificar que el nuevo producto Hermitiano es no degenerado. □

Vale la pena notar que este resultado fundamental es falso en situaciones mas generales. Por ejemplo,

1) si el grupo $G$ es infinito:

$$\begin{array}{rccc} \rho & :\mathbb{Z} & \longrightarrow & \mathrm{GL}_2(\mathbb{C}) \\ & 1 & \longmapsto & \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \end{array}$$

tiene como única subrepresentación de dimensión 1 el subespacio generado por el vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$;

2) si la característica del cuerpo divide al orden del grupo:

$$\begin{array}{rccc} \rho & :\mathbb{Z}/p\mathbb{Z} & \longrightarrow & \mathrm{GL}_2(\mathbb{F}_p) \\ & 1 & \mapsto & \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \end{array}$$

también tiene como única subrepresentación de dimensión 1 el subespacio generado por el vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Una observación importante sobre nuestras representaciones $\rho$ es la siguiente.

**Proposición 2.3.** *Sea $g \in G$ y $A := \rho(g) \in \mathrm{GL}_n(\mathbb{C})$. Entonces*

*1) $A$ es diagonalizable;*

*2) los autovalores de $A$ son raíces de la unidad de orden dividiendo $|G|$.*

**Definición 2.4.** Dada una representación $\rho$ definimos su *carácter* $\chi : G \to \mathbb{C}$ como la función

$$\chi(g) := \mathrm{Tr}(\rho(g)).$$

Es claro que $\chi$ sólo depende de la clase de isomorfía de $\rho$. De hecho, tenemos el siguiente resultado fundamental.

**Teorema 2.5.** *El carácter $\chi$ determina unívocamente la clase de isomorfía de $\rho$; mas precisamente, dos representaciones $\rho, \rho'$ son isomorfas si y sólo si sus respectivos caracteres $\chi, \chi'$ son iguales.*

Veamos algunas de las propiedades básicas del carácter de una representación.

**Proposición 2.6.** *Sea*

$$\rho : G \to \mathrm{GL}(V)$$

*una representación, $\chi$ su carácter y $n := \dim V$ su dimensión. Para todo $g \in G$ tenemos*

1. *$\chi$ es constante en clases de conjugación de $G$;*
2. *$\chi(g)$ es un entero algebraico (esto es $\chi(g)$ es raíz de un polinomio mónico con coeficientes enteros);*
3. *$|\chi(g)| \leq n$;*

4. *la igualdad vale en 3. si y sólo si* $\rho(g)$ *es la identidad;*
5. $\chi(1) = n$.

**2.2. Representaciones de permutación de grupos finitos.** Una forma natural en que la que ocurre un grupo es por medio de permutaciones de un conjunto finito; por ejemplo, caso central a estas notas, el grupo de Galois de un polinomio actúa como permutación de sus raíces.

Sea $G$ un grupo finito y $X$ un conjunto finito donde $G$ actúa por medio de permutaciones; i.e., tenemos $G \hookrightarrow S(X)$. Llamamos esto una *representación de permutación* de $G$ y le asociamos una representación lineal $\rho$ como sigue. Sea $V = \{\varphi : X \to \mathbb{C}\}$ el espacio vectorial de funciones en $X$ a valores complejos. Definimos para $x \in X$

$$\rho(g)\varphi(x) := \varphi(g^{-1}x).$$

Es fácil verificar que obtenemos efectivamente una representación de $G$

$$\rho : G \to \mathrm{GL}(V)$$

de dimensión $\dim(V) = \#X$. (Aquí es crucial que la definición usa $g^{-1}$ en el argumento de $\varphi$ para que se satisfaga que $\rho(g_1g_2) = \rho(g_1)\rho(g_2)$ y no $\rho(g_1g_2) = \rho(g_2)\rho(g_1)$.) Resulta también sencillo verificar que el carácter $\chi$ asociado a $\rho$ está dado por el número de puntos fijos de $g$:

$$\chi(g) = \#\{x \in X \mid gx = x\}.$$

En particular, el carácter toma valores enteros no negativos.

Notemos que el subespacio de $V$ de las funciones constantes es una subrepresentación de $\rho$ de dimensión uno. Por lo tanto, una representación lineal que proviene de una de permutación no es nunca irreducible si su dimensión es mayor que uno.

**Ejemplo 2.7.** Tomemos como ejemplo $G = S_n$ actuando en $\{1, \ldots, n\}$ en forma natural. Obtenemos una representación lineal de dimensión $n$. Concretamente, $\sigma \in S_n$ actúa en $\mathbb{C}^n$ por medio de la correspondiente matriz de permutación $A = (a_{i,j})$ donde

$$a_{i,j} = \begin{cases} 1, & \sigma^{-1}(i) = j \\ 0, & \sigma^{-1}(i) \neq j \end{cases},$$

ya que en la base canónica $e_1, \ldots, e_n$ de $V_0 := \mathbb{C}^n$ (pensando el vector $e_i$ como la función definida por $e_i(j) = \delta_{i,j}$ para $j = 1, \ldots, n$) tenemos

$$\sigma e_i = e_{\sigma(i)}.$$

Las matrices de permutación son aquellas que tienen un solo elemento no nulo igual a 1 en cada fila y columna.

Explícitamente, tomemos por ejemplo $n = 5$ y $\sigma = (235) \in S_5$. Entonces si $v = (a, b, c, d, e)$ tenemos

$$\sigma v = ae_1 + be_3 + ce_5 + de_4 + ee_2 = (a, e, b, d, c)$$

y la matriz de permutación correspondiente es

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

Para $n > 1$ la representación se descompone en la suma directa no trivial de representaciones

$$V_0 = U \oplus V$$

donde $U := \langle (1, 1, \ldots, 1) \rangle$ y

$$V := \{(x_1, \ldots, x_n) \,|\, x_1 + \cdots + x_n = 0\}.$$

Llamamos a $V$ la *representación estándar* de $S_n$.

**Teorema 2.8.** *La representación estándar de $S_n$ es irreducible.*

**Ejemplo 2.9.** Si $H < G$ es un subgrupo, $G$ actúa en las coclases $G/H$ a izquierda por multiplicación: $g \cdot xH := (gx)H$. Obtenemos entonces una representación de permutación y por lo tanto una representación lineal correspondiente de dimensión $[G : H]$. Si tomamos $H = \{1\}$ el subgrupo trivial la representación correspondiente se conoce como la *representación regular* $\rho_{\text{reg}}$ de $G$. Su dimensión es $|G|$ y su carácter asociado es la función

$$\chi_{\text{reg}}(g) = \begin{cases} |G|, & g = 1 \\ 0, & g \neq 1 \end{cases}.$$

En la representación de permutación asociada a un subgrupo $H \leqslant G$, el grupo $G$ actúa transitivamente (podemos pasar de una coclase $xH$ a la coclase $H$ multiplicando a izquierda por $x^{-1}$). Recíprocamente, toda acción transitiva de $G$ en un conjunto finito $X$ da lugar a una representación de permutación isomorfa a $G/H$ donde $H$ es el estabilizador de un punto cualquiera de $X$.

Por otro lado, toda acción de $G$ en un conjunto finito $X$ determina una partición de $X$ en órbitas. En cada órbita $G$ actúa transitivamente. Concluimos que las representaciones lineales de $G$ que provienen de una representación de permutación son suma directa de aquellas asociadas a subgrupos.

Digamos que dos representaciones de permutación de un grupo $G$ son *equivalentes* si existe una biyección equivariante entre los respectivos conjuntos donde $G$ actúa. En ese caso las representaciones lineales correspondientes son isomorfas. Veremos más adelante §4.3 que la recíproca no es cierta. Existen representaciones de permutación no equivalentes que dan lugar a representaciones lineales isomorfas. Este fenómeno esta relacionado con el problema de *can you hear the shape of a drum?* de variedades Riemannianas. En nuestro contexto da lugar a cuerpos de números no isomorfos con la misma función zeta.

En general clasificar representaciones de permutación módulo equivalencia es bien difícil mientras que clasificar representaciones lineales módulo isomorfismo es mucho más accesible, como veremos en la sección siguiente.

**2.3. Tabla de caracteres.** En el espacio vectorial de funciones $\phi : G \longrightarrow \mathbb{C}$ definimos el producto Hermitiano

$$(\phi, \psi) := \frac{1}{|G|} \sum_{g \in G} \phi(g)\overline{\psi(g)}.$$

Sea $\mathcal{C}$ el subespacio vectorial de todas las funciones $G \longrightarrow \mathbb{C}$ que son constantes en clases de conjugación, con el mismo producto Hermitiano. Una tal función se llama una *función de clase*. Como vimos (Proposición(2.6) i) el carácter $\chi$ de una representación de $G$ es un elemento de $\mathcal{C}$.

**Teorema 2.10.** *El conjunto de caracteres de representaciones irreducibles de $G$ es una base ortonormal de $\mathcal{C}$.*

**Corolario 2.11.**
1. *El número de representaciones irreducibles de $G$ no isomorfas entre sí es igual al número de clases de conjugación de $G$.*
2. *Sea $U_1, \dots, U_r$ una lista de todas las representaciones irreducibles de $G$ no isomorfas entre sí. Denotemos por $\chi_i$ el carácter de $U_i$. Sea $V$ una representación arbitraria de $G$ de carácter $\chi$. Entonces, $V$ admite una descomposición en suma directa*
$$V \simeq U_1^{m_1} \oplus \cdots \oplus U_r^{m_r},$$
*con enteros no negativos $m_1, \dots, m_r$ si y sólo si*
$$\chi = m_1\chi_1 + \cdots m_r\chi_r.$$

El entero $m_i$ se llama la *multiplicidad* de $U_i$ en $V$; notemos que
$$m_i = (\chi, \chi_i).$$

Deducimos que toda la información necesaria para descomponer representaciones de un grupo finito $G$ en suma de irreducibles está contenida en la *tabla de caracteres* de $G$. Ésta consiste en una matriz $r \times r$ con los valores $\chi_i(c_j)$, donde $c_1, \dots, c_r$ son representantes de las clases de conjugación de $G$. Conviene también listar el número de elementos de cada clase de conjugación para facilitar el cálculo del producto interno.

Veamos un ejemplo. Tomemos $G = S_4$. Su tabla de caracteres es la siguiente.

| | 1 | (1 2) | (1 2 3) | (1 2 3 4) | (1 2)(3 4) |
|---|---|---|---|---|---|
| | 1 | 6 | 8 | 6 | 3 |
| $U$ | 1 | 1 | 1 | 1 | 1 |
| $U'$ | 1 | −1 | 1 | −1 | 1 |
| $V$ | 3 | 1 | 0 | −1 | −1 |
| $V'$ | 3 | −1 | 0 | 1 | −1 |
| $W$ | 2 | 0 | −1 | 0 | 2 |

La tabla se puede calcular usando alguno de los varios programas existentes (por ejemplo MAGMA o GAP). También se puede calcular a mano sin gran dificultad. Aquí elegimos tomar la tabla como dada y entender las representaciones irreducibles de $S_4$ a partir de ella.

La primer fila da los representantes $c_1, \dots, c_5$ de las clases de conjugación. En general para $S_n$ las clases de conjugación están naturalmente indexadas por la descomposición en ciclos. La segunda fila tiene el número de elementos de la clase correspondiente (1 elemento trivial, 6 transposiciones, 8 triciclos, etc.).

Cada fila de la tabla corresponde al carácter de una representación irreducible de $S_4$. En general es bastante mas fácil dar el carácter de una representación irreducible que describir explícitamente la representación misma. Quizás parezca paradójico pero no es inusual obtener el carácter de una representación primero y solo luego construir trabajosamente la representación.

Las dos primeras representaciones $U, U'$ son la representación trivial (donde $\rho(g) = 1$ para todo $g \in G$) y la representación signo respectivamente (donde $\rho(\sigma) = \epsilon(\sigma)$ es el signo de la permutación $\sigma$). La representación $V$ es la representación estándar de $S_4$ ya mencionada y $V' = V \otimes U'$ es su producto tensorial con la representación signo (esto es $\rho_{V'}(\sigma) = \epsilon(\sigma)\rho_V(\sigma)$).

Nos queda entender la última representación $W$. Notemos que su carácter $\chi_W$ esta completamente determinado por el resto de la tabla. En efecto, $\chi_W$ es ortogonal a todas las otras filas (una condición que determina un espacio vectorial de dimensión 1 en $\mathcal{C}$), tiene norma $(\chi_W, \chi_W) = 1$ y $\chi_W(1) = \dim W$.

El grupo $S_4$ tiene un subgrupo normal $A \trianglelefteq S_4$ de orden 4, el famoso *Vierergruppe* de Klein. Consiste de la identidad y todos los productos de dos transposiciones disjuntas

$$A := \{1, (12)(34), (13)(24), (14)(23)\}.$$

El cociente $S_4/A$ es isomorfo a $S_3$. Componiendo el homomorfismo correspondiente $S_4 \longrightarrow S_3$ con la representación estándar de $S_3$ obtenemos una representación $\rho$ de dimensión 2 de $S_4$. Como la representación estándar de $S_3$ es irreducible (Teorema 2.8) $\rho$ también lo es.

Más directamente, hay una acción natural dada por conjugación de $S_4$ en el conjunto $\{(12)(34), (13)(24), (14)(23)\}$ y un cálculo simple muestra que la correspondiente representación es isomorfa a $U \oplus W$.

## 3. Aritmética en extensiones

**3.1. Factorización en primos.** Sea $K/\mathbb{Q}$ una extensión finita de grado $n$ y $\mathcal{O}_K$ su anillo de enteros, esto es el conjunto formado por todos los enteros algebraicos en $K$ (que se puede ver es un anillo). Un número primo $p$ genera un ideal primo de $\mathbb{Z}$ que se factoriza en $\mathcal{O}_k$ de la siguiente forma,

$$p\mathcal{O}_K = \mathcal{P}_1^{e_1} \cdots \mathcal{P}_r^{e_r}, \tag{3.1}$$

donde $\mathcal{P}_i$ son primos distintos de $\mathcal{O}_K$ y $e_i$ son enteros positivos. La factorización es única salvo reordenamiento. Definimos los enteros positivos $f_1, \ldots, f_r$ cómo

$$|\mathcal{O}_K/\mathcal{P}_i| = p^{f_i}, \qquad i = 1, \ldots, r.$$

Se tiene entonces que

$$e_1 f_1 + \cdots e_r f_r = n.$$

De hecho, salvo para un número finito de excepciones, los exponentes $e_i$ son siempre igual a uno. Tales primos $p$ para los que algún $e_i > 1$ se llaman *ramificados* en $K/\mathbb{Q}$.

**Teorema 3.1.** *(Dedekind) Sea $h \in \mathbb{Z}[x]$ un polinomio mónico irreducible tal que $K \simeq \mathbb{Q}[x]/(h)$. Sea $\Delta$ el discriminante de $h$ y $p \nmid \Delta$ un primo. Entonces $p$ es no ramificado en $K/\mathbb{Q}$ y, si*

$$h \equiv h_1 \cdots h_r \bmod p,$$

*es la factorización de $h$ en $\mathbb{Z}/p\mathbb{Z}[x]$ en producto de irreducibles distintos, se tiene que ordenando los factores apropiadamente el grado de $h_i$ es $f_i$.*

En general, $\Delta$ es divisible por más primos que los ramificados en $K$ y para estos primos el teorema deja sin poder calcular los $f_i$. Para buena parte de nuestra discusión esto no será un gran inconveniente.

**3.2. Automorfismo de Frobenius.** Sea $L/\mathbb{Q}$ una extensión finita de Galois de grado $n$, $\mathcal{O}_L$ su anillo de enteros y $G := \mathrm{Gal}(L/\mathbb{Q})$. Tomemos un primo $p$ y $\mathcal{P}$ un primo de $L$ en la factorización de $p$ en $\mathcal{O}_L$. El *grupo de descomposición* $D_{\mathcal{P}}$ de $\mathcal{P}$ es el subgrupo de $G$ de elementos que fijan $\mathcal{P}$. Tenemos un morfismo de reducción natural

$$\phi : D_{\mathcal{P}} \to \mathrm{Gal}(k_{\mathcal{P}}/k), \tag{3.2}$$

donde $k_{\mathcal{P}} := \mathcal{O}_L/\mathcal{P}$ y $k := \mathbb{Z}/p\mathbb{Z}$. Este morfismo $\phi$ es siempre sobreyectivo, su núcleo $I_{\mathcal{P}}$ se conoce como el *grupo de inercia* de $\mathcal{P}$. Si el primo $p$ es no ramificado en $L/\mathbb{Q}$ el grupo de inercia es trivial y $\phi$ es un isomorfismo.

El grupo de Galois $\mathrm{Gal}(k_{\mathcal{P}}/k)$ es cíclico generado por el automorfismo $\sigma_p : x \mapsto x^p$ (ver Teorema 1.32). En el caso de que $p$ sea no ramificado existe entonces un único automorfismo $\mathrm{Frob}_{\mathcal{P}} \in D_{\mathcal{P}}$, el *automorfismo de Frobenius* asociado a $\mathcal{P}$, tal que $\phi(\mathrm{Frob}_{\mathcal{P}}) = \sigma_p$.

Si $p$ es no ramificado y tomamos otro primo $\mathcal{P}'$ en la factorización de $p$ obtenemos un automorfismo $\mathrm{Frob}_{\mathcal{P}'}$ conjugado a $\mathrm{Frob}_{\mathcal{P}}$. Esto resulta del hecho que el grupo de Galois $G$ actúa transitivamente en los primos de $L$ que dividen a $p$. En definitiva tenemos que dado un primo $p$ no ramificado en $L/\mathbb{Q}$ existe una única clase de conjugación de $G$ asociada a $\mathrm{Frob}_p$: la clase de conjugación de $\mathrm{Frob}_{\mathcal{P}}$ para cualquier primo $\mathcal{P}$ de $L$ que divide a $p$.

Tomemos ahora una extensión finita arbitraria $K/\mathbb{Q}$ y sea $L/\mathbb{Q}$ su clausura Galoisiana (esto es la menor extensión de Galois que contiene a $K$). Un primo $p$ no ramificado en $K$ sera también no ramificado en $L$. La factorización de $p$ en $\mathcal{O}_K$ determina una *partición* de $n := [K : \mathbb{Q}]$: $\tau_p := [f_1, f_2, \ldots, f_r]$ donde $f_1 \geq f_2 \geq \cdots f_r$. Basta reordenar los números $f_i$ de mayor a menor. Llamaremos a $\tau_p$ el *tipo* de factorización de $p$ en $K$.

Sea $h \in \mathbb{Q}[x]$ un polinomio irreducible tal que $K = \mathbb{Q}[x]/(h(x))$ y $L$ su clausura de Galois. El grupo de Galois $G = \mathrm{Gal}(L/\mathbb{Q})$ actúa en las raíces de $h$ lo que nos da un homomorfismo $\iota : G \to S_n$.

**Teorema 3.2.** *Con la notación anterior, tenemos que la partición de $n$ dada por la descomposición en ciclos de $\iota(\mathrm{Frob}_p)$ es $\tau_p$.*

Combinando este resultado con el Teorema 3.1 obtenemos una forma práctica de conocer la descomposición en ciclos de los automorfismos de Frobenius como permutación de las raíces de $h$. Observemos que esto *no* es lo mismo que la clase de conjugación en $G$. En efecto al pasar de $G$ a $S_n$ vía $\iota$ dos elementos de $G$ pueden ser conjugados en $S_n$ sin ser conjugados en $G$. De todas formas este resultado es sumamente útil.

**3.3. Un ejemplo: el Teorema de Chebotarev.** Tomemos $h = x^4 + x + 1$ y $K := \mathbb{Q}(\theta)$ con $\theta \in \mathbb{C}$ una raíz cualquiera de $h$. El discriminante de $h$ es el primo 229, con lo cual $\mathrm{disc}(K) = 229$. La lista de $\tau_p$ para los primos $p \leq 50$ está dada en 1.

Vemos que aparecieron todos las posibles particiones de 4 excepto por $[1, 1, 1, 1]$. De hecho tenemos el siguiente resultado de Chebotarev. (En lo que sigue ignoramos tácitamente los primos ramificados o en general que dividan al discriminante del polinomio en cuestión).

**Teorema 3.3.** *Sea $C$ una clase de conjugación de $G$. Entonces*

$$\lim_{x\to\infty} \frac{\#\{p \leq x \mid \mathrm{Frob}_p = C\}}{\#\{p \leq x\}} = \frac{\#C}{|G|}. \tag{3.3}$$

**Corolario 3.4.** *Sea $\tau$ una partición de $n$ correspondiente a la descomposición de ciclos de $\iota(\sigma)$ para un $\sigma \in G$. Entonces existen infinitos primos $p$ tales que $\tau = \tau_p$.*

En nuestro ejemplo se puede verificar que $G = S_4$ (de hecho, dada la lista de $\tau_p$ que obtuvimos necesariamente $G = S_4$). Entonces de acuerdo al corolario toda partición de 4 es de la forma $\tau_p$ para infinitos primos $p$. Es típico que la clase

| $p$ | $\tau_p$ |
|---|---|
| 2 | [4] |
| 3 | [3, 1] |
| 5 | [3, 1] |
| 7 | [4] |
| 11 | [3, 1] |
| 13 | [4] |
| 17 | [3, 1] |
| 19 | [3, 1] |
| 23 | [2, 1, 1] |
| 29 | [2, 1, 1] |
| 31 | [4] |
| 37 | [2, 2] |
| 41 | [4] |
| 43 | [3, 1] |
| 47 | [2, 1, 1] |

FIGURA 1. Partición del polinomio $x^4 + x + 1$ para los primeros 50 primos

$[1, 1, \ldots, 1]$ que corresponde a $\sigma = 1$ (o alternativamente a primos $p$ donde $h$ se factoriza en $n$ factores lineales módulo $p$) requiera primos $p$ más bien grandes (ver [15]). En nuestro caso el primer primo tal que $\tau_p = [1, 1, 1, 1]$ es $p = 193$ aunque la proporción de tales primos con $p \leq x$ para $x$ muy grande es aproximadamente $1/24$.

Para una partición $\tau$ de $n$ sea

$$\delta(\tau, x) := \frac{\#\{p \leq x \mid \tau_p = \tau\}}{\#\{p \leq x\}}.$$

En la Figura 2 damos una lista de $d(x) := 24\delta([1, 1, 1, 1], x)$ para varios valores de $x$. El Teorema 3.3 de Chebotarev garantiza que $\lim_{x\to\infty} d(x) = 1$.

| $k$ | $d(500k)$ |
|---|---|
| 1 | 0.252631579 |
| 2 | 0.428571429 |
| 3 | 0.702928870 |
| 4 | 0.633663366 |
| 5 | 0.719346049 |
| 6 | 0.837209302 |
| 7 | 0.785276074 |
| 8 | 0.872727273 |
| 9 | 0.944262295 |
| 10 | 0.932735426 |

| $k$ | $d(10^4k)$ |
|---|---|
| 1 | 0.937347437 |
| 2 | 0.933687003 |
| 3 | 0.902311248 |
| 4 | 0.896502498 |
| 5 | 0.921098773 |
| 6 | 0.911342249 |
| 7 | 0.934390771 |
| 8 | 0.958530050 |
| 9 | 0.980603696 |
| 10 | 0.988323603 |

| $k$ | $d(10^6k)$ |
|---|---|
| 1 | 0.989069785 |
| 2 | 0.994594885 |
| 3 | 0.995682975 |
| 4 | 0.995867856 |
| 5 | 0.997770528 |
| 6 | 0.994533110 |
| 7 | 0.998170558 |
| 8 | 0.996455944 |
| 9 | 0.996705334 |
| 10 | 0.996071197 |

FIGURA 2. Valores de $24\delta([1, 1, 1, 1], x)$

Como vemos la convergencia en (3.3) no es particularmente rápida; es importante tener buenas cotas para el error

$$\frac{\#\{p \leq x \mid \mathrm{Frob}_p = C\}}{\#\{p \leq x\}} - \frac{\#C}{|G|}$$

y esto está ligado a la hipótesis de Riemann generalizada. Ver [22] para una historia del Teorema de Chebotarev y [7] para una demostración con una cota para el error.

Desde otro punto de vista, el tipo de factorización $\tau_p$ codifica el número de raíces que tiene el polinomio en los cuerpos $\mathbb{F}_q$ con $q$ una potencia de $p$ (éste es el punto de vista adoptado en [13]). Más precisamente, sea

$$N(q) := \#\{\theta \in \mathbb{F}_q \mid f(\theta) = 0\}.$$

Entonces, es claro que

$$N(p^r) = \sum_{d|r} m_d$$

donde $m_d$ es el número de partes de $\tau_p$ iguales a $d$. Ver [20] para una discusión general sobre cómo varía el número de puntos de una variedad algebraica sobre un cuerpo finito.

## 4. Series $L$ de Artin

Ahora combinamos las representaciones de grupos finitos de la sección 2.1 con la aritmética de la sección 3 para definir las series $L$ de Artin. Estas series generalizan simultáneamente la función $\zeta$ de un cuerpo de números y las series $L(\chi, s)$ asociadas a un carácter de Dirichlet. Supongamos que $L/\mathbb{Q}$ es una extensión finita de Galois con $G := \mathrm{Gal}(L/\mathbb{Q})$ y que

$$\rho : \quad G \to GL(V)$$

es una representación compleja de $G$. Siguiendo a Artin vamos a definir una serie de Dirichlet, es decir una serie de la forma

$$L(\rho, s) = \sum_{n \geq 1} a_n n^{-s}, \qquad \Re(s) \gg 0, \tag{4.1}$$

asociada a la representación $\rho$.

Esta serie se define a partir de un producto de Euler de la forma de

$$L(\rho, s) = \prod_p L_p(\rho, p^{-s})^{-1}, \tag{4.2}$$

donde $p$ recorre todos los números primos y $L_p(\rho, T)$ son polinomios de grado acotado. Llamamos a $L_p(\rho, T)$ el *factor de Euler* en $p$.

**4.1. Factores de Euler.** Si $p$ es un primo no ramificado de $L/\mathbb{Q}$ definimos el factor de Euler asociado de la siguiente manera

$$L_p(\rho, T) := \det\left(1 - \rho(\mathrm{Frob}_p)T\right).$$

Hay un poco de abuso de notación aquí ya que $\mathrm{Frob}_p$ no es un elemento de $G$ sino una clase de conjugación. Pero basta tomar cualquier representante de esta clase para calcular el determinante. El resultado no dependerá de esta elección. Lo mismo sucede en lo que sigue al tomar la traza $\mathrm{Tr}(\rho(\mathrm{Frob}_p))$, por ejemplo. No insistiremos con el tema.

Tenemos que

$$L_p(\rho, T) = 1 - a_p\, T + \cdots$$

y por lo tanto

$$L_p(\rho, T)^{-1} = 1 + a_p\, T + \cdots$$

Vemos entonces que la traza de $\rho(\mathrm{Frob}_p)$ es el coeficiente de $p^{-s}$ en (4.1), es decir

$$a_p = \mathrm{Tr}(\rho(\mathrm{Frob}_p)).$$

Si el primo $p$ es ramificado y $\mathcal{P}$ es un divisor primo de $p$ en $L$, la preimagen de $x \mapsto x^p$ vía $\phi$ en (3.2) es único sólo módulo el grupo de inercia $I_{\mathcal{P}}$. Es decir, tenemos un elemento $\mathrm{Frob}_{\mathcal{P}} \in D_{\mathcal{P}}/I_{\mathcal{P}}$ bien definido. Este elemento da lugar a su vez a un elemento $\rho(\mathrm{Frob}_{\mathcal{P}}) \in \mathrm{GL}(V^{I_{\mathcal{P}}})$ bien definido. Aquí $V^{I_{\mathcal{P}}} \subseteq V$ es el subespacio de los elementos de $V$ que quedan fijos por los elementos de $I_{\mathcal{P}}$. Podemos definir el factor de Euler asociado a $p$ entonces como $L_p(\rho, T) := \det(1 - \rho(\mathrm{Frob}_{\mathcal{P}})T|_{V^{I_{\mathcal{P}}}})$. El resultado no depende de la elección del divisor primo $\mathcal{P}$ ya que eligiendo otro divisor da lugar a un elemento conjugado.

La serie de Artin (4.2) formada con estos factores de Euler converge absolutamente para $\Re(s) > 1$ (usando la Proposición 2.3 (3)). Notemos que el grado de los factores de Euler $L_p(\rho, T)$ es a lo sumo $\dim V$ y es igual a $\dim V$ salvo un número finito de casos (que serán algunos de los primos ramificados). Es una conjetura de Artin que para $\rho$ no-trivial la serie $L(\rho, s)$ se extiende a una función entera de la variable $s$ y satisface una ecuación funcional relacionando $L(\rho, s)$ con $L(\rho, 1 - s)$.

**4.2. Ejemplo I: polinomios de grado cuatro.** Sea $h \in \mathbb{Z}[x]$ un polinomio irreducible de grado $n = 4$ y $\theta \in \mathbb{C}$ una raíz de $h$. Definimos $K := \mathbb{Q}(\theta)$ y $L \subset \mathbb{C}$ su clausura Galois. Supongamos que $G := \mathrm{Gal}(L/\mathbb{Q})$ es isomorfo al grupo $S_4$ de permutaciones de cuatro elementos, digamos $X := \{1, 2, 3, 4\}$. Vamos a describir las series $L$ de Artin de $L/\mathbb{Q}$ asociadas a las representaciones irreducibles de $S_4$ (que ya describimos en §2.3).

Como vimos en §2.2 la acción natural de $S_4$ en $X$ da lugar a una representación lineal $V_0$ isomorfa a $U \oplus V$, donde $U$ es la representación trivial y $V$ la representación estándar. Numerando las raíces de $h$ en $\mathbb{C}$ apropiadamente la acción de $G$ es justamente esta acción natural de $S_4$ en $X$.

Una forma más conceptual y completa del Teorema 3.2 es la siguiente

$$L(V_0, s) = \zeta_K(s) := \prod_{\mathcal{P}} (1 - \mathbb{N}\mathcal{P}^{-s})^{-1}, \tag{4.3}$$

donde el producto recorre todos los primos $\mathcal{P}$ del anillo de enteros $\mathcal{O}_K$ de $K$. Notar que $\mathbb{N}\mathcal{P} = |\mathcal{O}_L/\mathcal{P}| = p^f$. La función $\zeta_K(s)$ es la función zeta del cuerpo $K$; coincide con la función zeta de Riemann $\zeta(s)$ cuando $K = \mathbb{Q}$.

En efecto, para un primo $p$ que no divide al discriminante de $h$, la factorización de $h$ módulo $p$ determina la factorización de $p$ en $K$. Esta factorización a su vez nos da el factor de Euler en $\zeta_K(s)$ donde $\mathbb{N}\mathcal{P}$ que no es otro que $L_p(V_0, T)$. Vemos que (4.3) extiende apropiadamente el Teorema 3.2 a *todos* los primos.

La descomposición de $\rho$ en representaciones irreducibles corresponde a la factorización

$$\zeta_K(s) = L(U, s)L(V, s),$$

de las respectivas funciones $L$ de Artin. Como $L(U, s) = \zeta_{\mathbb{Q}}(s) = \zeta(s)$ deducimos que

$$L(V, s) = \zeta_K(s)/\zeta(s).$$

Resumimos el argumento que usamos para el caso general

**Proposición 4.1.** *Sea $h \in \mathbb{Q}[x]$ un polinomio irreducible de grado $n$ y $\theta \in \mathbb{C}$ una raíz de $h$. Sea $L/\mathbb{Q}$ la clausura de Galois de $\mathbb{Q}(\theta)/\mathbb{Q}$ con grupo de Galois $G$. Sea $\rho$ la representación de $G$ proveniente de la acción en las raíces de $h$. Entonces*

$$L(\rho, s) = \zeta_K(s). \tag{4.4}$$

Para simplificar la discusión, fijemos ahora $h := x^4 + x + 1$ de discriminante 229 como en §3.3. ¿Qué podemos decir de la serie $L$ de Artin asociada a la representación de signo $U'$?

Si $\theta_1, \ldots, \theta_4$ son las raíces de $h$ en $\mathbb{C}$, su discriminante es $\mathrm{disc}(h) = \prod_{i<j}(\theta_i - \theta_j)^2$. Consideremos

$$\Delta := \prod_{i<j}(\theta_i - \theta_j).$$

Claramente $\Delta^2 = \mathrm{disc}(h) = 229$ con lo que $F := \mathbb{Q}\left(\sqrt{229}\right)$ es un subcuerpo de $L = \mathbb{Q}(\theta_1, \ldots, \theta_4)$. Por otro lado una permutación $\sigma \in S_4$ actúa en $\Delta$ como

$$\Delta^\sigma = \epsilon(\sigma)\Delta.$$

Aplicando la Proposición 4.1 al polinomio $x^2 - 229$ vemos que

$$\zeta_F(s) = \zeta(s)L(U', s)$$

y deducimos que

$$L(U', s) = L(\chi, s)$$

donde $\chi(p) := \left(\frac{229}{p}\right)$ es el símbolo de Kronecker módulo 229. Es decir, la serie $L$ de Artin de $U'$ es la serie clásica de Dirichlet asociada a $\chi$.

Notemos que para un primo $p \neq 229$ se tiene que

$$\epsilon(\mathrm{Frob}_p) = \chi(p)$$

o, lo que es lo mismo, $\mathrm{Frob}_p$ consiste de permutaciones pares si y sólo si (gracias a la reciprocidad cuadrática) $p$ es un cuadrado módulo 229.

Esto se puede ver en la siguiente tabla donde agregamos una columna con los valores de $\chi(p)$ a la tabla 1.

| $p$ | $\tau_p$ | $\chi(p)$ |
|---|---|---|
| 2 | [4] | -1 |
| 3 | [3, 1] | 1 |
| 5 | [3, 1] | 1 |
| 7 | [4] | -1 |
| 11 | [3, 1] | 1 |
| 13 | [4] | -1 |
| 17 | [3, 1] | 1 |
| 19 | [3, 1] | 1 |
| 23 | [2, 1, 1] | -1 |
| 29 | [2, 1, 1] | -1 |
| 31 | [4] | -1 |
| 37 | [2, 2] | 1 |
| 41 | [4] | -1 |
| 43 | [3, 1] | 1 |
| 47 | [2, 1, 1] | -1 |

La serie $L$ de Artin de $V'$ se deduce de lo que ya vimos usando que $V' = V \otimes U'$. Si

$$L(V, s) = \sum_{n \geq 1} a_n n^{-s}$$

entonces

$$L(V', s) = \sum_{n \geq 1} \chi(n) a_n n^{-s}.$$

Nos queda describir $L(W,s)$. En paralelo a la descripción de $W$ que dimos en § 2.3 definimos

$$\eta_1 := (\theta_1+\theta_2)(\theta_3+\theta_4), \qquad \eta_2 := (\theta_1+\theta_3)(\theta_2+\theta_4), \qquad \eta_3 := (\theta_1+\theta_4)(\theta_2+\theta_3)$$

y

$$g(x) := (x-\eta_1)(x-\eta_2)(x-\eta_3).$$

El grupo de Galois $G$ preserva el conjunto $\{\eta_1,\eta_2,\eta_3\}$ y la acción correspondiente da lugar al homomorfismo sobreyectivo $S_4 \to S_3$. En particular, $g$ tiene coeficientes racionales; es la *resolvente cúbica* de $h$. Encontramos que

$$g(x) = x^3 - 4x + 1$$

que es irreducible de discriminante 229. En general, si $h = x^4+a_3x^3+a_2x^2+a_1x+a_0$ la resolvente es

$$g(x) = x^3 - 2a_2x^2 + (a_2^2 + a_1a_3 - 4a_0)x + a_1^2 - a_1a_2a_3 + a_0a_3^2.$$

Sea $F = \mathbb{Q}(\eta_1) \subseteq L$. Aplicando la Proposición 4.1 al polinomio $g$ vemos que

$$L(W,s) = \zeta_F(s)/\zeta(s).$$

Denotemos con $\sigma_p$ el tipo de factorización de $g$ módulo $p$. Obtenemos los siguientes valores (tabulados junto a los datos anteriores para comparar)

| $p$ | $\sigma_p$ | $\tau_p$ | $\chi(p)$ |
|---|---|---|---|
| 2 | [2,1] | [4] | -1 |
| 3 | [3] | [3, 1] | 1 |
| 5 | [3] | [3, 1] | 1 |
| 7 | [2,1] | [4] | -1 |
| 11 | [3] | [3, 1] | 1 |
| 13 | [2,1] | [4] | -1 |
| 17 | [3] | [3, 1] | 1 |
| 19 | [3] | [3, 1] | 1 |
| 23 | [2,1] | [2, 1, 1] | -1 |
| 29 | [2,1] | [2, 1, 1] | -1 |
| 31 | [2,1] | [4] | -1 |
| 37 | [1,1,1] | [2, 2] | 1 |
| 41 | [2,1] | [4] | -1 |
| 43 | [3] | [3, 1] | 1 |
| 47 | [2,1] | [2, 1, 1] | -1 |

Vemos de la tabla que tienen el mismo signo dado por $\chi$.

**4.3. Ejemplo II: polinomio de Trinks.** Terminamos con el ejemplo siguiente: $h := x^7 - 7x + 3$. En los años 60 Trinks [24] descubrió, justamente calculando sus tipos de factorización $\tau_p$ para suficientes primos $p$, que $h$ tiene grupo de Galois $G \simeq \mathrm{PSL}_2(\mathbb{F}_7)$. Éste es el famoso grupo simple de orden 168 asociado a Klein. Notemos que a priori un polinomio al azar de grado 7 tiene grupo de Galois $S_7$ de orden $7! = 5040$. Es decir que $h$ está lejos de ser un polinomio genérico.

En efecto, calculando para los primos $p \leq 10^{10}$ vemos que sólo aparecen los siguientes tipos de factorización (rutinas en PARI-GP para estos cálculos se pueden encontrar en [13]):

$$[7], \quad [4,2,1], \quad [3,3,1], \quad [2,2,1,1,1].$$

(Como ya vimos en §3.3 la clase de la identidad $[1,1,1,1,1,1,1]$, aunque es igual a $\tau_p$ con densidad positiva $1/|G|$, típicamente requiere que $p$ sea muy grande en relación a los otros tipos.)

Podemos también calcular la densidad aproximada

$$\delta_\tau(x) := \frac{\#\{p \le x \,|\, \tau_p = \tau\}}{\#\{p \le x\}}$$

para cada tipo $\tau$. Con $x = 10^{10}$ la mejor aproximación racional de $\delta_\tau(x)$ nos da lo siguiente

| $\tau$ | $\delta_\tau(10^{10})$ |
|---|---|
| $[7]$ | $2/7$ |
| $[4,2,1]$ | $1/4$ |
| $[3,3,1]$ | $1/3$ |
| $[2,2,1,1,1]$ | $1/8$ |

Si sumamos todas estas densidades incluyendo $1/168$ correspondiendo al tipo que falta $\tau = [1,1,1,1,1,1,1]$ obtenemos: $2/7 + 1/4 + 1/3 + 1/8 + 1/168 = 1$. Todo esto es consistente con que $G \simeq \mathrm{PSL}_2(\mathbb{F}_7)$.

De hecho, $\mathrm{PSL}_2(\mathbb{F}_7)$ tiene seis clases de conjugación pero tenemos sólo cinco tipos de factorización. Esto es el fenómeno que mencionamos anteriormente en §3.2. Dos de estas clases de conjugación ($7A$ y $7B$ en notación estándar, de orden 7) son indistinguibles vistas en $S_7$ vía la acción del grupo de Galois $G$ en las raíces de $h$. Es decir, no hay forma de saber a cual de estas dos clases pertenece $\mathrm{Frob}_p$ sólo sabiendo que $\tau_p = [7]$.

Tenemos también el otro fenómeno que mencionamos en §2.2. El grupo $G = \mathrm{PSL}_2(\mathbb{F}_7)$ tiene dos representaciones de permutación no equivalentes pero isomorfas como representaciones lineales. Esto se puede ver más claro usando que $G \simeq \mathrm{PGL}_3(\mathbb{F}_2)$. Sea $U \leqslant G$ el estabilizador de un punto del plano proyectivo de Fano $\mathcal{P}^2(\mathbb{F}_2)$ y $U'$ el estabilizador de una recta. Entonces las representaciones en las coclases de $U$ y $U'$ tienen esta propiedad. Concretamente podemos tomar los subgrupos

$$U := \begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix}, \qquad U' := \begin{pmatrix} * & 0 & 0 \\ * & * & * \\ * & * & * \end{pmatrix}$$

Estos subgrupos no son conjugados; por lo tanto los cuerpos $K, K' \subseteq L$ correspondientes por teoría de Galois no son isomorfos. Existe sin embargo una biyección $\phi : U \longrightarrow U'$, dada por transposición $u \mapsto u^t$, con la propiedad que $u$ y $\phi(u)$ son conjugadas. (Para ser claro: dada $u$ existe $g \in G$ tal que $gug^{-1} = u^t$ pero no es cierto que existe tal $g$ para *toda* $u$ simultáneamente.)

Concluimos que para cada clase de conjugación $C$ de $G$ se tiene que

$$\#(C \cap U) = \#(C \cap U'). \tag{4.5}$$

Esto muestra que las representaciones lineales correspondientes son isomorfas. En particular, se tiene la igualdad

$$\zeta_K(s) = \zeta_{K'}(s).$$

No es difícil dar una ecuación para el cuerpo $K'$, por ejemplo,

$$h'(x) := x^7 + 14x^4 - 42x^2 - 21x + 9.$$

## 5. Extensiones de cuerpos no finitas

**5.1. Correspondencia de Galois.** Queremos extender la relación entre extensiones y grupos de manera que permita considerar extensiones que no sean finitas. Al hacer esto, aparece un nuevo ingrediente que pasa desapercibido al trabajar con extensiones finitas, a saber *la topología*. Recordar que dar una topología en un conjunto es determinar que subconjuntos son abiertos, y cuales cerrados, satisfaciendo un conjunto de axiomas.

**Definición 5.1.** Un *espacio topológico* es un par $(X, \mathcal{B})$, donde $X$ es un conjunto, y $\mathcal{B}$ es un conjunto de subconjuntos de $X$ que satisface las siguientes propiedades:

1. El conjunto vacío y $X$ están en $\mathcal{B}$.
2. La unión de elementos de $\mathcal{B}$ es un elemento de $\mathcal{B}$.
3. La intersección finita de elementos de $\mathcal{B}$, es un elemento de $\mathcal{B}$.

Los elementos de $\mathcal{B}$ son los llamados *abiertos* del conjunto $X$, y los cerrados son por definición complementos de abiertos.

**Ejemplos 5.2.**
1. Los espacios métricos son espacios topológicos (tomando $\mathcal{B}$ el conjunto de abiertos). Por ejemplo, $(\mathbb{R}^n, |\,.\,|_2)$ es un espacio topológico.
2. Si $X$ es un conjunto arbitrario, podemos tomar $\mathcal{B}$ el conjunto de partes de $X$. Así, todo subconjunto es abierto y cerrado. Esta es la llamada topología discreta en $X$.
3. Si $X$ es un conjunto arbitrario, podemos tomar $\mathcal{B} = \{\emptyset, X\}$. Esta es la llamada topología trivial en $X$.

**Definición 5.3.** Si $G$ es un espacio topológico, y $* : G \times G \to G$ es una operación binaria tal que $(G, *)$ es un grupo, decimos que $(G, *)$ es un grupo topológico si vale que:

- El producto: $G \times G \to G$ es continuo.
- La función inversa: $\iota : G \to G$, $\iota(g) = g^{-1}$ es continua.

Recordar que una función es continua si vale que la preimagen de un conjunto abierto es abierto, donde en $G \times G$ tomamos la topología producto (o sea un abierto es unión de productos de abiertos de $G$).

**Ejemplos 5.4.**
1. Si $(G, *)$ es un grupo, entonces $(G, *)$ es un grupo topológico con la topología discreta (dado que todo subconjunto de $G$ es abierto).
2. $(\mathbb{R}, +)$ es un grupo topológico. Para ver esto, tomemos un abierto de $\mathbb{R}$, digamos $(a, b)$, y calculemos su preimagen por la operación suma. Luego estamos buscando todos los pares $(x, y) \in \mathbb{R}^2$ tales que $a < x + y < b$. Esto claramente es un abierto (que corresponde a la parte pintada de amarillo en la Figura 3).

*Ejercicio* 5.5. Probar que $(\mathbb{R} \setminus \{0\}, *)$ es un grupo topológico.

Si $G$ es un grupo topológico, y $g \in G$ es un elemento cualquiera, la función dada por traslación por $g$ es la función $m_g : G \to G$ dada por $m_g(h) = gh$. Dicha función es un homeomorfismo. La razón es que trasladar siempre es una función biyectiva (por tener inversa), y es continua por ser composición de las siguientes funciones

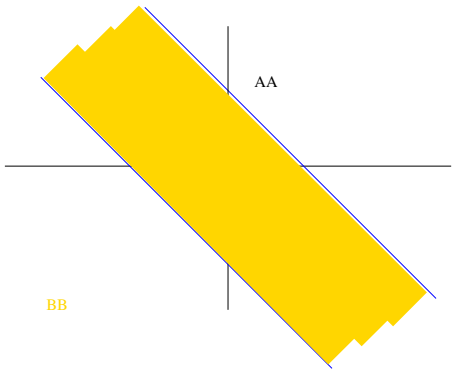



FIGURA 3. Continuidad suma

continuas:

$$G \hookrightarrow G \times G \xrightarrow{*} G$$

$$h \longrightarrow (g,h) \longrightarrow gh$$

En particular, si $H < G$ es un subgrupo abierto, todas sus coclases (que son de la forma $gH = m_g(H)$) son también abiertas. Luego, el grupo $G$ lo podemos escribir como la unión

$$G = H \cup \bigcup_{\substack{g \in G/H \\ g \notin H}} gH.$$

El último conjunto es abierto (por ser unión de abiertos), con lo cual $H$ es cerrado. Con esto hemos probado lo siguiente.

**Proposición 5.6.** *Si $H < G$ es un subgrupo abierto, entonces es cerrado.*

Si $L/K$ es una extensión de Galois (no necesariamente finita), y $S \subset L$ es un conjunto finito, definimos el conjunto

$$G(S) := \mathrm{Gal}(L/K)^S = \{\sigma \in \mathrm{Gal}(L/K) \ : \ \sigma(s) = s \ \forall s \in S\}. \tag{5.1}$$

**Lema 5.7.** *Los conjuntos $G(S)$ satisfacen las siguientes propiedades:*

1. *$G(S)$ es un subgrupo de $\mathrm{Gal}(L/K)$.*
2. *Si $S_1$ y $S_2$ son conjuntos finitos, $G(S_1) \cap G(S_2) = G(S_1 \cup S_2)$.*
3. *Si $S_1 \subset S_2$, $G(S_2) \subset G(S_1)$.*
4. *Si $S$ satisface que para todo $\tau \in \mathrm{Gal}(L/K)$, $\tau(S) = S$, entonces $G(S)$ es un subgrupo normal de $\mathrm{Gal}(L/K)$.*
5. *El subgrupo $G(S)$ tiene índice finito en $\mathrm{Gal}(L/K)$.*

*Demostración.* Las tres primeras propiedades son inmediatas de la definición. Para ver la cuarta afirmación, tomemos $\tau \in \mathrm{Gal}(L/K)$ y $\sigma \in G(S)$. Entonces $\tau(s) = \tilde{s} \in S$ por hipótesis. Luego

$$\tau^{-1}\sigma\tau(s) = \tau^{-1}(\sigma(\tilde{s})) = \tau^{-1}(\tilde{s}) = s.$$

La última afirmación se deduce de lo siguiente:

**Afirmación:** sin pérdida de generalidad, podemos suponer que $\tau(s) \in S$ para todo $\tau \in \mathrm{Gal}(L/K)$.

Si esto no fuera así, para cada elemento $s \in S$, el conjunto $\{\sigma(s) : \sigma \in \mathrm{Gal}(L/K)\}$ es finito, pues como $L/K$ es algebraica, y $s \in L$, $s$ es raíz de un polinomio con coeficientes en $K[x]$. En particular, $\sigma(s)$ debe ser otra raíz del mismo polinomio. Denotemos por $\overline{S} = \{\sigma(s) : s \in S, \sigma \in \mathrm{Gal}(L/K)\}$. Dicho conjunto es finito por ser $S$ finito, y el ítem 3 del lema implica que $G(\overline{S}) \subset G(S)$, con lo cual si probamos que $G(\overline{S})$ tiene índice finito en $\mathrm{Gal}(L/K)$, $G(S)$ también lo tiene.

Consideremos la extensión $K(S)$, que es la mínima subextensión de $L$ que contiene a $K$ y a todos los elementos de $S$. La extensión $K(S)/K$ es finita (¿por qué?) y Galois por la hipótesis en $S$. Notar que todo elemento de $G(S)$ es la identidad sobre los elementos de $K(S)$ (o sea la restricción de los elementos de $G(S)$ a $K(S)$ es la identidad). Luego obtenemos una sucesión exacta:

$$0 \longrightarrow G(S) \longrightarrow \mathrm{Gal}(L/K) \longrightarrow \mathrm{Gal}(K(S)/K)$$

$$\sigma \longrightarrow \sigma|_{K(S)}$$

Luego el índice de $G(S)$ en $\mathrm{Gal}(L/K)$ es menor o igual que $[K(S) : K] < \infty$. En realidad no es difícil ver que la sucesión anterior es exacta en todos lados (o sea restringir es suryectivo), pero no precisamos este resultado. □

**Proposición 5.8.** *Existe en* $\mathrm{Gal}(L/K)$ *una única topología para la cual los conjuntos* $G(S)$*, con* $S \subset L$ *finito, forman una base de entornos abiertos de la identidad. Esta topología hace de* $\mathrm{Gal}(L/K)$ *un grupo topológico.*

Antes de demostrar la proposición, entendamos que sucede cuando $L/K$ es finita (donde el Teorema de Galois no incluye topología alguna). En tal caso, si $\{l_1, \ldots, l_n\}$ es una base de $L$ como $K$-espacio vectorial, podemos tomar $S = \{l_1, \ldots, l_n\}$. Así, $G(S) = \{1\}$ (la identidad), o sea que cada punto del conjunto (finito) $\mathrm{Gal}(L/K)$ es abierto, con lo cual obtenemos la topología discreta. Luego es claro que dicha topología nos da una estructura de grupo topológico (dado que toda función es continua para ella).

Cuando la extensión $L/K$ no es finita, la topología obtenida no es la trivial, y es cuando realmente obtenemos algo distinto de la teoría clásica.

*Demostración de la Proposición 5.8.* Para abreviar la notación, denotemos por $G = \mathrm{Gal}(L/K)$, por $*$ la operación (en nuestro caso la composición) y por 1 el elemento neutro.

Por el Lema anterior, si $S_1$ y $S_2$ son conjuntos finitos, $G(S_1) \cap G(S_2) = G(S_1 \cup S_2)$. Luego $G(S)$ da una base de entornos de la identidad. Recordar que luego un entorno de un punto $g \in G$ está dado por $g * G(S)$ para $S \subset L$ finito. Debemos verificar que el producto con esta topología es continuo. Para ello debemos verificar las siguientes propiedades:

1. Si $S$ es finito, sea $g \in G(S)$, tal que $g = \sigma\tau$, con $\sigma, \tau \in G$. Entonces existen $S_1, S_2 \subset L$ finitos tales que $(\sigma * G(S_1)) * (\tau * G(S_2)) \subset G(S)$.
2. Si $S$ es finito y $g \in G(S)$, entonces existe $\tilde{S} \subset L$ finito tal que $(g * G(\tilde{S}))^{-1} = G(\tilde{S})^{-1} * g^{-1} \subset G(S)$.

La primer propiedad dice que el producto es continuo en un entorno del 1, mientras que la segunda dice que la función $g \mapsto g^{-1}$ es continua en un entorno del 1.

Para demostrar (1), consideremos el conjunto $G(\overline{S})$ (ver la demostración del Lema 5.7 para la notación). Claramente $G(\overline{S}) \subset G(S)$ (por la segunda propiedad del lema), y $G(\overline{S}) \lhd G$ (o sea es un subgrupo normal). Luego,

$$(\sigma * G(\overline{S})) * (\tau * G(\overline{S})) = \sigma * \tau * (\tau^{-1} * G(\overline{S}) * \tau) * G(\overline{S}) = \sigma * \tau * G(\overline{S}) = g * G(\overline{S}) \subset G(S).$$

La demostración de (2) es inmediata, dado que $G(S)$ es un subgrupo, luego $G(S)^{-1} = G(S)$ y claramente $g^{-1}G(S) \subset G(S)$ (por definición). □

**Definición 5.9.** Si $L/K$ es una extensión de Galois, el grupo de Galois topológico de la extensión (que haciendo abuso de notación lo denotamos también $\mathrm{Gal}(L/K)$) es el grupo $\mathrm{Gal}(L/K)$, con la única topología para la cual una base de entornos de la identidad son los conjuntos $G(S)$. Dicha topología se denomina la *topología de Krull*.

**Proposición 5.10.** *Si $L/K$ es una extensión de Galois, entonces el grupo topológico $\mathrm{Gal}(L/K)$ es Hausdorff, compacto y totalmente disconexo.*

*Demostración.* Veamos que es Hausdorff: sean $\sigma, \tau \in \mathrm{Gal}(L/K)$ distintos. Luego, existe $\alpha \in L$ tal que $\sigma(\alpha) \neq \tau(\alpha)$. Miremos $S = \{\alpha\}$. Luego $\sigma G(S)$ y $\tau G(S)$ son abiertos, y claramente disjuntos.

Veamos que es totalmente disconexo: alcanza con ver que la componente conexa de la identidad, denotada $\mathcal{C}_{\mathbb{1}}$, es sólo la identidad. Sea $S$ un conjunto finito y $\mathrm{Gal}(L/K)$ estable (o sea $\sigma(S) = S$ para todo $\sigma \in \mathrm{Gal}(L/K)$). Entonces $G(S)$ es un abierto que contiene a la identidad. Como $G(S)$ es abierto, por la Proposición 5.6 $G(S)$ es también cerrado, con lo cual la componente conexa de la identidad está contenida en $G(S)$. Luego

$$\mathcal{C}_{\mathbb{1}} \subset \bigcap_{\substack{S \subset L \\ \#S < \infty}} G(S) = \{\mathbb{1}\}.$$

Probar que $\mathrm{Gal}(L/K)$ es compacto es más complicado, y proviene de dar dicho grupo de Galois como un límite inverso sobre todas las subextensiones finitas, y utilizar el Teorema de Tychonoff (ver el capítulo 1 de [3] por ejemplo). □

Por completitud, enunciamos la correspondencia de Galois para extensiones de Galois arbitrarias.

**Teorema 5.11** (Galois)**.** *Sea $L/K$ una extensión Galois. Entonces existe una correspondencia biyectiva entre el conjunto de extensiones: $\{N : K \subset N \subset L\}$ y el conjunto de subgrupos cerrados de $\mathrm{Gal}(L/K)$. La biyección esta dada por:*

$$N \quad \to \quad \mathrm{Aut}_N(L) \subset \mathrm{Gal}(L/K) \tag{5.2}$$

$$L^H := \{\alpha \in L \, : \, \sigma(\alpha) = \alpha \, \forall \sigma \in H\} \quad \leftarrow \quad H \tag{5.3}$$

*Además, la extensión $N/K$ es Galois si y sólo si el grupo $\mathrm{Gal}(L/N)$ es un subgrupo normal de $\mathrm{Gal}(L/K)$. En el caso que la extensión $N/K$ sea Galois, $\mathrm{Gal}(N/K) \simeq \mathrm{Gal}(L/K)/\mathrm{Gal}(L/N)$.*

**5.2. Grupo de descomposición y Frobenius en extensiones infinitas.** De manera análoga a lo hecho en la sección 3.2, podemos definir el anillo de enteros de $\overline{\mathbb{Q}}$, esto es:

$$\overline{\mathbb{Z}} := \{\alpha \in \overline{\mathbb{Q}} \ : \ \alpha \text{ es entero algebraico}\}. \tag{5.4}$$

Recordar que $\alpha$ es entero algebraico si es raíz de un polinomio con coeficientes enteros mónico. Como en el caso de extensiones finitas, $\overline{\mathbb{Z}}$ es un anillo. Sea $p \in \mathbb{Z}$ un primo, y sea $\mathfrak{p} \subset \overline{\mathbb{Z}}$ un ideal maximal que contiene a $p$.

*Ejercicio* 5.12. Probar que $\overline{\mathbb{Z}}/\mathfrak{p} \simeq \overline{\mathbb{F}_p}$.

**Definición 5.13.** El grupo de descomposición de $\mathfrak{p}$ es

$$D_{\mathfrak{p}} := \{\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \ : \ \sigma(\mathfrak{p}) = \mathfrak{p}\}.$$

Claramente $D_{\mathfrak{p}}$ es un subgrupo de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. A pesar de que el grupo de descomposición $D_{\mathfrak{p}}$ depende del ideal $\mathfrak{p}$, si $\mathfrak{p}, \tilde{\mathfrak{p}}$ son dos ideales que contienen a $p$, entonces existe $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ tal que $\sigma(\mathfrak{p}) = \tilde{\mathfrak{p}}$. Luego $D_{\mathfrak{p}} = \sigma^{-1} D_{\tilde{\mathfrak{p}}} \sigma$.

Si $\sigma \in D_{\mathfrak{p}}$, $\sigma$ induce una función en $\overline{\mathbb{Z}}/\mathfrak{p}$, que es la identidad sobre $\mathbb{F}_p$, con lo cual obtenemos una función

$$\varphi : D_{\mathfrak{p}} \mapsto \mathrm{Aut}_{\mathbb{F}_p}(\overline{\mathbb{F}_p}) = \mathrm{Gal}(\overline{\mathbb{F}_p}/\mathbb{F}_p). \tag{5.5}$$

*Ejercicio* 5.14. Verificar las siguientes propiedades:

1. El subgrupo $D_{\mathfrak{p}}$ es un subgrupo cerrado (equivalentemente, su complemento es abierto). Luego por la correspondencia de Galois $D_{\mathfrak{p}} = \mathrm{Gal}(\overline{\mathbb{Q}}/L)$, donde $L = \overline{\mathbb{Q}}^{D_{\mathfrak{p}}}$.
2. El morfismo $\varphi$ es continuo, donde miramos a $D_{\mathfrak{p}}$ como grupo topológico con la topología de subgrupo (o con la topología de Krull identificándolo con $\mathrm{Gal}(\overline{\mathbb{Q}}/\overline{\mathbb{Q}}^{D_{\mathfrak{p}}})$).
3. El morfismo $\varphi$ es suryectivo. Sugerencia: probar primero que las funciones $\varphi_n : D_{\mathfrak{p}} \to \mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p)$ (componer $\varphi$ con la función cociente) es suryectivo para todo $n$ utilizando lo visto en extensiones finitas. Deducir que la imagen de $\varphi$ es densa en $\mathrm{Gal}(\overline{\mathbb{F}_p}/\mathbb{F}_p)$, y utilizar el ítem anterior.

Se puede ver que el grupo $D_{\mathfrak{p}}$ es isomorfo al grupo de Galois $\mathrm{Gal}(\overline{\mathbb{Q}_p}/\mathbb{Q}_p)$.

**Definición 5.15.** El grupo de inercia $I_{\mathfrak{p}}$ de $\mathfrak{p}$ es el núcleo del morfismo $\varphi$, o sea

$$I_{\mathfrak{p}} = \{\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \ : \ \sigma(x) \equiv x \pmod{\mathfrak{p}} \text{ para todo } x \in \overline{\mathbb{Z}}\}.$$

Llamamos un *Frobenius absoluto sobre* $p$ a cualquier elemento $\mathrm{Frob}_{\mathfrak{p}} \in D_{\mathfrak{p}}$ en la preimagen de $\mathrm{Frob}_p \in \mathrm{Gal}(\overline{\mathbb{F}_p}/\mathbb{F}_p)$. Claramente dos Frobenius absolutos difieren por un elemento en el subgrupo de inercia.

Como sucede con el grupo de descomposición, si $\mathfrak{p}, \tilde{\mathfrak{p}}$ son dos ideales de $\overline{\mathbb{Z}}$ que contienen a un primo $p$, entonces existe $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ tal que $\sigma(\mathfrak{p}) = \tilde{\mathfrak{p}}$ y entonces $I_{\mathfrak{p}} = \sigma^{-1} I_{\tilde{\mathfrak{p}}} \sigma$.

Por último, obtenemos la siguiente versión del Teorema de Chebotarev.

**Teorema 5.16** (Chebotarev)**.** *Para todo número primo $p$, salvo finitos, tomemos para cada ideal primo $\mathfrak{p}$ que contiene a $p$ un Frobenius absoluto* $\mathrm{Frob}_{\mathfrak{p}}$*. Entonces el conjunto* $\{\mathrm{Frob}_{\mathfrak{p}}\}_{\mathfrak{p}|p} \subset \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ *es denso.*

## 6. Representaciones de Galois

El objetivo deseado es el de poder entender el grupo $G_{\mathbb{Q}} := \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ (o su análogo para otros cuerpos como extensiones finitas de $\mathbb{Q}$ por ejemplo). Lamentablemente, dicho grupo es muy difícil de manejar (y ni siquiera está bien definido, dado que depende de elegir una clausura algebraica de $\mathbb{Q}$). Una manera de entender un grupo es mediante sus representaciones, esto es entender como actúa en espacios vectoriales de dimensión finita. En varios casos, conocer dichas representaciones es suficiente para determinar el grupo (y sus propiedades) unívocamente, por ejemplo es lo que sucede al considerar grupos finitos. A pesar de que $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ no es finito, el mismo es un límite (inverso) de grupos finitos, con lo cual es de esperar poder recuperar bastante información si logramos entender todas sus representaciones.

Durante este capítulo, $K$ va a denotar un cuerpo topológico (principalmente nos concentraremos en los casos $\mathbb{C}$, $\mathbb{Q}_p$ o una extensión finita de $\mathbb{Q}_p$), y $V$ un $K$-espacio vectorial de dimensión finita $d$. En particular, $V$ admite una topología a partir de la de $K$. Lo mismo sucede para el grupo $\mathrm{End}_K(V)$ (grupo de transformaciones lineales de $V$ en sí mismo), ya que una vez que elegimos una base para $V$, podemos identificar $\mathrm{End}_K(V)$ con el conjunto $M_{d\times d}(K)$ de matrices con $d$ filas y $d$ columnas, y este es isomorfo con $K^{d^2}$.

**Definición 6.1.** Una representación de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ es un morfismo continuo

$$\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_K(V).$$

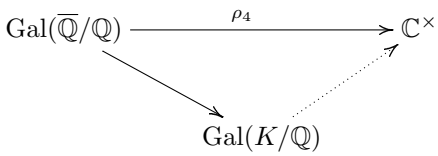
Para poder comenzar a entender que es una representación de Galois, miremos algunos ejemplos sencillos.

**Ejemplos 6.2.** 1. En el caso en que $V$ tiene dimensión 1, cualquier representación es un morfismo entre $G_{\mathbb{Q}}$ y $K^\times$ (dado que a un elemento $v$ de $V$, lo manda a un múltiplo suyo no nulo).

Sea $K = \mathbb{Q}(i)$ el cuerpo de raíces cuartas de la unidad. Si $\sigma \in G_{\mathbb{Q}}$, $\sigma(i) = \pm i$, con lo cual podemos definir una representación $\rho_4 : G_{\mathbb{Q}} \to \mathbb{C}^\times$ por

$$\rho_4(\sigma) = \frac{\sigma(i)}{i} \in \{\pm 1\}. \tag{6.1}$$

Notar que $\rho_4$ es trivial en el subgrupo $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$, dado que dichos elementos actúan trivialmente en $i$. Luego $\rho_4$ factoriza por $\mathrm{Gal}(K/\mathbb{Q})$, o sea tenemos el siguiente diagrama




Notar que $\mathrm{Gal}(K/\mathbb{Q}) \simeq \mathbb{Z}/2$, y $\rho_4$ es justamente la representación no trivial de dicho grupo. Esto demuestra que efectivamente $\rho_4$ es un morfismo de grupos. Nos resta ver que $\rho_4$ es continua.

El conjunto $\{\pm 1\} \subset \mathbb{C}^\times$ es discreto, con lo cual su topología es la discreta. Para verificar que $\rho_4$ es continua, basta con ver que $\ker(\rho_4)$ (su núcleo) es abierto. Pero $\mathrm{Gal}(\overline{\mathbb{Q}}/K) = G(\{i\})$ (siguiendo la notación del capítulo anterior), con lo cual es abierto y tiene índice finito en $G_{\mathbb{Q}}$ (ver Lema 5.7). De esto se deduce que $\ker(\rho_4)$ es abierto, como queríamos ver.

2. Mas generalmente, si $n \in \mathbb{N}$, $n \geq 3$, y tomamos $K = \mathbb{Q}(\xi_n)$ (el cuerpo ciclotómico de raíces $n$-ésimas de la unidad), el grupo $\mathrm{Gal}(K/\mathbb{Q})$ es un grupo abeliano finito. En particular, todas sus representaciones irreducibles son de dimensión 1. Si $\chi : \mathrm{Gal}(K/\mathbb{Q}) \to \mathbb{C}^\times$ es una tal representación, podemos definir $\rho_\chi : G_\mathbb{Q} \to \mathbb{C}^\times$, mediante el diagrama

$$\begin{array}{ccc} \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) & \xrightarrow{\quad\rho_\chi\quad} & \mathbb{C}^\times \\ & \searrow \quad \nearrow_{\chi} & \\ & \mathrm{Gal}(K/\mathbb{Q}) & \end{array}$$

Dejamos como ejercicio probar que dicho morfismo es continuo.

**Proposición 6.3.** *Sean $\rho_1, \rho_2$ dos representaciones de Galois. Si para todo número primo $p$, salvo finitos, y para cada ideal primo $\mathfrak{p}$ que contiene a $p$ vale que $\rho_1(\mathrm{Frob}_\mathfrak{p}) = \rho_2(\mathrm{Frob}_\mathfrak{p})$ entonces $\rho_1 = \rho_2$.*

*Demostración.* Por el Teorema 5.16 el conjunto $\{\mathrm{Frob}_\mathfrak{p}\}$ es denso en $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Como las representaciones son continuas, si coinciden en un conjunto denso, coinciden en todos los elementos. □

Mas adelante veremos cómo reemplazar primos de $\overline{\mathbb{Z}}$ por primos de $\mathbb{Z}$.

**Definición 6.4.** Si $\rho_1, \rho_2$ son dos representaciones de Galois de $G_\mathbb{Q}$ en $\mathrm{Aut}_K(V)$, decimos que son equivalentes, si existe una transformación lineal $\psi \in \mathrm{Aut}_K(V)$ tal que para todo $\sigma \in G_\mathbb{Q}$ y todo $v \in V$,

$$\rho_1(\sigma)(v) = \psi^{-1}(\rho_2(\sigma)(\psi(v))).$$

Sea $\rho$ una representación de Galois de $G_\mathbb{Q}$ en $\mathrm{Aut}_K(V)$, donde $V$ es un $K$-espacio vectorial de dimensión $d$. Elegir una base $B$ de $V$ induce un isomorfismo $V \simeq K^d$ y una representación de Galois

$$\rho_B : G_\mathbb{Q} \to \mathrm{GL}_d(K).$$

Elegir otra base de $V$ da representaciones equivalentes, dado que si $B'$ es otra base, y $\mathcal{C}$ es la matriz de cambio de base de $B$ a $B'$, entonces las representaciones $\rho_B$ y $\rho_{B'}$ difieren en conjugar por $\mathcal{C}$. Es por esto que no vamos a distinguir entre representaciones de $G_\mathbb{Q}$ en $\mathrm{Aut}_K(V)$ o en $\mathrm{GL}_d(K)$.

**Teorema 6.5.** *Si $V$ es un $\mathbb{C}$-espacio vectorial de dimensión finita, y $\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_\mathbb{C}(V)$ es una representación de Galois, entonces $\rho$ factoriza por una extensión finita; o sea existe $L$ extensión finita de $\mathbb{Q}$, y $\tilde{\rho} : \mathrm{Gal}(L/\mathbb{Q}) \to \mathrm{Aut}_\mathbb{C}(V)$ tal que el siguiente diagrama conmuta*

$$\begin{array}{ccc} \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) & \xrightarrow{\quad\rho\quad} & \mathrm{Aut}_\mathbb{C}(V) \\ & \searrow \quad \nearrow_{\tilde{\rho}} & \\ & \mathrm{Gal}(L/\mathbb{Q}) & \end{array}$$

*Demostración.* Supongamos primero que $d = 1$, o sea que tenemos un morfismo continuo de $G_\mathbb{Q}$ en $\mathbb{C}^\times$. Si tomamos un abierto de $\mathbb{C}^\times$ cerca del 1, digamos $U = \{z \in \mathbb{C} \,:\, |z-1| < \frac{1}{2}\}$, entonces su preimagen es un abierto de $G_\mathbb{Q}$ (por ser $\rho$ continua).

Como la función identidad está en $\rho^{-1}(U) \subset G_{\mathbb{Q}}$, hay un abierto centrado en la matriz identidad que esta contenido en la preimagen, o sea existe un conjunto finito $S \subset \overline{\mathbb{Q}}$ (que podemos suponer $G_{\mathbb{Q}}$ estable) tal que $G(S) \subset \rho^{-1}(U)$. En particular, $\rho(G(S)) \subset U$. Como $G(S)$ es un subgrupo de $G_{\mathbb{Q}}$ y $\rho$ es un morfismo, $\rho(G(S))$ es un subgrupo de $U$. Pero el único subgrupo de $U$ es $\{1\}$ (convencerse). En particular, $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(S))$ está en el núcleo de $\rho$ y por lo tanto $\rho$ factoriza por $\mathrm{Gal}(\mathbb{Q}(S)/\mathbb{Q})$ como queríamos ver.

Cuando $d > 1$, la demostración es similar, con la dificultad de que tenemos que demostrar que existe un entorno de la matriz identidad (de $d \times d$) que no contiene subgrupos no triviales. Supongamos que tomando una bola (abierta) centrada en la identidad de radio $\epsilon$ (pequeño y fijo) tenemos un subgrupo $H$ no trivial. Es fácil ver que todos los elementos de $H$ son diagonalizables (¿por qué?). Si $M$ es un elemento de $H$, y $\lambda$ es un autovalor de $M$, entonces $|\lambda - 1| \leq \frac{1}{2}$ (para la elección adecuada de $\epsilon$, que no depende de $M$, ver Ejercicio 6.6). Pero como $M \in H$, sus potencias también lo están, con lo cual $|\lambda^n - 1| \leq \frac{1}{2}$ para todo $n$ entero. Luego $\lambda = 1$, y $M$ es la matriz identidad. □

*Ejercicio* 6.6. Para el alumno que no vio los detalles de métricas utilizados en la última demostración, les dejamos algunas propiedades que cumplen los espacios vectoriales en dimensión finita.

1. Probar que dos normas cualesquiera en un espacio vectorial $V$ de dimensión finita son equivalentes, esto es, si $\|.\|_1$ , $\|.\|_2$ son normas en $V$, entonces existe una constante $C > 0$ tal que $\|v\|_1 \leq C\|v\|_2$ para todo $v \in V$.
2. Si $K$ es un cuerpo con un valor absoluto $|.|$, y tomamos cualquiera de las normas usuales en $K^d$, entonces podemos asociar una norma de *operadores* en $M_{d\times d}(K)$ de la siguiente manera: si $M \in M_{d\times d}(K)$ definimos
$$\|M\| := \sup_{\substack{v\in K^d \\ \|v\|=1}} \|Mv\|.$$
Probar que $\|.\|$ es una norma.
3. Probar que si $\lambda \in K$ es un autovalor de $M$, entonces $|\lambda| \leq \|M\|$.
4. Completar los detalles de la demostración anterior para $d > 1$.

El Teorema 6.5 nos dice que las representaciones de Galois complejas son interesantes, pero si trabajamos exclusivamente con ellas, sólo vamos a obtener información sobre las extensiones finitas de $\mathbb{Q}$. Es por esto que hay que considerar también representaciones en espacios vectoriales sobre $\mathbb{Q}_p$ (las llamadas representaciones de Galois $p$-ádicas).

Si $\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_K(V)$ es una representación de Galois, decimos que un subespacio $W \subset V$ es *$\rho$-invariante* (o que $W$ es una subrepresentación) si $\rho(\sigma)(w) \in W$ para todo $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ y todo $w \in W$.

**Definición 6.7.** Una representación $\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_K(V)$ se dice *irreducible* si no existe $W \subset V$ subespacio propio y no trivial $\rho$-invariante. Una representación $\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_K(V)$ se dice *semi-simple* si $V$ se puede descomponer como
$$V = \bigoplus_i V_i,$$
donde cada $V_i$ es un subespacio $\rho$-invariante, y las representaciones $\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_K(V_i)$ son irreducibles.

**Teorema 6.8.** *Si $\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_{\mathbb{C}}(V)$ una representación de Galois, donde $V$ un $\mathbb{C}$-espacio vectorial de dimensión finita, $\rho$ es semi-simple.*

*Demostración.* Por el Teorema 6.5 sabemos que $\rho$ factoriza por una extensión finita, o sea existe $L/\mathbb{Q}$ Galois y finita y $\tilde{\rho} : \mathrm{Gal}(L/\mathbb{Q}) \to \mathrm{Aut}_{\mathbb{C}}(V)$ tal que $\rho = \tilde{\rho} \circ \Pi$, donde $\Pi : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Gal}(L/\mathbb{Q})$ es la proyección (que corresponde a restringir un automorfismo al cuerpo $L$). Luego el resultado se sigue del Teorema 2.2. □

*Observación* 6.9. Como sucedía con grupos abelianos no finitos, no vale en general que toda representación de Galois sea semi-simple, usamos fuertemente que el cuerpo $K$ era el cuerpo de números complejos.

**Definición 6.10.** Sea $\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_K(V)$ una representación de Galois. Dado un primo $p$, decimos que $\rho$ es *no-ramificada* en $p$ si $\rho(I_{\mathfrak{p}}) = 1$ (la identidad) para cualquier primo $\mathfrak{p} \subset \overline{\mathbb{Z}}$ que contiene a $p$.

Veamos que efectivamente la condición de $\rho$ ser no-ramificada en $p$ no depende del primo $\mathfrak{p}$. Notar que si $\mathfrak{p}$, $\tilde{\mathfrak{p}}$ son dos ideales que contienen al mismo primo $p$, entonces los grupos $I_{\mathfrak{p}}$, $I_{\tilde{\mathfrak{p}}}$ son conjugados, con lo cual $\rho$ es trivial en uno de ellos si y sólo si lo es en el otro. Decimos que $\rho$ es ramificada en $p$ si no es no-ramificada (equivalente, la imagen del grupo de inercia es no-trivial).

Recordar que dado $\mathfrak{p} \subset \overline{\mathbb{Z}}$, un Frobenius absoluto sobre $p$, denotado por $\mathrm{Frob}_{\mathfrak{p}}$, es un elemento cualquiera en el grupo de descomposición $D_{\mathfrak{p}}$ cuya imagen residual corresponde al automorfismo de Frobenius. Dos tales morfismos difieren en un elemento del grupo de inercia $I_{\mathfrak{p}}$. Ahora si $\rho$ es no ramificada en $p$, entonces $\rho(I_{\mathfrak{p}}) = 1$, con lo cual el valor $\rho(\mathrm{Frob}_{\mathfrak{p}})$ depende solamente del ideal $\mathfrak{p}$. Además, si $\mathfrak{p}, \tilde{\mathfrak{p}}$ son dos ideales en $\overline{\mathbb{Z}}$ que contienen a $p$, entonces uno es conjugado del otro. Luego $\rho(\mathrm{Frob}_{\mathfrak{p}})$ y $\rho(\mathrm{Frob}_{\tilde{\mathfrak{p}}})$ son matrices conjugadas (identificando $\mathrm{Aut}_K(V)$ con el grupo de matrices). En particular, sus polinomios característicos son iguales.

Dada una matriz cuadrada $M$, denotemos por $\mathrm{car}(M) = \det(1 - T \cdot M)$ a su polinomio característico. De lo expuesto anteriormente, tenemos que si $\rho$ es no ramificada en $p$, podemos definir

$$\mathrm{car}(\rho(\mathrm{Frob}_p)) := \mathrm{car}(\rho(\mathrm{Frob}_{\mathfrak{p}})), \tag{6.2}$$

donde $\mathfrak{p} \subset \overline{\mathbb{Z}}$ es cualquier ideal maximal tal que $p \in \mathfrak{p}$.

Para los primos ramificados, la situación es igual que lo expuesto en la Sección 4 para extensiones finitas: dado un primo $\mathfrak{p} \subset \overline{\mathbb{Z}}$, el grupo de inercia $I_{\mathfrak{p}}$ actúa en $V$. Denotemos por $V^{I_{\mathfrak{p}}}$ al subespacio de vectores donde la inercia actúa trivialmente. Por ejemplo, si $p$ es un primo no ramificado de $\rho$, entonces $V^{I_{\mathfrak{p}}} = V$.

*Ejercicio* 6.11. Probar que el subespacio $V^{I_{\mathfrak{p}}}$ es invariante por la acción de $D_{\mathfrak{p}}$, o sea si $v \in V^{I_{\mathfrak{p}}}$ y $\sigma \in D_{\mathfrak{p}}$ entonces $\rho(\sigma)(v) \in V^{I_{\mathfrak{p}}}$.

La restricción $\rho|_{V^{I_{\mathfrak{p}}}}$ da una acción de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ en el subespacio $V^{I_{\mathfrak{p}}}$. Por definición, $I_{\mathfrak{p}}$ actúa trivialmente en $V^{I_{\mathfrak{p}}}$, con lo cual la matriz $\rho|_{V^{I_{\mathfrak{p}}}}(\mathrm{Frob}_{\mathfrak{p}})$ está bien definida (o sea no depende de la preimagen de $\mathrm{Frob}_{\mathfrak{p}}$ escogida). Mas aún, dicho valor depende exclusivamente del primo $p$ contenido en $\mathfrak{p}$.

*Ejercicio* 6.12. Sean $\mathfrak{p}$ y $\tilde{\mathfrak{p}}$ son dos primos en $\overline{\mathbb{Z}}$ que contienen al mismo primo $p$. En particular, existe $\tau \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ tal que $\tau\mathfrak{p} = \tilde{\mathfrak{p}}$.

- Probar que $\tau I_{\mathfrak{p}} \tau^{-1} = I_{\tilde{\mathfrak{p}}}$ y que $\tau V^{I_{\mathfrak{p}}} = V^{I_{\tilde{\mathfrak{p}}}}$.

- Probar que el polinomio $\mathrm{car}(\rho|_{V^{I_\mathfrak{p}}}(\mathrm{Frob}_\mathfrak{p}))$ no depende de $\mathfrak{p}$.

Luego podemos definir $\mathrm{car}(\rho|_{V^{I_p}}(\mathrm{Frob}_p))$ como $\mathrm{car}(\rho|_{V^{I_\mathfrak{p}}}(\mathrm{Frob}_\mathfrak{p}))$ para cualquier $\mathfrak{p} \subset \overline{\mathbb{Z}}$ maximal que contenga a $p$.

**Teorema 6.13** (Brauer-Nesbitt)**.** *Sean $\rho_i : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_K(V)$, $i = 1, 2$ dos representaciones semi-simples. Si para todo elemento $s$ de un conjunto denso $S \subset \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ vale*

$$\mathrm{car}(\rho_1(s)) = \mathrm{car}(\rho_2(s)),$$

*entonces $\rho_1$ y $\rho_2$ son isomorfas.*

*Demostración.* Ver [2, 30.16]. Ver también el Teorema 2.4.6 de las notas de Gabor Wiese (en `math.uni.lu/~wiese/notes/GalRep.pdf`). La demostración involucra todos los elementos de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, pero por ser las representaciones continuas, basta verificarla en un conjunto denso. □

Si el cuerpo $K$ tiene característica cero, además basta con verificar que las trazas de las representaciones son iguales, sin necesidad de calcular todo el polinomio característico.

**Corolario 6.14.** *Sean $\rho_i : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_K(V)$, $i = 1, 2$ dos representaciones semi-simples. Supongamos que ambas representaciones son ramificadas solamente en un conjunto finito de primos $S$. Si*

$$\mathrm{car}(\rho_1(\mathrm{Frob}_p)) = \mathrm{car}(\rho_2(\mathrm{Frob}_p)),$$

*para todo primo $p \notin S$, entonces $\rho_1$ y $\rho_2$ son isomorfas. Además, si $K$ tiene característica 0, basta verificar que $\mathrm{Tr}(\rho_1(\mathrm{Frob}_p)) = \mathrm{Tr}(\rho_2(\mathrm{Frob}_p))$ para todo $p \notin S$.*

Una pregunta natural es cuando las representaciones de Galois son ramificadas solamente en un conjunto finito de primos.

**Proposición 6.15.** *Si $\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_\mathbb{C}(V)$ es una representación de Galois, entonces $\rho$ es no ramificada fuera de un conjunto finito de primos.*

*Demostración.* Por Teorema 6.5 $\rho$ factoriza por una extensión Galois finita $L/\mathbb{Q}$, o sea existe una representación $\tilde{\rho} : \mathrm{Gal}(L/\mathbb{Q}) \to \mathrm{Aut}_\mathbb{C}(V)$ tal que $\rho = \tilde{\rho} \circ \Pi$, donde $\Pi : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Gal}(L/\mathbb{Q})$ es la proyección. Es conocido que en una extensión finita hay un conjunto finito de primos que ramifican (precisamente aquellos que dividen al discriminante de la extensión). Luego $\tilde{\rho}$ es no ramificada fuera de un conjunto finito de primos y por consiguiente $\rho$ también lo es. □

Para un cuerpo general, no es cierto que una representación de Galois semi-simple sea no ramificada fuera de un conjunto finito de primos (ver por ejemplo [12]). Sin embargo, el conjunto de primos ramificados siempre tiene densidad cero (ver [5]), con lo cual el Teorema de Brauer-Nesbitt se puede aplicar en general.

### 6.1. Series $L$ asociadas a representaciones del grupo de Galois absoluto.

Las series $L$ son objetos analíticos, que se utilizan para *codificar* información asociada a diversos objetos geométricos/aritméticos/algebraicos. Un primer ejemplo es la función zeta de Riemann, definida como

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \qquad s \in \mathbb{C}.$$

Es fácil ver que dicha serie converge en el semiplano $\Re(s) > 1$. Riemann en 1859 demostró que dicha función se puede extender de forma analítica a todo el plano complejo, probando que la misma satisface una ecuación funcional que relaciona el valor en $s$ con el valor en $1 - s$. La importancia de la función zeta de Riemann es que admite una descomposición como

$$\zeta(s) = \prod_{p \text{ primo}} \frac{1}{1 - p^{-s}},$$

con lo cual la misma da mucha información sobre la distribución de los números primos. Existen muchas generalizaciones de la función zeta de Riemann, como la asociada a caracteres por Dirichlet, la asociada a cuerpos de números por Dedekind y muchas otras (como vimos en la Sección 4). El objetivo nuestro es asociarle a una representación de Galois una función analítica similar.

**Definición 6.16.** Dada $\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_K(V)$, definimos la $L$-serie $L(\rho, s)$ asociada a $\rho$ como:

$$L(\rho, s) = \prod_{p \text{ primo}} L_p(\rho, p^{-s})^{-1} = \prod_{p \text{ primo}} \frac{1}{\mathrm{car}(\rho|_{V^{I_p}}(\mathrm{Frob}_p))(p^{-s})}, \tag{6.3}$$

donde $\mathrm{car}(\rho|_{V^{I_p}}(\mathrm{Frob}_p))$ es el polinomio característico definido en el Ejercicio 6.12, evaluado en $p^{-s}$.

**Ejemplo 6.17.** Sea $\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_K(V)$ la representación trivial, con $\dim(V) = 1$, o sea $\rho(\sigma)(v) = v$. Entonces todos los primos son no ramificados y vale que $\chi(\rho(\mathrm{Frob}_p)) = 1 - t$. Luego,

$$L(\rho, s) = \prod_{p \text{ primo}} \frac{1}{(1 - p^{-s})} = \zeta(s).$$

Como sucedió en el ejemplo de la Sección 4.2, se puede ver que las $L$-series de Dedekind y de Dirichlet se obtienen como series L asociadas a representaciones complejas de Galois (las series L de Dirichlet corresponden a representaciones de dimensión 1, mientras que la función zeta de Dedekind corresponden a tomar la representación regular del grupo de Galois $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$, para $K/\mathbb{Q}$ finita como se explicó en la Proposición 4.1).

Las $L$-series mas estudiadas son las que corresponden a los siguientes dos casos: $K = \mathbb{C}$, que corresponden a representaciones de Galois complejas y fueron estudiadas por Artin (vistas en la Sección 4); y $K = \mathbb{Q}_p$, las llamadas representaciones de Galois $p$-ádicas. En general dichas $L$-series convergen sólo en un semiplano como sucede con la función zeta de Riemann. Un problema muy difícil es estudiar si dicha función se puede extender a todo el plano complejo, y otro problema muy interesante es poder calcular valores en puntos *especiales*, donde se espera que el valor de $L(\rho, s_0)$ contenga información del objeto geométrico/aritmético/algebraico a partir del cual se construyó $\rho$ (para el lector interesado, ver el Teorema de Dedekind, y la conjetura de Birch y Swinnerton-Dyer).

En lo que respecta a las representaciones de Galois complejas, tenemos la siguiente conjetura.

**Conjetura 6.18** (Artin)**.** *Si $\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_{\mathbb{C}}(V)$ es una representación irreducible y no trivial, entonces $\rho$ se extiende de manera holomorfa a todo el plano complejo.*

Al día de hoy no se conoce ninguna demostración de dicha conjetura. Brauer (en 1946) usando teoría de caracteres inducidos, logró demostrar que la $L$-serie se puede extender de manera meromorfa a todo el plano complejo, pero a priori podría tener polos. Los únicos casos que se conocen de la conjetura es cuando $V$ tiene dimensión 2, y la representación es "impar" (esto quiere decir que el determinante de la imagen de conjugación compleja es $-1$), como corolario de la demostración de las conjeturas de Serre dado por Khare y Wintenberger.

En el próximo capítulo veremos cómo construir algunas representaciones de Galois $p$-ádicas. Dichas construcciones provienen en general de variedades algebraicas proyectivas (veremos el caso de curvas elípticas). La conjetura de Fontaine-Mazur predice que todas las representaciones $p$-ádicas que ramifiquen en un conjunto finito y satisfagan una cierta condición técnica, se obtienen a partir de variedades algebraicas. Al día de hoy se conocen muy pocos casos de dicha conjetura.

## 7. Curvas algebraicas

En el presente capítulo vamos a estudiar la *aritmética de las curvas elípticas*. La palabra *aritmética* está directamente asociada a los números enteros y sus propiedades, mientras que una *curva* es un objeto geométrico de dimensión 1, que en nuestro caso estará contenida en un plano. También podríamos decir que lo que vamos a hacer es *teoría de números de ciertas curvas*.

Las curvas que consideraremos serán *curvas algebraicas planas*, y aunque no vamos a dar la definición formal de las mismas, momentáneamente llamaremos así a aquellas curvas $C$ descritas por una ecuación polinómica en dos variables $F(x, y) = 0$, donde los coeficientes del polinomio $F$ serán números de algún cuerpo, como bien podrían ser $\mathbb{R}$ o $\mathbb{C}$, pero como vamos a hacer aritmética nos interesará sobre todo que estén en $\mathbb{Q}$.

Aunque inevitablemente al esbozar la gráfica de una curva tengamos que "dibujar" *todos* sus puntos, es decir, tanto los racionales como aquellos con alguna o ambas coordenadas irracionales, momentáneamente estos últimos no existirán para nuestros fines (serán como si no estuvieran).

También tendremos que considerar para una curva definida sobre los racionales sus puntos con coordenadas que sean enteros algebraicos, al menos, algunos de tales puntos.

**7.1. Cónicas.** Podríamos comenzar considerando *rectas*, en vista de que son el tipo de curvas más simple, pero las dejaremos de lado y pasaremos directamente a las *cónicas*, que son las curvas descritas por un polinomio en dos variables y de segundo grado. Implícitamente estamos midiendo el *nivel de complejidad* de una curva por el grado del polinomio $F$ cuya ecuación polinómica $F(x, y) = 0$ la define. Así, las rectas están descritas por polinomios de primer grado y las cónicas (circunferencia, elipse, parábola e hipérbola) por polinomios de segundo grado. En un nivel de complejidad posterior al de las cónicas se ubican las *curvas elípticas* y en el siguiente y último escalón están las *curvas de mayor complejidad*, que ni siquiera tienen un nombre específico. En realidad, para ser un poco más precisos, las singularidades (puntos dobles por ejemplo) también deben tenerse en cuenta para medir esta complejidad, pero eso ya lo veremos más adelante.

Detengámonos entonces en las cónicas, que en relación a nuestras necesidades, se comportan todas de igual modo, la principal diferencia será que, sobre los racionales,

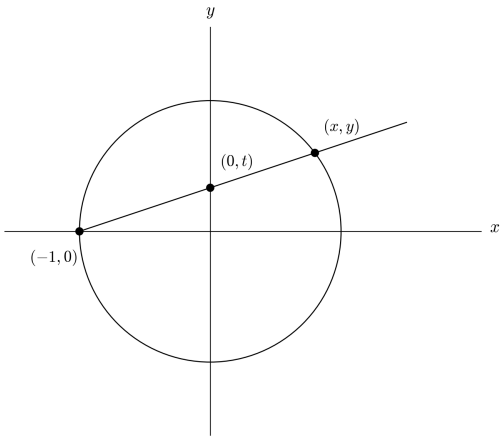



FIGURA 4. Cónica

algunas tienen puntos y otras no. Apoyaremos esta última afirmación viendo cómo se pueden poner de manifiesto los puntos racionales de una cónica, y llegaremos a la conclusión que en todas, éstos se "manifiestan" igual.

Sea, por ejemplo, la cónica $C$ descrita por la siguiente ecuación:

$$C: \qquad x^2 + y^2 = -1. \tag{7.1}$$

Trivialmente, en este caso es $C(\mathbb{Q}) = \emptyset$. Pero veremos a continuación que si hay algún punto racional en una cónica, entonces hay infinitos. Para ello consideremos la circunferencia $C$ con centro en el origen de coordenadas y radio 1, junto con uno de sus puntos racionales, como por ejemplo $P = (-1, 0)$. Si elegimos sobre el eje $y$ un punto racional $T = (0, t)$, con $t \in \mathbb{Q}$ y trazamos la recta $\ell$ que pasa por $P$ y $T$, veremos que la misma corta a $C$ en un punto $R$ racional (ver Figura 4).

En efecto, la ecuación de $\ell$ es $y = tx + t$ y su intersección con

$$C: \qquad x^2 + y^2 = 1 \tag{7.2}$$

es el punto $R = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)$, que claramente es racional. Recíprocamente, dado un punto racional $R$ de $C$, la recta que pasa por él y $P$ corta al eje $y$ en un punto $T$ racional. De modo que a partir de $P$ podemos *generar todos* los puntos racionales de $C$ considerando todas las rectas de pendiente racional que pasan por $P$:

$$C(\mathbb{Q}) = \left\{ \left( \frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2} \right) : t \in \mathbb{Q} \right\} \cup \{P\}. \tag{7.3}$$

Por otro lado, digamos que siempre una recta corta a una cónica en dos puntos (reales o complejos), contados con su multiplicidad (más generalmente, un teorema de la Geometría Proyectiva debido a E. Bézout afirma que: una curva de grado $m$ y una curva de grado $n$ siempre se cortan en $mn$ puntos). Esto último nos permite interpretar que:

- una *recta tangente* a una circunferencia, por ejemplo, también corta a ésta en dos puntos, o como diremos ahora, en un punto *doble* (o *de multiplicidad* 2); y que

- una *recta exterior* a una cónica también corta a ésta en dos puntos, aunque imaginarios, pues sus coordenadas serán números complejos.

Observemos ahora que el problema *geométrico* resuelto recién al hallar todos los puntos del conjunto $C(\mathbb{Q})$ en el caso de la circunferencia unitaria también nos da la solución de un problema *aritmético* de vieja data, el de obtener todas las ternas de números enteros tales que la suma de los cuadrados de los dos primeros sea igual al cuadrado del tercero. Tales ternas se denominan ternas Pitagóricas, dada su relación evidente con las longitudes de los lados de los triángulos rectángulos (en los cuales se verifica el teorema de Pitágoras). En efecto, el problema de hallar todas las ternas $(x, y, z)$ de números enteros tales que $x^2 + y^2 = z^2$ puede considerarse resuelto por el problema anterior dado que la precedente ecuación homogénea (en tres variables y de grado 2) puede deshomogeneizarse respecto de $z$ (esto es: podemos en ella pasar dividiendo $z^2$ al primer miembro si desechamos la solución trivial $x = y = z = 0$) para obtener la ecuación

$$\left(\frac{x}{z}\right)^2 + \left(\frac{y}{z}\right)^2 = 1, \tag{7.4}$$

la cual, en vista de los elementos de $C(\mathbb{Q})$ dados por (7.3), tiene las soluciones: $\frac{x}{z} = \frac{1-t^2}{1+t^2}$ y $\frac{y}{z} = \frac{2t}{1+t^2}$, con $t \in \mathbb{Q}$, lo que nos termina por dar (suponiendo $t = \frac{r}{s}, 0 \leq t \leq 1$):

$$x = s^2 - r^2, \qquad y = 2rs, \qquad z = s^2 + r^2, \qquad \text{con } r, s \in \mathbb{Z}, \tag{7.5}$$

como familia de ternas Pitagóricas (primitivas si $r$ y $s$ son primos relativos y de distinta paridad).

El procedimiento que acabamos de describir para la circunferencia unitaria puede extenderse al resto de las cónicas: partiendo de un punto racional en cualquiera de ellas (si lo hubiera) y trazando por él todas las rectas con pendiente racional, los puntos en que éstas cortan a la cónica serán todos los puntos racionales de la misma.

Como vimos en el caso de la cónica descrita por la ecuación $x^2 + y^2 = -1$, a veces una curva algebraica no tiene puntos racionales. Un caso menos evidente es el de la siguiente cónica:

$$C: \qquad x^2 + y^2 = 3, \tag{7.6}$$

para la cual ya no es obvio que $C(\mathbb{Q}) = \emptyset$. Veamos que sin embargo es así: expresando los posibles puntos racionales de $C$ como cociente de enteros y homogeneizando vía común denominador obtenemos que $C(\mathbb{Q}) = \emptyset$ si y sólo si la ecuación

$$x^2 + y^2 = 3z^2 \tag{7.7}$$

no tiene soluciones en números enteros (distintas de la trivial $(0, 0, 0)$).

Podemos ver esto último, que es un resultado negativo, encontrando un número primo $p$ para el cual la congruencia

$$x^2 + y^2 \equiv 3z^2 \qquad (\text{mód } p) \tag{7.8}$$

no tenga solución (dado que, evidentemente, una congruencia es más débil que una igualdad: dos enteros pueden ser congruentes sin ser iguales, pero no al revés). Este enfoque, que a priori podría parecer arbitrario, sorpresivamente funciona siempre para polinomios homogéneos de grado 2. En general tenemos el siguiente resultado (que generaliza un teorema de Legendre): si $F(x_1, \ldots, x_n) = 0$ no tiene solución en

números enteros (donde $F$ es un polinomio homogéneo de grado 2) entonces o no hay soluciones reales (algo fácil de verificar), o existe al menos un primo $p$ y un natural $r$ (generalmente igual a 1) para los cuales la congruencia

$$F(x_1, \ldots, x_n) \equiv \quad (\text{mód } p^r) \tag{7.9}$$

tampoco tiene solución. La recíproca es trivial (y vale para cualquier función $F$), pero la veracidad de la afirmación precedente es una "grata sorpresa". Decimos esto último porque hay igualdades (de grado superior al 2) que no tienen solución y sin embargo todas sus correspondientes congruencias (módulo todos los primos) sí la tienen. Un ejemplo concreto de esta última afirmación fue encontrado en 1951 por E. Selmer (y publicado en la referencia [17], pero realmente se "terminó de entender" en la década de 1980) y lo constituye la siguiente curva algebraica:

$$A: \qquad 3x^3 + 4y^3 + 5z^3 = 0. \tag{7.10}$$

La misma no tiene puntos racionales distintos del $(0, 0, 0)$ a pesar de que para todo primo $p$ hay puntos módulo $p$ en ella.

Retomando nuestro problema de ver que la forma cuadrática $x^2 + y^2 - 3z^2$ no representa a cero para ninguna terna de racionales $(x, y, z) \neq (0, 0, 0)$ utilizando la estrategia de hallar un primo $p$ (adecuado) para el cual la congruencia $x^2+y^2-3z^2 \equiv 0$ no tenga solución módulo $p$, debemos empezar por buscar dicho primo. Así por ejemplo, si elegimos arbitrariamente $p = 7$ veremos que la congruencia se cumple (existen, por ejemplo, $x = 1, y = 2$ y $z = 2$ tales que $x^2 + y^2 - 3z^2 = 1 + 4 - 12 = -7 \equiv 0 \pmod 7$). De modo que $p = 7$ no es adecuado para nuestro propósito. Sí lo será $p = 3$.

Empecemos observando que si $a \not\equiv 0 \pmod 3$ entonces $a^2 \equiv 1 \pmod 3$. De modo que si $x \not\equiv 0$ e $y \not\equiv 0 \pmod 3$ entonces $x^2 + y^2 \equiv 2 \pmod 3$, y como para todo $z$ es $3z^2 \equiv 0 \pmod 3$ resulta que $x^2 + y^2 - 3z^2 \not\equiv 0 \pmod 3$.

Resta ver el caso $x \equiv 0$ ó $y \equiv 0 \pmod 3$. Pero (siempre trabajando módulo 3): $x \equiv 0 \Rightarrow x^2 \equiv 0$, lo cual, junto con el hecho de que $3z^2 \equiv 0$ nos da que $x^2 - 3z^2 \equiv 0$, es decir, que $y^2 \equiv 0 \pmod 3$), lo cual a su vez implica que $y \equiv 0 \pmod 3$. Hemos probado que en nuestra ecuación: $x \equiv 0 \pmod 3 \Leftrightarrow y \equiv 0 \pmod 3$. Luego, en el caso que nos resta ver, sólo pueden ser simultáneamente $x \equiv 0$ e $y \equiv 0 \pmod 3$. Además, bajo estas condiciones también $z$ debe ser múltiplo de 3, porque si no lo fuera tendríamos la contradicción de que $x^2$ e $y^2$ serían múltiplos de 9 (y por ende también su suma), sin serlo su "igual" $3z^2$. Por otro lado, como el polinomio es homogéneo, si admite una solución no trivial, también admite una solución primitiva, esto es, una en la cual $\text{mcd}\{x, y, z\} = 1$. Y además, como acabamos de explicar, en una tal solución primitiva no pueden ser $x$ e $y$ múltiplos de 3 a la vez (porque eso también forzaría a que lo sea $z$, dejando de ser, entonces, primitiva). Y si uno no lo es, ninguno puede serlo (ya visto), pero en tal caso (también visto) la congruencia no tiene solución.

Lo que hemos dado en llamar una "grata sorpresa" es un profundo teorema, extremadamente fuerte, debido a Hasse-Minkowski (que denotaremos por *teorema de H-M* en adelante. El mismo está demostrado en la referencia [18], IV, Theorem 8), el cual afirma que $C(\mathbb{Q}) \neq \emptyset$ si y sólo si se verifican estas dos condiciones:

- para todo primo $p$ hay puntos módulo $p$ en la cónica $C$, y
- hay al menos un punto real (esto es: de coordenadas reales) en $C$.

Se suele decir que la segunda condición es la primera aplicada al *primo al infinito*.

Hay otra forma de enunciar el teorema de H-M en términos de los cuerpos $p$-ádicos $\mathbb{Q}_p$, diciendo que $C(\mathbb{Q}) \neq \emptyset$ si y sólo si se verifican estas dos condiciones:

- para todo primo $p$ hay puntos de $\mathbb{Q}_p$ en la cónica $C$, y
- hay al menos un punto real (esto es: de coordenadas reales) en $C$.

También se suele expresar la segunda condición en lenguaje análogo a la primera, diciendo que debe haber puntos de $\mathbb{Q}_\infty$ en la curva. El teorema de H-M es un ejemplo de aplicación de lo que suele llamarse *"principio global* versus *local"*.

*Ejercicio* 7.1. Mostrar que para cada número natural $e$ existe un entero $x$ solución de la congruencia $x^2 + 1 \equiv 0 \pmod{5^e}$.

Digamos también, y sólo a título informativo, que la "lista" (o sucesión) de las infinitas soluciones de la congruencia del Ejercicio 7.1 (una para cada $e \in \mathbb{N}$) es un número 5-ádico, esto es, es un elemento del cuerpo $\mathbb{Q}_5$.

El teorema de H-M es válido para cónicas (y así lo hemos enunciado), pero deja de serlo para curvas más complejas (como la del ejemplo encontrado por Selmer).

**7.2. Curvas Elípticas.** Como dijimos anteriormente, las cónicas son curvas que poseen un primer nivel de dificultad (en lo referido a la forma de hallar sus puntos racionales). Hay una definición topológica del nivel de complejidad de una curva, que es el llamado *género*. Así, las cónicas (como las rectas) tienen género 0, mientras que las curvas elípticas poseen el siguiente nivel, o sea género 1. De modo que las curvas elípticas estarán definidas por ecuaciones polinómicas de grado 3. Pero ¡cuidado!, porque no toda ecuación de grado 3 define una verdadera curva elíptica.

Con mayor precisión, llamaremos curva elíptica a toda aquella curva plana $E$ cuyos puntos $(x, y)$ verifiquen una ecuación de la forma

$$E : y^2 = f(x) = x^3 + ax^2 + bx + c, \tag{7.11}$$

en la cual $f$ es un polinomio con coeficientes enteros, de grado 3 y sin raíces múltiples (o, como también se suele decir: sin *singularidades*). Simbolizaremos $E(\mathbb{Q})$ al conjunto de los puntos racionales de una curva elíptica $E$.

No es elíptica, por ejemplo, la curva definida por la ecuación $y^2 = x^2(x+1)$, pues a pesar de ser cúbico, el polinomio $f(x) = x^2(x+1)$ tiene una *raíz doble* en $x = 0$. Este hecho (el de poseer $f$ raíces múltiples) hace que la curva se parezca más a una cónica que a una curva elíptica (con esto último queremos decir que su género es 0, como en las cónicas, en lugar de 1 como en las verdaderas curvas elípticas).

A continuación veremos una forma "analítica" para determinar singularidades. Consideremos una curva $C$ definida por una ecuación de la forma $y^2 = f(x)$. Si escribimos la misma como $F(x, y) = y^2 - f(x) = 0$ y calculamos sus derivadas parciales,

$$\frac{\partial F}{\partial x} = -f'(x), \qquad \frac{\partial F}{\partial y} = 2y \tag{7.12}$$

entonces por definición la curva es no singular si y sólo si no existen puntos sobre la misma que anulen simultáneamente ambas derivadas parciales. Geométricamente, eso significará que todo punto sobre la curva tiene una recta tangente bien definida.

- Si las derivadas parciales se anulan simultáneamente en un punto $(x_0, y_0)$ de la curva entonces $y_0 = 0$ y $f'(x_0) = 0$, como asimismo $f(x_0) = 0$ (pues

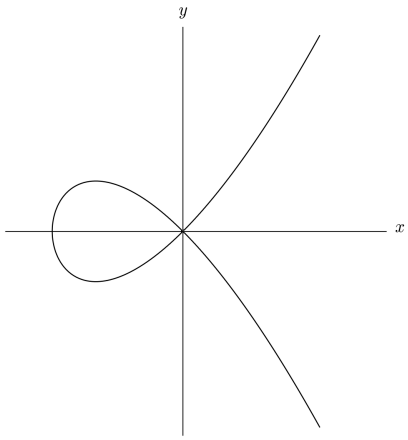



FIGURA 5. Cúbica Singular

$0 = y_0^2 = f(x_0)$). Ahora bien: el hecho de que $f(x_0) = f'(x_0) = 0$ implica que $x_0$ es una raíz doble de $f(x)$.

- Recíprocamente, si $f$ tuviera una raíz doble en $x_0$, entonces el punto $(x_0, 0)$ sería un punto singular de la curva (pues su multiplicidad 2 implica que $f(x_0) = f'(x_0) = 0$).

En resumen, una curva $C$ definida por una ecuación polinómica $F(x, y) = 0$ posee una singularidad en $(x_0, y_0)$ si y sólo si:

- $F(x_0, y_0) = 0$ (esto es: $(x_0, y_0) \in C$), y
- $\frac{\partial F}{\partial x}(x_0, y_0) = \frac{\partial F}{\partial y}(x_0, y_0) = 0$.

En el caso de la curva precedente tenemos que $F(x, y) = y^2 - x^2(x + 1)$ y un punto singular de la misma es $(x_0, y_0) = (0, 0)$. Procediendo como en las cónicas, a partir de éste (que es racional) también podemos hallar infinitos puntos racionales de la curva (que en este caso los habrá) haciendo pasar por $(0, 0)$ rectas con pendiente racional: aquellos puntos donde éstas corten a la curva serán puntos racionales. De modo que ahora la condición "nueva" es que la recta debe pasar por el punto singular (que asumiremos que será racional). Y vemos que el procedimiento es muy similar al ya descrito para las cónicas (y habíamos ejemplificado con cierto detalle en la circunferencia unitaria). Ver la Figura 5.

En el caso que ahora nos ocupa, tales rectas tienen ecuación $y = rx$, con $r \in \mathbb{Q}$, y si buscamos sus puntos de intersección con $C$ obtendremos los puntos $(x, y)$ con $x = r^2 - 1$ e $y = r\left(r^2 - 1\right)$, que evidentemente son racionales.

De modo que las dos "patologías" que pueden hacer que una ecuación de la forma $y^2 = f(x)$, con $f$ polinomio a coeficientes enteros de tercer grado, no sea una curva elíptica son:

- que $f$ posea dos raíces iguales, o
- que $f$ posea sus tres raíces iguales.

*Ejercicio* 7.2. Consideremos la ecuación $y^2 = x^3$.

1. Dibujar su conjunto de puntos reales en el plano,
2. Indicar si dicha curva es singular (esto es: si posee un punto singular),

3. Indicar si la curva posee puntos racionales, y en caso afirmativo, calcularlos.

Para que una ecuación polinómica en dos variables $F(x,y)=0$ represente a una curva elíptica, también es suficiente que exista un cambio de variables (adecuado) que permita llevarla a la llamada *forma normal de Weierstrass* (o simplemente *forma normalizada* o *forma tipo*) como en (7.11) donde $f$ es un polinomio cúbico con coeficientes enteros y sus tres raíces distintas.

Otra manera de concluir lo mismo es verificando que $F(x,y)=0$ es una ecuación cúbica no singular, con el agregado (y este es un requisito *sine qua non*) de que sea $C(\mathbb{Q}) \neq \emptyset$.

Dada una cúbica $F(x,y)=0$ no singular con un punto racional, se puede ver (usando por ejemplo el Teorema de Riemann-Roch) que existe un función que manda la cúbica en una ecuación normalizada como en (7.11), con lo cual nos restringiremos a trabajar únicamente con dicho tipo de ecuaciones.

Para poder unificar el estudio de curvas, es preciso salir de la geometría clásica y trabajar con la llamada *Geometría Proyectiva*. En ella una curva consta no sólo de sus *puntos afines*, sino también de sus *puntos proyectivos*. Así por ejemplo el *plano proyectivo* se obtiene agregando al plano afín un *punto al infinito* por cada dirección. Luego en el plano proyectivo todas las rectas se cortan: las que no en un punto afín (antes llamadas *paralelas*), sí en el punto al infinito (correspondiente a la dirección de las mismas). Para los puntos al infinito interesa la dirección pero no el sentido: yendo "para arriba" o "para abajo" por una misma recta, nos encontramos con el mismo punto al infinito, que en adelante simbolizaremos $\mathcal{O}$. En las curvas elípticas normalizadas, se dice que las rectas verticales son rectas "que pasan por $\mathcal{O}$". Además $\mathcal{O}$ tiene coordenadas ¡proyectivas!, y es racional: $\mathcal{O} \in E(\mathbb{Q})$.

La ventaja de trabajar con el plano proyectivo es que en él, toda recta corta a una curva elíptica $E$ en tres puntos (contados con multiplicidad).

Así como en las cónicas y en las cúbicas con puntos singulares, en las curvas elípticas también existen *métodos propagadores* que permiten, a partir de uno o más de sus puntos racionales conocidos, generar otros. Lo importante es, siempre, encontrar el o los puntos de partida y saber que los que generaremos por estos métodos pueden no ser todos. Dado que una recta corta a una curva elíptica en tres puntos, necesitaremos partir de dos de sus puntos racionales (dos puntos de $E(\mathbb{Q})$) para ir generando otro u otros puntos racionales de las mismas. Esto dará lugar a una "operación", a dos puntos de la curva le podremos asociar un tercero, que le otorgará al conjunto $E(\mathbb{Q})$ la estructura, antes anunciada, de grupo Abeliano.

Dados los dos puntos (racionales) de partida sobre $E$, que llamaremos $P$ y $Q$, comenzaremos por definir una operación $*$ (que leeremos *estrella*) entre ellos cuyo resultado se obtiene trazando la recta pasante por $P$ y $Q$ e intersectando a $E$ : dicho punto de intersección $R^*$ será el resultado de $P*Q$. Si ésta fuera la "operación" en $E(\mathbb{Q})$, no habría estructura de grupo (la estructura todavía está ausente). A continuación "unimos" $R^*$ con $\mathcal{O}$ trazando por $R^*$ una recta vertical (y en general: paralela a la asíntota) y el punto de intersección de esta última con $E$, punto que llamaremos $R$, será el resultado de la "verdadera" operación entre $P$ y $Q$, operación que denominaremos *suma* y le otorga a $E(\mathbb{Q})$ estructura de grupo Abeliano. En síntesis: $P+Q=(P*Q)*\mathcal{O}=R$ (ver Figura 6). La conmutatividad de la operación suma entre $P$ y $Q$ es evidente, no así su asociatividad.

*Ejercicio* 7.3. Probar que $\mathcal{O}$ es el neutro para la suma en la curva $E$.

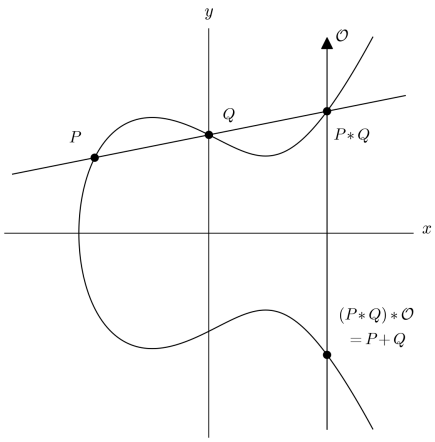



FIGURA 6. Suma en Curvas Elípticas

Supongamos ahora que una curva elíptica $E$ está dada por la ecuación normalizada (7.11), y que los puntos racionales de partida sobre $E$ sean los puntos $P$ y $Q$, de coordenadas (afines) $P = (x_1, y_1), Q = (x_2, y_2)$. La recta que pasa por ellos tiene por ecuación $y = \lambda x + \nu$, donde $\lambda = \frac{y_2 - y_1}{x_2 - x_1}$ y $\nu = y_1 - \lambda x_1$ (desde luego, el caso $x_2 = x_1$ hay que hacerlo aparte. Es un caso más fácil y lo consideraremos más adelante). Reemplazando en la ecuación de $E$, con el fin de hallar $R^* = (x_3, y_3)$, resulta la igualdad

$$(\lambda x + \nu)^2 = x^3 + ax^2 + bx + c, \tag{7.13}$$

y la abscisa $x_3$ de $R^*$ es la "tercera" raíz (diferente de $x_1$ y $x_2$) del polinomio mónico de tercer grado

$$x^3 + \left(a - \lambda^2\right) x^2 + (b - 2\lambda\nu)\, x + \left(c - \nu^2\right) = 0. \tag{7.14}$$

Si recordamos la relación que existe entre las raíces de un polinomio y los coeficientes del mismo, observaremos que:

$$x_1 + x_2 + x_3 = \lambda^2 - a, \tag{7.15}$$

$$x_1x_2 + x_1x_3 + x_2x_3 = b - 2\lambda\nu, \tag{7.16}$$

$$x_1x_2x_3 = \nu^2 - c, \tag{7.17}$$

planteado lo cual, podemos hallar sus raíces resolviendo un sistema de ecuaciones. En este caso resultará:

$$x_1 + x_2 + x_3 = \lambda^2 - a \quad \Rightarrow \quad x_3 = \lambda^2 - a - x_1 - x_2. \tag{7.18}$$

Una vez obtenida $x_3$ hallamos la ordenada $y_3$ de $R^*$ vía la ecuación de la recta: $y_3 = \lambda x_3 + \nu$. Finalmente, como $R\,(x_r, y_r)$ es simétrico a $R^*$ respecto del eje de las abscisas, sus coordenadas serán $x_r = x_3$ e $y_r = -y_3$. Denominaremos a la expresión (7.18) la *fórmula de la adición*.

A los efectos de ver el caso pendiente ($x_1 = x_2$), correspondiente a la fórmula para $x_3$ cuando sumamos un punto $S$ con sí mismo, suma que abreviaremos $2S := S + S$ conviene proceder como a continuación indicaremos (pues la expresión anterior de la pendiente $\lambda$ exigía que los puntos tuvieran abscisas distintas).

Dado $S = (x_0, y_0) \in E$, queremos encontrar las coordenadas de $2S$ a las que llamaremos $(x_3, y_3)$. Primero necesitamos encontrar la ecuación de la recta que une $S$ con $S$, es decir, de la recta tangente a la cúbica en $S$. Para ello, a partir de la relación

$$F(x,y) = y^2 - f(x) = y^2 - x^3 - ax^2 - bx - c = 0, \tag{7.19}$$

encontramos, derivando en forma implícita, que $\lambda = \frac{dy}{dx} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} = -\frac{-f'(x)}{2y} = \frac{f'(x)}{2y}$, y valorizando en $S$ obtenemos dicha pendiente:

$$\lambda = \frac{f'(x_0)}{2y_0}, \tag{7.20}$$

y la ecuación de la recta tangente será entonces $y = \lambda x + \nu$ con $\nu = y_0 - \lambda x_0$. A los efectos de obtener una fórmula explícita para el valor de las coordenadas de $2S$ en términos de las coordenadas de $S$, sustituimos la expresión (7.20) de $\lambda$ en la expresión (7.15). Sacando común denominador y reemplazando $y_0^2$ por $f(x_0)$ encontraremos que:

$$x_3 = \frac{x_0^4 - 2bx_0^2 - 8cx_0 + \left(b^2 - 4ac\right)}{4x_0^3 + 4ax_0^2 + 4bx_0 + 4c}, \tag{7.21}$$

obteniéndose su ordenada $y_3$ en consecuencia. La expresión (7.21) es la llamada *fórmula de duplicación.*

Lo que nos interesa destacar de todas estas expresiones, es que las coordenadas del punto "suma" de otros dos, son *funciones racionales* (esto es: cociente de polinomios, y ¡a coeficientes enteros!) en las coordenadas de los sumandos, de modo que si $P, Q \in E(\mathbb{Q})$ entonces $P + Q \in E(\mathbb{Q})$.

**Teorema 7.4** (Mordell)**.** *Sea $E$ una curva elíptica sobre $\mathbb{Q}$, y $E(\mathbb{Q})$ el grupo Abeliano de sus puntos racionales. Entonces $E(\mathbb{Q})$ es finitamente generado.*

Recordar que cuando un grupo Abeliano está finitamente generado su estructura se puede "controlar", pues al existir un número finito de *generadores* (o "puntos privilegiados"), conociéndolos (esto es: conociendo una cantidad "finita" de información referida al grupo), podemos calcular ¡todo! Desde ya que lo anterior no quiere decir que vaya a haber un número finito de puntos racionales en la curva, como tampoco lo contrario.

Los resultados del siguiente ejercicio nos permitirán exhibir dos pares de números (ciertamente "no triviales") tales que la diferencia entre el cubo de uno de ellos y el cuadrado del otro es igual a 17.

*Ejercicio* 7.5. Dada la curva elíptica definida por la ecuación normalizada $y^2 = x^3 + 17$ y sabiendo que $P_1 = (-1, 4)$ y $P_2 = (2, 5)$ son dos de sus puntos racionales, calcular (por el procedimiento descrito de las tangentes y secantes):

1. $P_1 + P_2$,
2. $2P_1$

**7.3. Puntos de Torsión.** Consideremos la ecuación cúbica $x^3 + y^3 = 1$, que como afirmó P. de Fermat y demostró L. Euler sólo tiene un número finito de puntos racionales, que corresponden a sus dos soluciones "triviales" $(x = 0, y = 1)$ y $(x = 1, y = 0)$ y al punto al infinito $\mathcal{O}$ (que también es racional por ser racional la pendiente de la asíntota, recta cuya dirección define a $\mathcal{O}$). De modo que estos tres son los únicos puntos racionales de la curva definida por $x^3 + y^3 = 1$ (ver Figura 7).

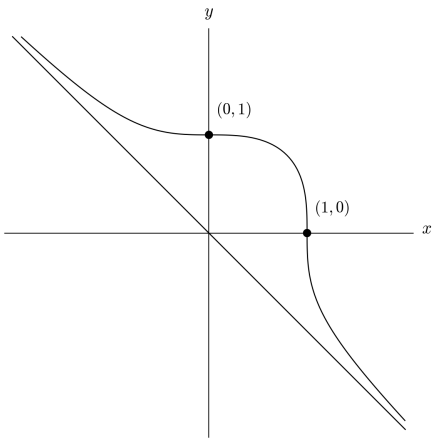



FIGURA 7. Cúbica Fermat

Puede verse que $x^3+y^3=1$ representa una curva elíptica $E$. En efecto: el cambio de variables que nos permite expresarla en la forma normal es el siguiente:

$$x=\frac{36+Y}{6X}, \qquad y=\frac{36-Y}{6X}. \tag{7.22}$$

Luego de llevarlo a cabo (¡hacerlo!) obtendremos $Y^2=X^3-432$.

Pero a pesar de la existencia de tal cambio de variable (o cambio de coordenadas, ¡y racional!) que la lleva a la forma normalizada de Weierstrass, la dejaremos simplemente expresada así, como $x^3+y^3=1$ (digamos que "en honor a su fama").

El hecho de que $P=(0,1), Q=(1,0)$ y $\mathcal{O}$ sean sus tres únicos puntos racionales (debido al resultado de L. Euler), hace que si operamos con ellos como aprendimos (esto es, si calculamos $P+Q, 2P, 2Q$, etc.) el resultado siempre será uno de ellos tres y consecuentemente estará en el conjunto:

$$E(\mathbb{Q})=\{P,Q,\mathcal{O}\}. \tag{7.23}$$

También hemos dicho que toda recta corta a una curva elíptica en tres puntos y pareciera que en este caso, una recta tangente a $E$ por $P$ cortará a la curva en sólo dos (porque los puntos de tangencia de una recta, son dobles). Pero en esta curva se da el caso que $P$ y $Q$ son, además, *puntos de inflexión* de la misma, lo que hace que en realidad sean puntos *triples*.

Sabemos que $E(\mathbb{Q})$ tiene estructura de grupo Abeliano, de modo que en este caso $E(\mathbb{Q})$ es el (único) grupo de orden tres, el cual es cíclico y por lo tanto Abeliano. Concretamente

$$2P=Q, \qquad 2Q=P, \qquad 3P=\mathcal{O}, \qquad 3Q=\mathcal{O}. \tag{7.24}$$

De modo que $P$ y $Q$ (y desde ya, también $\mathcal{O}$) son los puntos de la curva que tienen *orden finito*. Tales puntos son raros (porque no es frecuente que un punto sumado una cantidad finita de veces consigo mismo tenga por resultado él mismo) y se denominan *puntos de torsión* de la curva elíptica. Como en este caso $P$ y $Q$ tienen orden tres, esto se expresa diciendo que son puntos de 3-torsión de $E$.

Ampliemos momentáneamente el campo de variación de los puntos de una curva permitiendo que tengan coordenadas reales e incluso complejas. Y lo interesante es

que los tres conjuntos de puntos $E(\mathbb{Q}), E(\mathbb{R})$ y $E(\mathbb{C})$ tienen estructura de grupo Abeliano con la misma operación "suma de puntos". De modo que ahora nos van a interesar, por ejemplo, no sólo los puntos racionales de orden 3 de $E$, sino *todos* los puntos de $E$ de orden 3 (con coordenadas reales o complejas). Para nuestros fines, considerar el conjunto más amplio $\mathbb{C}$, tendrá una ventaja doble: la de tratar con un conjunto que es a la vez *algebraicamente cerrado* y *topológicamente completo* (recordemos que $\mathbb{Q}$ no posee ninguna de ambas propiedades, y que $\mathbb{R}$ reúne sólo la segunda).

Ante todo vamos a hacer una observación respecto del lenguaje que adoptaremos. Es frecuente encontrar en algunos libros la definición de *orden de un elemento* $P$ de un grupo como el *mínimo* natural $m$ tal que

$$\underbrace{P+\ldots+P}_{m \text{ veces}} = mP = \mathcal{O}. \tag{7.25}$$

Los *puntos de m-torsión* de una curva elíptica $E$ (con $m \in \mathbb{N}$) serán aquellos para los cuales $mP = \mathcal{O}$. La curva elíptica del ejemplo precedente ($x^3 + y^3 = 1$) tiene tres puntos de 3-torsión racionales ($P, Q$ y $\mathcal{O}$).

Analicemos propiedades que caracterizan a los puntos de 2-torsión de cualquier curva elíptica $E$, esto es, de aquellos puntos $P \in E(\mathbb{C})$ tales que $2P = \mathcal{O}$. Por lo pronto, de esto se deduce que $P = -P$. Luego: si un punto $P$ es de 2-torsión entonces está en el eje de simetría de la curva, que será el eje $x$ si la curva esta normalizada. Si simbolizamos con $E(\mathbb{C})[m]$ al conjunto de todos los puntos de $m$-torsión de $E$, resulta que el mismo también tiene estructura de grupo Abeliano. Concretamente, en el caso que estamos considerando, $E(\mathbb{C})[2]$ es un grupo Abeliano con cuatro elementos: $\mathcal{O}$ y los tres puntos con ordenada $y = 0$, que son $(\alpha_1, 0), (\alpha_2, 0)$ y $(\alpha_3, 0)$, donde $\alpha_1, \alpha_2$ y $\alpha_3$ son las raíces del polinomio $f(x)$. Pero no es el grupo *cíclico* de cuatro elementos, pues éstos deben ser de orden 2 (¡por ser puntos de 2-torsión!); sino que se trata del otro grupo de cuatro elementos, el isomorfo a la *suma directa* $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$, grupo cuyo conjunto subyacente, lo decimos rápidamente, es el producto cartesiano $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} = \{\overline{0}, \overline{1}\} \times \{\overline{0}, \overline{1}\} = \{(\overline{0}, \overline{0}), (\overline{0}, \overline{1}), (\overline{1}, \overline{0}), (\overline{1}, \overline{1})\}$ y cuya "suma" (la operación que le otorga estructura de grupo) está definida por componentes. En resumidas cuentas:

$$E(\mathbb{C})[2] \cong \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}, \tag{7.26}$$

(con $\cong$ indicamos *isomorfismo de grupos*).

Pasemos a continuación a analizar los puntos de 3-torsión de una curva elíptica $E$. Éstos tienen la propiedad de que $3P = \mathcal{O}$, o equivalentemente, que $2P = -P$. Recordando que la abscisa de $-P$ es igual que la abscisa de $P$, resulta que $2P$ tiene la misma abscisa que $P$ y por lo tanto está sobre la curva, en la misma vertical que $P$. Dicho brevemente: si $P$ es un punto de 3-torsión de $E$ entonces $2P$ es el otro punto de $E$ que está en la misma vertical que $P$.

Recordemos la fórmula de duplicación (7.21) que nos permite calcular la abscisa de la suma de un punto con sí mismo y apliquémosla a $2P$. Como en este caso ésta (la abscisa de $2P$) coincide con la de $P$, igualándolas llegaremos a que:

**Teorema 7.6.** *$P(x, y) \in E(\mathbb{C})[3]$ si y sólo si $P = \mathcal{O}$ o su abscisa es raíz del polinomio*

$$\psi_3(x) = 3x^4 + 4ax^3 + 6bx^2 + 12cx + \left(4ac - b^2\right). \tag{7.27}$$

Puede probarse que el polinomio $\psi_3(x)$ posee 4 raíces distintas (el hecho de que $\psi_3(x)$ no posee raíces múltiples no es un resultado trivial). Combinando estos cuatro valores de $x$ con los dos posibles valores de $y$ (el dado por la fórmula $y = \lambda x + \nu$ y su opuesto), tenemos los ocho puntos afines de 3-torsión de $E$, que junto con el punto al infinito $\mathcal{O}$ nos dan los 9 puntos de 3-torsión de $E$. De modo que $E(\mathbb{C})[3]$ es un grupo Abeliano de 9 elementos ($|E(\mathbb{C})[3]| = 9$) e isomorfo (dado que todos sus elementos son de orden 3) a la suma directa de los *enteros módulo* 3:

$$E(\mathbb{C})[3] \cong \mathbb{Z}/3\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z}. \tag{7.28}$$

(así que no es, por ejemplo, isomorfo al grupo cíclico de 9 elementos $\mathbb{Z}/9\mathbb{Z}$, etc.)

*Ejercicio* 7.7. Calcular todos los puntos de 3-torsión (racionales o no) de la curva elíptica definida por la ecuación $x^3 + y^3 = 1$.

El análisis que hemos llevado a cabo sobre los puntos de 2 y de 3-torsión de una curva elíptica $E$ nos induce a pensar que puede haber un resultado general al respecto, y efectivamente es así.

**Teorema 7.8.** *Sea $E$ una curva elíptica sobre $\mathbb{Q}$, y $m \in \mathbb{N}$. Entonces:*

1. $|E(\mathbb{C})[m]| = m^2$, *y*
2. $E(\mathbb{C})[m] \cong \mathbb{Z}/m\mathbb{Z} \oplus \mathbb{Z}/m\mathbb{Z}$.

El teorema precedente describe la estructura del grupo de los puntos de $m$-torsión (o de *$m$-división*, como también se los denomina) de una curva elíptica cuya ecuación (normalizada) tiene coeficientes racionales: es la suma directa de dos grupos cíclicos de orden $m$. En dicho grupo puede o no haber puntos racionales.

Recordar que el Teorema de Mordell (7.4) afirma que $E(\mathbb{Q})$ es finitamente generado. Pero sucede que en Teoría de Grupos existe un teorema que describe la estructura de cualquier grupo Abeliano finitamente generado, el cual informalmente afirma que: todo grupo Abeliano finitamente generado está generado por un número finito $r$ de elementos de orden infinito y por otro número finito $c$ de elementos de orden finito. Esto quiere decir que todo grupo Abeliano finitamente generado es isomorfo a $r$ copias de $\mathbb{Z}$ (que es un grupo libre de rango 1 con respecto a la suma: más precisamente su generador es el entero 1, pues sumando el número 1 con sí mismo cualquier cantidad de veces, siempre obtenemos un número distinto de $\mathbb{Z}$), más $c$ elementos de torsión (que en nuestro caso concreto son los $c$ elementos que contenga dicho grupo finito, que simbolizaremos $E_{tor}(\mathbb{Q})$). Simbólicamente, vale el siguiente isomorfismo:

$$E(\mathbb{Q}) \cong \underbrace{\mathbb{Z} \oplus \ldots \oplus \mathbb{Z}}_{r \text{ veces}} \oplus E_{tor}(\mathbb{Q}) = \mathbb{Z}^{\oplus r} \oplus E_{tor}(\mathbb{Q}) \tag{7.29}$$

Como acabamos de decir, con $E_{tor}(\mathbb{Q})$ estamos representando a los puntos racionales de torsión (que son "unos pocos", esto es: un número finito que sólo se generan entre ellos mismos). Por ejemplo, en el caso de la curva elíptica $x^3 + y^3 = 1$, de sus nueve puntos de torsión, sólo tres son racionales.

La expresión (7.29) nos está diciendo que $E(\mathbb{Q})$ tiene "dos partes": una libre y una de torsión, y "la parte de torsión, se puede calcular", pero "calcular la parte libre" es más difícil.

Digamos también que hay curvas elípticas en las cuales $r = 0$ y también las hay en las cuales $r > 0$ (y en principio, tan grande como se quiera). Si por ejemplo

fuera $r = 20$ (y ya se han encontrado curvas elípticas con este valor de $r$), nos expresaremos diciendo que la curva es *de rango* 20.

Lo bueno, para nosotros, es que la parte de torsión es bien conocida. Y puntos de $m$-torsión los hay para todo $m \in \mathbb{N}$. En relación con esto, Beppo Levi conjeturó (en el año 1908, ver [8]) que, aunque puntos de $m$-torsión puede haber para cualquier $m$, los que son racionales "no pueden ser tantos", "ni darse para cualquier valor de $m$". Este trabajo de B. Levi quedó en el olvido y más de 40 años después la misma propiedad fue re-conjeturada por T. Nagell [10], y aún posteriormente, en la década de 1970 era ampliamente conocida como conjetura de Ogg [11]. Finalmente, en el año 1976, B. Mazur [9] confirmó la *conjetura de B. Levi* (como en realidad mereció haberse llamado) demostrando el Teorema 7.9 que a continuación enunciaremos.

En relación con la *conjetura de B. Levi* , su historia y su alcance resulta muy instructiva la lectura del trabajo de Norbert Schappacher y René Schoof titulado *Beppo Levi and the Arithmetic of Elliptic Curves* [16][1]. En el mismo está claramente documentado que B. Levi había intuido lo mismo que otros sospecharon mucho después que él, sólo que a principios del siglo XX todo esto se expresaba en un lenguaje algo diferente del actual.

**Teorema 7.9** (Mazur)**.** *El conjunto $E_{tor}(\mathbb{Q})$ de los puntos racionales de torsión de una curva elíptica $E$ sólo puede ser isomorfo a los siguientes grupos finitos:*

1. $\mathbb{Z}/n\mathbb{Z}$ *si* $1 \leq n \leq 10$ *o* $n = 12$*, o bien a*
2. $\mathbb{Z}/n\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$ *si* $n = 2, 4, 6$ *u* $8$*.*

Pero volvamos ahora al Teorema 7.8 y al isomorfismo de su tesis:

$$E(\mathbb{C})[m] \cong \mathbb{Z}/m\mathbb{Z} \oplus \mathbb{Z}/m\mathbb{Z}. \tag{7.30}$$

Para tratar de "entenderlo" del todo y ver porqué siempre es cierto (o ver más claramente su significado geométrico, o ver una interpretación donde este isomorfismo sea "evidente") nos conviene mirar todos los puntos *complejos* de la curva elíptica $E$ y efectuar la "construcción" de $E(\mathbb{C})[m]$ desde el punto de vista del Análisis Complejo. Esto significa, en principio, que las variables $x$ e $y$ de la ecuación de $E$ pueden tomar libremente valores en $\mathbb{C}$.

La genuina y original ecuación de K. Weierstrass normalizada para una curva elíptica $E$ es:

$$E: \qquad y^2 = 4x^3 - g_2 x - g_3, \tag{7.31}$$

---

[1]En el mismo podremos leer pasajes como los siguientes, en los cuales refiriéndose a B. Levi, expresan los autores:

- *... his work on the arithmetic of elliptic curves has not received the attention it deserves. He occupied himself with this subject from 1906 to 1908. His investigations, although duly reported by him at the 1908 International Congress of Mathematicians in Rome, appear to be all but forgotten. This is striking because in this work Beppo Levi anticipated explicitly, by more than 60 years, a famous conjecture made again by Andrew P. Ogg in 1970, and proved by Barry Mazur in 1976.* (página 57);
- *More than 40 years later T. Nagell made the same conjecture* [10]. *In our days the conjecture became widely known as Ogg's conjecture, after Andrew Ogg, who formulated it 60 years after Beppo Levi. The problem studied by Beppo Levi and later by Billing, Mahler, Nagell and Ogg has been very important in the development of arithmetic algebraic geometry.* (página 65).

donde $g_2$ y $g_3$ serán, como veremos, dos funciones de "peso" 2 y 3 respectivamente (de ahí los símbolos asignados). A partir de esta $E$ se demuestra que existen dos números complejos $\omega_1$ y $\omega_2$, linealmente independientes sobre $\mathbb{R}$, tales que si consideramos el *retículo* (*lattice* en inglés) generado por el conjunto $\{\omega_1, \omega_2\}$, retículo que se simboliza $L[\omega_1, \omega_2] = \omega_1\mathbb{Z} + \omega_2\mathbb{Z}$ (y que se denomina *retículo asociado a* $E$), entonces las funciones

$$g_2 = 60 \sum_{\omega \in L, \omega \neq 0} \frac{1}{\omega^4} \qquad \text{y} \qquad g_3 = 140 \sum_{\omega \in L, \omega \neq 0} \frac{1}{\omega^6}$$

hacen que $x = \wp(u)$ e $y = \wp'(u)$, con $u \in \mathbb{C}$, verifiquen la ecuación de $E$, donde a su vez $\wp(u)$ (que se lee *p de u*, siendo $\wp$ la llamada *función* $\wp$ *de Weierstrass*) es la siguiente *función elíptica con respecto al retículo* $L$ (lo cual significa que es una función *doblemente periódica*, de períodos $\omega_1$ y $\omega_2$) :

$$\wp(u) = \frac{1}{u^2} + \sum_{\omega \in L, \omega \neq 0} \left[ \frac{1}{(u-\omega)^2} - \frac{1}{\omega^2} \right], \qquad u \in \mathbb{C}.$$

Dado un retículo $L[\omega_1, \omega_2]$ existe todo un "universo" de funciones doblemente periódicas de períodos $\omega_1$ y $\omega_2$, pero la función $\wp$ recién definida "controla" a todas ellas (¡es la principal!).

Que $\wp(u)$ sea doblemente periódica con respecto al retículo $L[\omega_1, \omega_2]$ significa que

$$\forall u \in \mathbb{C}, \omega \in L : \qquad \wp(u + \omega) = \wp(u). \tag{7.32}$$

La ecuación diferencial que verifica $\wp$ es, por lo tanto

$$(\wp')^2 = 4\wp^3 - g_2\wp - g_3 \tag{7.33}$$

y es ella la que nos dice que los puntos $P(u)$ de coordenadas $(\wp(u), \wp'(u))$ están en la curva definida por (7.31):

$$P(u) = (\wp(u), \wp'(u)) \in E. \tag{7.34}$$

La precedente, constituye una parametrización con parámetro $u \in \mathbb{C}$, de todos los puntos ¡complejos! de la curva elíptica $E$ (así como (7.3), con parámetro $t \in \mathbb{Q}$, era una parametrización de todos los puntos ¡racionales! de la circunferencia unitaria).

De modo que la correspondencia

$$\mathbb{C} \longrightarrow E(\mathbb{C}), u \mapsto P(u) = (\wp(u), \wp'(u)) \tag{7.35}$$

es un homomorfismo de grupos, que se transforma en isomorfismo si pasamos al cociente entre $\mathbb{C}$ y el retículo $L$ (y aquí utilizamos la doble periodicidad de $\wp$ respecto de $L$):

$$\mathbb{C}/L \longrightarrow E(\mathbb{C}), u \mapsto P(u) = (\wp(u), \wp'(u)). \tag{7.36}$$

El cociente $\mathbb{C}/L$ no es más que el *toro* (o más precisamente: isomorfo al toro, debido a la doble periodicidad de $\wp$), que es una superficie de Riemann de género 1 (porque tiene 1 "agujero"). La *esfera de Riemann* (sobre la cual hemos considerado las rectas y las cónicas) tiene género 0 (porque no tiene agujeros).

La relación de equivalencia en $\mathbb{C}$ que se considera al pasar al espacio cociente se basa en la doble periodicidad de la función $\wp$ de Weierstrass, y define a dos puntos $u_1, u_2 \in \mathbb{C}$ como *equivalentes* si y sólo si $\wp(u_1) = \wp(u_2)$.

Además, al ser el toro una superficie compacta, resulta que todas las funciones meromorfas en él definidas son *funciones racionales* y por lo tanto están en correspondencia con una *curva algebraica* (en el toro, el punto al infinito de $\mathbb{C}$ aparece como un punto más ¡y cualquiera! de su superficie).

Ahora bien: si calculamos $P(u_1+u_2)$ aplicando (7.34) (donde $u_1+u_2$ es la suma habitual de dos números complejos) obtendremos ¡oh coincidencia! la función $P(u_1)+P(u_2)$, donde ahora el signo $+$ se refiere a la suma de puntos sobre $E$ definida por el proceso geométrico descrito en la primera clase *de las tangentes y secantes*. Algo informalmente podríamos escribir que:

$$P(u_1+u_2)=P(u_1)+P(u_2). \tag{7.37}$$

Finalmente (y para llegar a esto hicimos la presentación de la función $\wp$) los $m^2$ puntos de $m$-división de $E$ (recordemos la tesis 1 del Teorema 7.8) se corresponden con los $m^2$ números complejos $z$ tales que

$$mz\equiv 0 \pmod{L}, \tag{7.38}$$

esto es, con los $z\in\mathbb{C}$ tales que sumados con sí mismo $m$ veces "caen" en el mismo lugar "relativo" dentro del retículo $L$. Una imagen geométrica que puede ayudar a ver esto que estamos diciendo (o tal vez: que estamos "tratando de decir"), es considerar un *paralelogramo fundamental* de $L$, dividir dos de sus lados no paralelos en $m$ partes iguales cada uno y trazar rectas paralelas a los lados por cada uno de estos $2m$ puntos. En el paralelogramo fundamental quedan así determinados $m^2$ puntos $z$ tales que si los sumamos con sí mismos $m$ veces a cada uno de ellos, el resultado $mz$ de esta suma será un punto de otro paralelogramo del retículo (no "el fundamental" de partida), pero dentro de este otro paralelogramo el punto $mz$ ocupará la misma posición (relativa) que ocupa el 0 en el fundamental.

## 8. Curvas Elípticas sobre cuerpos finitos

Sea $E$ una curva elíptica dada por

$$E: \qquad y^2=x^3+Bx+C, \qquad \text{con } B,C\in\mathbb{Z}, \tag{8.1}$$

ecuación a la que le "falta" el término en $x^2$, pero puede verse que siempre hay un cambio de variables que permite expresar toda curva elíptica $E$ sobre los racionales con una ecuación de Weierstrass de esta forma. Pero ahora estamos considerando el caso en que los coeficientes $B$ y $C$ del polinomio cúbico $f(x)=x^3+Bx+C$ son números enteros, y como siempre (para que realmente se trate de una curva elíptica) pedimos que $f$ no tenga raíces dobles ni triples (o como también se suele decir: que $f$ sea no singular). Esta última condición equivale a pedir la no anulación de su *discriminante* $D$ :

$$D(f)=\prod_{\alpha_i\neq\alpha_j,\, i<j}(\alpha_i-\alpha_j)^2=(\alpha_1-\alpha_2)^2(\alpha_1-\alpha_3)^2(\alpha_2-\alpha_3)^2\neq 0,$$

donde $\alpha_i\,(i=1,2,3)$ representan las raíces de $f$ (el discriminante de un polinomio es un buen *invariante*). Por otro lado, también tenemos el *discriminante* $\Delta$ de la curva elíptica $E$ definido por:

$$\Delta(E)=-16D(f)=-16\left(4B^3+27C^2\right), \tag{8.2}$$

el cual también va a ser distinto de cero si el discriminante de $f$ lo es.

Vamos a empezar a *controlar módulo $p$*, con $p$ primo (o como también se dice: a *reducir módulo $p$*) una curva elíptica. Esto es, dado un primo $p$, en lugar de $E$ como está dada en (8.1), vamos a trabajar con $E \pmod p$, que no será más que la propia $E$ pero considerando que tanto los coeficientes $B$ y $C$ que la definen, como sus puntos $(x, y)$, están en el conjunto $\mathbb{Z}/p\mathbb{Z}$ de los *enteros módulo $p$*. Recordemos que este conjunto, al que simbolizaremos $\mathbb{F}_p$, tiene estructura de cuerpo (pues $p$ es primo): es un cuerpo finito con $p$ elementos. Si lo pensamos "proyectivamente", en él también podemos "agregar" un *punto al infinito*. Usaremos el símbolo $E \pmod p$ para indicar una tal curva elíptica, en cuyo caso, la "igualdad" de (8.1) pasa a ser (aún cuando por comodidad o costumbre sigamos poniendo el signo igual) una congruencia módulo $p$ :

$$E \pmod p: \qquad y^2 \equiv x^3 + Bx + C, \qquad \text{con } B, C \in \mathbb{F}_p, \tag{8.3}$$

y también, como estamos diciendo: $x, y \in \mathbb{F}_p$ (esto es: no miramos a los enteros $x$ e $y$ sino a sus restos luego de dividirlos por $p$). Ahora bien: luego de este drástico cambio de punto de vista puede suceder que, módulo $p$, ya no sea $\Delta(E) \neq 0$ (o más precisamente: no sea $\Delta(E) \not\equiv 0 \pmod p$) y la curva, que pensada en $\mathbb{Z}$ es elíptica, deje de serlo en $\mathbb{F}_p$ por fallar dicha condición.

Este hecho ya "divide las aguas" entre los primos, y diremos que $p$ será un *primo de buena reducción* si al considerar $E \pmod p$ se verifica o mantiene la condición $\Delta(E) \neq 0$ (o más precisamente: $\Delta(E) \not\equiv 0 \pmod p$). En caso contrario $p$ será *de mala reducción* (estos últimos primos son, entonces, aquellos que dividen (módulo $p$) al discriminante de la curva elíptica, esto es, aquellos primos $p$ tales que $p|\Delta$).

Fijada una curva elíptica de partida $E$ del tipo (8.1), a la misma podemos reducirla módulo cada primo $p$ y resultará que (sin variar la $E$ de partida) la ecuación reducida $E \pmod p$ seguirá siendo una curva elíptica (ahora sobre $\mathbb{F}_p$) si y sólo si $p$ es un primo de buena reducción para $E$. Se sabe que para cada curva del tipo (8.1), los primos de mala reducción sólo son un número finito (dado que sólo puede haber un número finito de primos que dividen al discriminante $\Delta(E)$). Para ser más precisos, conviene aclarar que antes de reducir modulo $p$ una ecuación de una curva elíptica $E$, hay que asegurarse que la ecuación escogida es una ecuación minimal en $p$, pero omitiremos los detalles de la construcción de una tal ecuación minimal.

En lo que sigue, nos interesaremos por los primos de buena reducción. Una primera pregunta y que nos parece bastante natural es: si $p$ es un primo de buena reducción (y por lo tanto $E \pmod p$ es una curva elíptica sobre $\mathbb{F}_p$) ¿cuántos puntos $(x, y)$ con $x, y \in \mathbb{F}_p$ posee $E \pmod p$? Si simbolizamos $E/\mathbb{F}_p$ a tal conjunto, lo que nos preguntamos es: ¿cuál es la cardinalidad del conjunto $E/\mathbb{F}_p$?, esto es: ¿cuánto vale $\#E/\mathbb{F}_p$? Un primer procedimiento para determinarlo podría consistir en darle a $x$ todos sus posibles valores en $\mathbb{F}_p : 0, 1, \ldots, p-1$, reemplazar cada uno de ellos en la ecuación de $E \pmod p$ y ver para cuáles existe un $y \in \mathbb{F}_p$ tal que la congruencia (8.3) se cumple. Por cada valor de $x$ pueden haber a lo sumo dos valores de $y$ (debido a que está elevada al cuadrado), de modo que una primera cota (la más burda) para el número de puntos de $E/\mathbb{F}_p$ resulta ser $2p + 1$ y así tenemos que $\#E/\mathbb{F}_p \leq 2p + 1$ (ponemos $2p + 1$ en lugar de $2p$ porque estamos teniendo en cuenta el punto al infinito). Apenas nos detenemos un poco más en la cuestión podemos hacer el siguiente razonamiento heurístico (que luego nos lo confirmará el Teorema 8.2): como $y$ está elevada al cuadrado, esta potencia 2 de $y$ "pega" la mitad de las $y$, o dicho más precisamente (dejamos como ejercicio la prueba de esta propiedad):

*Ejercicio* 8.1. Dado un primo $p$, los residuos distintos de 0 módulo $p$ son la mitad cuadrados y la mitad no. Esto es: existen $r = \frac{p-1}{2}$ elementos de $\mathbb{F}_p^*$ tales que $r \equiv x^2 \pmod p$ al variar $x$ en $\mathbb{F}_p$.

Entonces, de los $p$ valores que puede asumir el segundo miembro al variar $x$ en $\mathbb{F}_p$ (supongamos en principio que todos son posibles), la mitad no van a estar en correspondencia con ningún $y$ que al cuadrado sea congruente con ellos, de modo que sólo "para la otra mitad" habrá dos $y$ por cada valor, y en definitiva la cota se reduce aproximadamente a $p+1$, esto es: $\#E/\mathbb{F}_p \approx p+1$.

**Teorema 8.2** (Hasse)**.** *Dada una curva elíptica $E$ y un primo $p$ de buena reducción para $E$, entonces para $E \pmod p$ se verifica que:*

$$|\#E/\mathbb{F}_p - (p+1)| < 2\sqrt{p}. \tag{8.4}$$

Digamos, antes de continuar, que:

- al número de puntos módulo $p$ que están en la curva elíptica para cada primo $p$ de buena reducción se lo simboliza $N_p$, esto es: $N_p := \#E/\mathbb{F}_p$, y
- se simboliza $a_p$ a la siguiente diferencia: $a_p := (p+1) - N_p$, que representa el número de puntos de la curva "controlados" por la cota $2\sqrt{p}$.

**Ejemplo 8.3.** Una curva elíptica peculiar

$$E: \qquad y^2 = x^3 - x. \tag{8.5}$$

Esta curva tiene una particularidad, que es que sobre $\mathbb{C}$ (o cualquier cuerpo donde $-1$ admita una raíz cuadrada) la curva admite la transformación $(x,y) \to (-x, iy)$ (que se puede verificar preserva la estructura de grupo abeliano).

Veremos a continuación la fórmula para calcular exactamente (esto es: ¡no sólo acotarlos!) los $a_p$ de esta curva elíptica. Tal fórmula se debe a P. de Fermat, quien alrededor de 1630 demostró cuáles enteros se pueden escribir como suma de los cuadrados de otros dos, y cuáles no. A nosotros nos interesará saber qué *primos* son igual a la suma de dos cuadrados, cuestión a la que P. de Fermat dio la siguiente respuesta: si $p$ es primo, entonces

$$p = x^2 + y^2 \iff p = 2 \quad \text{ó} \quad p \equiv 1 \pmod 4, \tag{8.6}$$

siendo esta descomposición, además, esencialmente única (salvo signos y/u orden de los términos). Concretamente, entonces, para la curva peculiar en consideración, el 2 es el único (¡confirmar la unicidad como ejercicio!) primo de mala reducción, y resulta que

$$E \pmod p \rightsquigarrow a_p = \begin{cases} 2\alpha & \text{si} \quad p \equiv 1 \pmod 4 \text{ y } p = \alpha^2 + \beta^2, \\ 0 & \text{si} \quad p \equiv 3 \pmod 4. \end{cases} \tag{8.7}$$

Para verificar la igualdad, notemos que si $p \equiv 3 \pmod 4$, entonces $-1$ no es un cuadrado, con lo cual si al evaluar el polinomio $p(x) = x^3 - x$ en un $x_0 \in \mathbb{F}_p$ obtenemos un valor no nulo, entonces sólo uno de los valores $p(x_0)$ o $p(-x_0)$ puede ser un cuadrado. En particular, el número de puntos de $E/\mathbb{F}_p$ es $3 + 1 + 2 \cdot \frac{p-3}{2}$, donde el 3 corresponde a las raíces de $p(x)$ $\{0, 1, -1\}$, el 1 corresponde al punto del infinito, y los otros a los valores donde $p(x_0)$ es un cuadrado no nulo (que aporta dos posibilidades para $y$). En particular, $E/\mathbb{F}_p$ tiene $p+1$ puntos con lo cual $a_p = 0$. La otra demostración es un poco mas avanzada, ver [1, Theorem 14.16], donde siguiendo su notación, $\mathcal{O} = \mathbb{Z}[i]$, y la condición $p = \pi\bar{\pi}$ es exactamente escribir

$p = (\alpha + \beta i)(\alpha - \beta i)$. Notar la ambigüedad entre $\alpha$ y $\beta$, una sola corresponde a la elección correcta.

Digamos de paso, que este resultado de P. de Fermat también viene a decirnos qué naturales pueden ser *normas* de enteros algebraicos del cuerpo $\mathbb{Q}(i)$ y, en el fondo, con esto último tiene que ver (luego de que C. Gauss publicara sus *Disquisitiones Arithmeticae*, en 1801) el teorema de P. de Fermat sobre los números que se pueden escribir como suma de dos cuadrados.

*Ejercicio* 8.4. Calcular los puntos que tiene la curva elíptica $E : y^2 = x^3 + x + 1$ sobre $\mathbb{F}_5$.

Además, sobre un cuerpo finito, obviamente todos los puntos de la curva son de torsión.

## 9. Acción del grupo de Galois en puntos de torsión

Tengamos presente estos tres resultados ya vistos:

1. $E(\mathbb{C})[m] \cong \mathbb{Z}/m\mathbb{Z} \oplus \mathbb{Z}/m\mathbb{Z}$ (Teorema 7.8),
2. $E(\mathbb{Q})$ es un grupo Abeliano finitamente generado (Teorema 7.4),
3. $E_{tor}(\mathbb{Q})$ es un grupo Abeliano cuyo orden $|E_{tor}(\mathbb{Q})|$ está acotado "universalmente" (Teorema 7.9).

A pesar de que la demostración del Teorema 7.8 involucra trabajar en el cuerpo $\mathbb{C}$ como cuerpo de variación de los puntos de una curva elíptica, en realidad las coordenadas de cualquier punto de $m$-torsión $P \in E(\mathbb{C})[m]$ estarán en la *clausura algebraica de* $\mathbb{Q}$. De esta última observación y también de las propiedades de los puntos de $m$-torsión de $E$ saldrán las representaciones del grupo de Galois.

Como los únicos números de $\mathbb{C}$ que hemos manejado y manejaremos son sus *números algebraicos* (cuyo conjunto se simboliza $\overline{\mathbb{Q}}$, al que también se lo denomina la *clausura algebraica de* $\mathbb{Q}$), simbolizaremos $E(\overline{\mathbb{Q}})[m]$ (en lugar de $E(\mathbb{C})[m]$) al conjunto de (todos) los puntos de $m$-torsión de una curva elíptica $E$ con coordenadas algebraicas (ya sean reales o complejas). Sabemos que dicho conjunto contiene $m^2$ números algebraicos, tiene estructura de grupo Abeliano, y es isomorfo a la suma directa de los enteros módulo $m$ : $E(\overline{\mathbb{Q}})[m] \cong \mathbb{Z}/m\mathbb{Z} \oplus \mathbb{Z}/m\mathbb{Z}$.

Sin pretender desarrollar la *Teoría de Galois* en toda su generalidad, ni mucho menos, lo que haremos será construir, a partir de los puntos de $m$-división de una curva elíptica $E$, su respectivo grupo de Galois.

Recordemos que con cada $N_p$ había asociado un $a_p$ y que dada una curva elíptica $E/\mathbb{Q}$ considerábamos un "primo bueno" $p$ y efectuábamos la "reducción módulo $p$" de $E$, obteniendo así $E$ (mód $p$): una curva elíptica tal que el polinomio que la define tiene, ahora, todos sus coeficientes en el cuerpo finito $\mathbb{F}_p$. Recordemos también que la relación que vincula los $N_p$ y los $a_p$ es la siguiente:

$$a_p = p + 1 - N_p. \tag{9.1}$$

Sea ahora $K$ un cuerpo de números Galoisiano sobre $\mathbb{Q}$, sea $G$ el grupo de Galois de $K$ ($G = \mathrm{Gal}(K/\mathbb{Q})$) y como hasta ahora, $E(K)$ el conjunto de todos los puntos de una curva elíptica $E$ con coordenadas en $K$. Sea $P(x, y) \in E(K)$ y $\sigma \in G$. Entonces:

1. $\sigma(P) \in E(K)$, donde obviamente $\sigma(P) = (\sigma(x), \sigma(y))$. Estamos afirmando que si $P \in E(K)$ entonces $\sigma(P) \in E(K)$ o lo que es lo mismo: si $P$ es un

punto de $E$, también lo es su transformado por cualquier automorfismo del grupo de Galois de $K$.
2. Por ser $\sigma$ un automorfismo de cuerpos, $\sigma(\mathcal{O}) = \mathcal{O}$.

**Teorema 9.1.** *Sea $E$ una curva elíptica sobre $\mathbb{Q}$, $K$ un cuerpo Galoisiano sobre $\mathbb{Q}$ y $G = \mathrm{Gal}(K/\mathbb{Q})$ su grupo de Galois. Entonces:*

1. $E(K)$ *es un subgrupo de* $E(\mathbb{C})$,
2. $P \in E(K) \Longrightarrow \forall \sigma \in G : \sigma(P) \in E(K)$,
3. $\forall P \in E(K), \forall \sigma, \tau \in G : (\sigma \circ \tau)(P) = \sigma(\tau(P)), \qquad id \in G, id(P) = P$
4. *Interacción con la estructura de grupo.* $\forall P, Q \in E(K)$ :
   a) $\sigma(P+Q) = \sigma(P) + \sigma(Q)$,
   b) $\sigma(-P) = -\sigma(P)$,
   c) $\sigma(nP) = n\sigma(P)$.
5. $P \in E(K)[m] \Longrightarrow \sigma(P) \in E(K)[m]$, *donde queremos decir lo siguiente: si $P$ es un punto de $m$-torsión de $E$ y de orden exactamente $m$ (en esta propiedad $m$ sí es minimal: es el orden mínimo de $P$) entonces $\sigma(P)$ también es un punto de $m$-torsión de $E$ y de orden exactamente $m$. O sea que no sólo se preserva lo propiedad de "ser punto de torsión", sino que también se preserva su "orden" (exacto).*

*Demostración.* Veremos la prueba de las dos últimas partes de Teorema, dejando las otras como ejercicio.

4. Esta parte del Teorema nos está diciendo que la estructura de grupo de los puntos de la curva elíptica $E(K)$ (con la suma de puntos definida geométricamente por el método de las tangentes y secantes) es compatible con la acción del grupo de Galois de $K$.

   Sean $P = (x_1, y_1), Q = (x_2, y_2)$ y $(P+Q) = (x_3, y_3)$. Entonces, por la fórmula de adición (7.18) es $x_3 = \lambda^2 - a - x_1 - x_2$, con $\lambda = \frac{y_2 - y_1}{x_2 - x_1}$, de modo que:

$$x_3 = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)^2 - a - x_1 - x_2, \qquad y_3 = \lambda x_3 + \nu \tag{9.2}$$

   (donde, recordemos, $\nu = y_1 - \lambda x_1$, y por lo tanto determinando la abscisa $x_3$, la ordenada $y_3$ queda automáticamente determinada), y por ser $\sigma$ un automorfismo resultará:

$$\sigma(x_3) = \ldots = \left(\frac{\sigma(y_2) - \sigma(y_1)}{\sigma(x_2) - \sigma(x_1)}\right)^2 - a - \sigma(x_1) - \sigma(x_2), \qquad \sigma(y_3) = \ldots \tag{9.3}$$

   de modo que:

$$\sigma(P+Q) = (\sigma(x_3), \sigma(y_3)), \qquad \text{y} \tag{9.4}$$

$$\sigma(P) + \sigma(Q) = (\sigma(x_1), \sigma(y_1)) + (\sigma(x_2), \sigma(y_2)) = \ldots \tag{9.5}$$

   y por lo tanto $\sigma(P+Q) = \sigma(P) + \sigma(Q)$, como queríamos ver.

   Las partes b) y c) se dejan como ejercicio.
5. Sabemos que $mP = \mathcal{O}$ con $m$ minimal. Entonces por la parte 4. recién demostrada:

$$m\sigma(P) = \sigma(mP) = \sigma(\mathcal{O}) = \mathcal{O}, \tag{9.6}$$

de modo que $\sigma(P)$ tiene orden dividiendo a $m$. Para ver el orden de $\sigma(P)$ es también $m$, debemos considerar $\sigma^{-1}(\sigma(P))$ y proceder en consecuencia (lo cual se deja como Ejercicio). □

Para entender el grupo de Galois que se obtiene agregando a $\mathbb{Q}$ todos los puntos de $m$-torsión de una curva elíptica $E$, sólo hacen falta las partes 4. y 5. del Teorema 9.1.

En el grupo $E(\mathbb{C})$ se define el siguiente morfismo (que no sólo es un homomorfismo de grupos, sino también un morfismo en el sentido más general de la Geometría Algebraica, en donde se definen los morfismos sobre curvas), llamado *multiplicación por* $m$ (con $m \in \mathbb{N}$):

$$\lambda_m : E(\mathbb{C}) \longrightarrow E(\mathbb{C}), P \mapsto \lambda_m(P) = mP, \tag{9.7}$$

el cual, como tal, respeta la estructura de grupo de $E(\mathbb{C})$. Pues bien: el grupo de los puntos de $m$-torsión de la curva elíptica $E$ no es más que el núcleo del morfismo $\lambda_m$ precedente:

$$\ker(\lambda_m) = E(\mathbb{C})[m]. \tag{9.8}$$

Estos morfismos $\lambda_m$ se denominan *morfismos triviales* y los poseen todas las curvas elípticas (por tener estructura de grupo abeliano). Sea $E[m]$ el conjunto de todos sus puntos de $m$-torsión distintos del punto al infinito $\mathcal{O}$, y sea $S$ el conjunto de las $2\left(m^2-1\right)$ coordenadas $(x_i, y_i)$, $1 \leq i \leq m^2-1$, de estos puntos (esto es: $S$ es un conjunto de $2\left(m^2-1\right)$ números algebraicos), entonces el cuerpo $K$ de números algebraicos en consideración será la siguiente extensión algebraica de los racionales: $K = \mathbb{Q}(S)$. También se lo simboliza $K = \mathbb{Q}(E[m])$.

**Teorema 9.2.** *Bajo las hipótesis precedentes,* $K/\mathbb{Q}$ *es una extensión de Galois.*

*Demostración.* Por lo pronto, observemos que todo automorfismo $\sigma : K \longrightarrow \mathbb{C}$ queda determinado por sus valores sobre los números algebraicos $x_i, y_i$ del conjunto $S$.

Por el Teorema 9.1, parte 5., si $P_i \in E[m]$ entonces $\sigma(P_i) \in E[m]$ y tiene el mismo orden exacto que el propio punto $P_i$ de partida. Además: $P_i \neq \mathcal{O} \Longrightarrow \sigma(P_i) \neq \mathcal{O}$, de modo que si $P_i \neq \mathcal{O}$ entonces $P_i$ es alguno de los listados en $E[m]$ y su imagen por $\sigma$ también será uno de los listados en $E[m] : \sigma(P_i) = P_j$, donde las coordenadas de $P_j$, que son $(x_j, y_j)$ están en $S$ y por lo tanto en $K$. De modo que $\sigma(K) \subset K$ y por lo tanto la extensión $K/\mathbb{Q}$ es Galois. □

Ahora nos interesaremos en describir el grupo de Galois $G = \mathrm{Gal}(K/\mathbb{Q})$ de esta extensión. Tales grupos de Galois, en general, no van a ser conmutativos. Como

$$(P \in E[m], \sigma \in G) \Longrightarrow \sigma(P) \in E[m], \tag{9.9}$$

en particular el automorfismo $\sigma$ (y por lo tanto: inyectivo) induce una permutación de los puntos de $E[n]$ que respeta la estructura de grupo:

$$\sigma(P+Q) = \sigma(P) + \sigma(Q), \qquad \sigma(-P) = -\sigma(P). \tag{9.10}$$

De modo que cada $\sigma \in G$ da lugar a un homomorfismo del grupo $E[m]$ en sí mismo.

En general, el conjunto $\mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z}$ de los enteros módulo $m$ es un anillo conmutativo (con la suma y producto habitual de clases) y con divisores de cero si $m$ no es primo (y consecuentemente $\mathbb{Z}_m$ no es un cuerpo si $m$ no es primo). Pero si $m = p$ es primo, entonces $\mathbb{Z}/p\mathbb{Z}$ es un cuerpo finito de $p$ elementos, el cual

acostumbra simbolizarse $\mathbb{F}_p$ (como en las secciones anteriores). Esto hace que el grupo Abeliano $E[m]$ sea:

- un *módulo libre* de dimensión 2 sobre $\mathbb{Z}_m$ (cuando $m$ no es primo), y
- un *espacio vectorial* de dimensión 2 sobre $\mathbb{F}_p$ (cuando $p$ es primo).

Luego, $\sigma \in G$ actuando sobre $E[p]$ es, para $p$ primo, una transformación lineal de un espacio vectorial. Y esto es lo que se denomina una *representación* del grupo $G$ (como las vistas en la sección 2.2). No todo grupo admite una representación como en este caso, y que $G$ la admita es una gran ventaja. Para toda $\sigma \in G : \sigma^{-1}$ induce el morfismo inverso, y por lo tanto la representación tiene inversa.

Cuando $m$ no es primo lo que estamos haciendo es "álgebra lineal sobre un anillo", y entonces las cosas se complican, porque allí hay elementos que no tienen inverso. Miremos entonces el caso $m = p$ primo (como veremos, si entendemos este caso vamos a entender todo). Si tomamos en $E[p]$ dos puntos $P_1$ y $P_2$ "independientes", obtenemos una base de $E[p]$, dado que

$$E[p] = \{a_1 P_1 + a_2 P_2 : a_1, a_2 \in \mathbb{F}_p\}, \tag{9.11}$$

y entonces toda transformación lineal $h : E[p] \longrightarrow E[p]$ queda completamente definida por sus valores sobre $P_1$ y $P_2$ (que, como acabamos de decir, constituyen una base para el espacio vectorial $E[p]$) :

$$h(P_1) = \alpha_h P_1 + \gamma_h P_2, \qquad h(P_2) = \beta_h P_1 + \delta_h P_2, \tag{9.12}$$

de modo que $h$ queda caracterizada por la matriz $M \in \mathrm{GL}_2(\mathbb{F}_p)$ de la transformación lineal en la base $\{P_1, P_2\}$.

Volvemos ahora al caso general ($m$ primo o no). Dada $\sigma \in G$, induce una transformación lineal $\sigma : E[m] \longrightarrow E[m], P \mapsto \sigma(P)$ (que es la acción de $\sigma$ en $E[n]$). Como dicha transformación lineal tiene inversa, si elegimos una base de $E[m]$ (como $\mathbb{Z}_m$-módulo) le podemos asociar una matriz $M \in \mathrm{GL}_2(\mathbb{Z}/m\mathbb{Z})$. En particular $\det(M) \in (\mathbb{Z}/m\mathbb{Z})^\times$, donde con $(\mathbb{Z}/m\mathbb{Z})^\times$ representamos el grupo multiplicativo de los elementos inversibles del anillo $\mathbb{Z}/m\mathbb{Z}$.

Estas representaciones que hemos construido se enmarcan dentro de las representaciones de Galois definidas anteriormente, pero provienen de la curva elíptica $E$ en cuestión.

$$G^m := \mathrm{Gal}(\mathbb{Q}(E[m])/\mathbb{Q}) \rightsquigarrow \rho_m : G^m \longrightarrow \mathrm{GL}_2(\mathbb{Z}/m\mathbb{Z}). \tag{9.13}$$

Además, la anterior también se denomina *construcción fiel*, dado que las $\rho_m$ resultan inyectivas. Pero en general no son exhaustivas (o suryectivas).

Recordar que vimos que el anillo de números $p$-ádicos se podía obtener como "pegar" los anillos $\mathbb{Z}/p^r$ para distintos valores de $r$ (esto es truncar las series (1.3)). Queremos hacer un proceso similar para construir representaciones de Galois con coeficientes en los enteros $p$-ádicos $\mathbb{Z}_p$ a partir de las representaciones $\rho_{p^r}$. Pues bien, para los anillos $\mathbb{Z}/p^r$ tenemos definidas las respectivas representaciones $\rho_{p^r}$ :

$$\begin{aligned}
\rho_p &: G^p \longrightarrow \mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z}) \\
\rho_{p^2} &: G^{p^2} \longrightarrow \mathrm{GL}_2(\mathbb{Z}/p^2\mathbb{Z}) \\
&\vdots \\
\rho_{p^r} &: G^{p^r} \longrightarrow \mathrm{GL}_2(\mathbb{Z}/p^r\mathbb{Z}) \\
&\vdots
\end{aligned}$$

Notar que $\mathbb{Q}(E[p]) \subset \mathbb{Q}(E[p^2]) \subset \ldots$. En particular, $G^p$ es un cociente de $G^{p^2}$ (dado por restringir los automorfismos al cuerpo $\mathbb{Q}(E[p])$) y así sucesivamente. Si elegimos las bases de $E[p]$, $E[p^2]$, etc de manera "compatible", o sea que si $\{Q_1, Q_2\}$ es base de $E[p^2]$ entonces la base elegida en $\rho_p$ es exactamente $\{pQ_1, pQ_2\}$, entonces si a los coeficientes de la representación $\rho_{p^2}$ los reducimos módulo $p$, obtenemos $\rho_p$. Esto se debe a que si $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, entonces para construir $\rho_{p^2}$ debemos escribir $\sigma(Q_1) = \alpha Q_1 + \beta Q_2$, y así obtenemos la primer columna de la matriz. Pero como $\sigma$ preserva suma, $\sigma(pQ_1) = \alpha pQ_1 + \beta pQ_2$, lo que da la primer columna de la matriz de $\rho_p$. Mirando todas las representaciones a juntas, podemos obtener una matriz en $\mathrm{GL}_2(\mathbb{Z}_p)$ a partir de ellas (de la misma manera como construimos los números $p$-ádicos en (1.3). Este proceso como ya mencionamos corresponde a tomar un *límite inverso* (que es una construcción general). Así obtenemos una representación, que se simboliza $\rho_{p^\infty}$ :

$$\rho_{p^\infty} : G^{p^\infty} \longrightarrow \mathrm{GL}_2(\mathbb{Z}_p). \tag{9.14}$$

Otra forma de obtener dicha representación, es a partir de los $\mathbb{Z}/p^r$-módulos $E[p^r]$, hacer una construcción similar, tomando límite inverso, lo que da un $\mathbb{Z}_p$-módulo de rango 2, llamado el *módulo de Tate* de $E$, y que se denota por $T_p(E)$. Así, podemos pensar la representación como $\rho_{p^\infty} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}(T_p(E))$. Como vimos en la sección de representaciones de Galois, a veces es preferible con trabajar con espacios vectoriales en lugar de módulos. Como $\mathbb{Z}_p \subset \mathbb{Q}_p$, simplemente podemos pensar a nuestra representación como

$$\rho_{p^\infty} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{Q}_p). \tag{9.15}$$

El criterio de Néron-Ogg-Shafarevich dice que la extensión $\mathbb{Q}(E[p^r])/\mathbb{Q}$ es no ramificada fuera de $p$ y de los primos que dividen a $\Delta(E)$, para todo valor de $r$. Luego por el Teorema de Chebotarev, el conjunto $\{\mathrm{Frob}_\ell\}$, para todo primo $\ell$ con $\ell \neq p$ es denso en $G^{p^\infty}$.

**Teorema 9.3.** *Sean $E/\mathbb{Q}$ es una curva elíptica, $p$ un número primo y $\rho_{p^\infty} : G^{p^\infty} \to \mathrm{GL}_2(\mathbb{Z}_p)$. Si $\ell$ es un primo bueno para $E$, entonces el polinomio característico de $\rho_{p^\infty}(\mathrm{Frob}_\ell) = x^2 - a_\ell x + \ell$.*

*Demostración.* Ver por ejemplo [21] Proposición 8.6. □

Lo sorprendente de este resultado está en que este polinomio característico no depende del primo $p$, sino solamente del primo $\ell$, y que la traza coincide con el invariante $a_\ell$ que definimos asociado al número de puntos de una curva elíptica sobre el cuerpo $\mathbb{F}_\ell$. De modo que si hubiéramos montado toda la construcción precedente a partir de otro primo $q \neq p$, al pasar al límite (proyectivo) hubiéramos llegado a la aplicación $\rho_{q^\infty}$ que a los Frobenius $\{Frob_\ell\}$ para $\ell$ primo, $\ell \neq p$ y $\ell \neq q$, le hubiera hecho corresponder la matriz de $\mathrm{GL}_2(\mathbb{Z}_q)$ cuyo polinomio característico hubiera coincidido con el anterior.

*Ejercicio* 9.4. A la representación de Galois $\rho_{p^\infty}$ le podemos definir una L-serie denotada $L(\rho_{p^\infty}, s)$. Probar que dicha L-serie converge absolutamente si $\Re(s) > \frac{3}{2}$.

**Pregunta:** dada la representación $\rho_{p^\infty}$, ¿que podemos decir de su imagen?

En general, los únicos morfismos de curvas elípticas son los triviales $\lambda_m$ (9.7). Cuando hay más endomorfismos, además de ellos (que siempre existen), se dice que

la curva elíptica $E$ tiene *multiplicación compleja* (MC en adelante). Estas últimas son curvas "raras". Una de ellas, por ejemplo, es $y^2 = x^3 + x$. Las curvas que tienen MC tiene la propiedad de que las representaciones de Galois $\rho_{p^\infty}$ asociadas tienen imagen "no muy grande" para todo $p$. De hecho, se sabe que la mitad de los $a_p$ son iguales a 0, y esto fuerza al grupo de matrices imagen a ser "casi abeliano", más precisamente, es un grupo que contiene un subgrupo abeliano normal de índice 2 .

Para enunciar el teorema de Serre nos ponemos, por el contrario, en el caso (el más frecuente) de curvas sin MC. Nos preguntamos:

- $\rho_p$ ¿es suryectiva ?( es decir, su imagen es todo $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$),
- $\rho_{p^\infty}$ ¿es suryectiva?

En primer lugar, se tiene el siguiente resultado (ver [19]):

**Lema 9.5** (Serre)**.** *Si $p > 3$, $\rho_p$ es sobreyectiva si y sólo si $\rho_{p^\infty}$ es sobreyectiva.*

A partir de este lema, basta estudiar el problema de las imágenes para las representaciones $\rho_p$, que tienen la ventaja de que su imagen es un grupo finito, si se desea probar que las imágenes son grandes. Es así como comienza la prueba del siguiente resultado.

**Teorema 9.6** (Serre)**.** *si $E$ no tiene MC entonces $\rho_{p^\infty}$ es suryectiva para casi todo $p$ (esto es: para todo primo $p$, salvo un número finito). El conjunto finito donde esto falla es un conjunto a veces llamado de primos excepcionales.*

El Ejemplo (8.5), de la *curva elíptica peculiar* $y^2 = x^3 - x$, lo era porque la misma ¡tiene multiplicación compleja! Otro ejemplo similar es el de la curva dada por la ecuación $y^2 = x^3 + x$, que también tiene MC, el endomorfismo extra es el siguiente: $\phi(x, y) = (-x, iy)$.

El teorema de Serre nos dice por lo tanto que las curvas sin MC tienen grupos de Galois asociados $\mathrm{Gal}(\mathbb{Q}(E[p^\infty]/\mathbb{Q})) \cong \mathrm{GL}_2(\mathbb{Z}_p)$ para casi todo primo $p$. En cambio, las curvas con MC (y esto se sabía con anterioridad al resultado de Serre) tienen grupo de Galois $\mathrm{Gal}(\mathbb{Q}(E[p^\infty]/\mathbb{Q}))$ "casi conmutativo" para todo primo $p$.

## 10. Puntos racionales en curvas de género mayor que 1

En la década de 1920, L. Mordell conjeturó que todas las curvas de complejidad superior (las de nivel siguiente al de las elípticas) poseen sólo un número finito de puntos racionales, esto es, que para ellas $\#C(\mathbb{Q}) < \infty$. Esta propiedad resultó ser cierta y fue probada por G. Faltings en 1983 (quien en 1986 recibió la medalla Fields por demostrar esta conjetura de L. Mordell y otras, debidas a J. Tate, a I. Shafarevich, etc.).

La famosa afirmación de P. de Fermat, conocida como "la Conjetura de Fermat" ó "el Último Teorema de Fermat" (se supone que del año 1637) de que para todo $n \geq 3$, no existen soluciones en números enteros para la ecuación $x^n + y^n = z^n$ fuera de las triviales (que son: $(1, 0, 1)$ y $(0, 1, 1)$ cuando $n$ es impar, y $(\pm 1, 0, \pm 1), (\pm 1, 0, \mp 1), (0, \pm 1 \pm 1)$ y $(0, \pm 1 \mp 1)$ cuando $n$ es par, además del $(0, 0, 0)$ en ambos casos), equivale (deshomogeneizando) a la afirmación de que para todo $n \geq 3$, no existen soluciones en números racionales (fuera de las triviales) para la ecuación $x^n + y^n = 1$. Esta familia de ecuaciones define: una curva elíptica (cuando $n = 3$) e infinitas curvas de complejidad superior (cuando $n \geq 4$).

Aunque en la época en que G. Faltings demostró la conjetura de Mordell, el Último Teorema de Fermat no había sido probado, la misma significó un avance

en favor de la veracidad de este último, pues podía asegurarse, al menos, que de existir soluciones para la ecuación $x^n + y^n = z^n$ (con $n \geq 4$), éstas sólo podrían ser un número finito para cada $n$.

Hoy ya sabemos que la Conjetura de Fermat es cierta. Fue probada en sus primeros casos particulares por el propio P. de Fermat (para $n = 4$, quien indicó el método a seguir en una carta), por L. Euler (alrededor de 1740, para $n = 3$), por A-M. Legendre y G. Dirichlet (en 1823, para $n = 5$), etc.; y en toda su generalidad por A. Wiles, en 1994 (ver [25]), quien en 1998 recibió un premio especial, otorgado por la Unión Matemática Internacional, en reconocimiento a su trabajo.

Concretamente, por ejemplo, la curva $C$ de complejidad superior definida por la ecuación $x^5 + y^5 = 1$ no tiene puntos racionales aparte de los triviales $(0, 1)$ y $(1, 0)$ que constituyen $C(\mathbb{Q})$ en este caso. Además, el conjunto finito $C(\mathbb{Q})$ de las curvas de complejidad superior no posee ningún tipo de "estructura" (esto es: sus puntos no forman grupo, etc.) y también por eso se hace realmente difícil abordar el estudio de los mismos en dichas curvas.

## Referencias


[1] D. A. Cox, *Galois theory*, second ed., Pure and Applied Mathematics (Hoboken), John Wiley & Sons, Inc., Hoboken, NJ, 2012.

[2] C. W. Curtis and I. Reiner, *Representation theory of finite groups and associative algebras*, AMS Chelsea Publishing, Providence, RI, 2006, Reprint of the 1962 original.

[3] M. D. Fried and M. Jarden, *Field arithmetic*, third ed., Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 11, Springer-Verlag, Berlin, 2008, Revised by Jarden.

[4] W. Fulton and J. Harris, *Representation theory*, Graduate Texts in Mathematics, vol. 129, Springer-Verlag, New York, 1991, A first course, Readings in Mathematics.

[5] C. Khare and C. S. Rajan, *The density of ramified primes in semisimple p-adic Galois representations*, Internat. Math. Res. Notices (2001), no. 12, 601–607.

[6] N. Koblitz, *p-adic numbers, p-adic analysis, and zeta-functions*, second ed., Graduate Texts in Mathematics, vol. 58, Springer-Verlag, New York, 1984.

[7] J. C. Lagarias and A. M. Odlyzko, *Effective versions of the Chebotarev density theorem*, Algebraic number fields: $L$-functions and Galois properties (Proc. Sympos., Univ. Durham, Durham, 1975), 1977, pp. 409–464.

[8] B. Levi, *Sull'equazione indeterminata del 3° ordine.*, Atti del IV Congresso Internaz. dei Matematici Roma 6-11 Aprile 1908 **II** (1909), 173–177.

[9] B. Mazur, *Modular curves and the Eisenstein ideal*, Inst. Hautes Études Sci. Publ. Math. (1977), no. 47, 33–186 (1978).

[10] T. Nagell, *Problems in the theory of exceptional points on plane cubics of genus one*, Den 11te Skandinaviske Matematikerkongress, Trondheim, 1949, Johan Grundt Tanums Forlag, Oslo, 1952, pp. 71–76.

[11] A. P. Ogg, *Rational points of finite order on elliptic curves*, Invent. Math. **12** (1971), 105–111.

[12] R. Ramakrishna, *Infinitely ramified Galois representations*, Ann. of Math. (2) **151** (2000), no. 2, 793–815.

[13] F. Rodriguez Villegas, *Experimental number theory*, Oxford Graduate Texts in Mathematics, vol. 13, Oxford University Press, Oxford, 2007.

[14] J. Rotman, *Galois theory*, second ed., Universitext, Springer-Verlag, New York, 1998.

[15] M. Rubinstein and P. Sarnak, *Chebyshev's bias*, Experiment. Math. **3** (1994), no. 3, 173–197.

[16] N. Schappacher and R. Schoof, *Beppo Levi and the arithmetic of elliptic curves*, Math. Intelligencer **18** (1996), no. 1, 57–69.

[17] E. S. Selmer, *The Diophantine equation $ax^3 + by^3 + cz^3 = 0$*, Acta Math. **85** (1951), 203–362 (1 plate).

[18] J.-P. Serre, *A course in arithmetic*, Springer-Verlag, New York-Heidelberg, 1973, Translated from the French, Graduate Texts in Mathematics, No. 7.

[19] ———, *Abelian l-adic representations and elliptic curves*, second ed., Advanced Book Classics, Addison-Wesley Publishing Company, Advanced Book Program, Redwood City, CA, 1989, With the collaboration of Willem Kuyk and John Labute.
[20] ———, *Lectures on $N_X(p)$*, Chapman & Hall/CRC Research Notes in Mathematics, vol. 11, CRC Press, Boca Raton, FL, 2012.
[21] J. H. Silverman, *The arithmetic of elliptic curves*, second ed., Graduate Texts in Mathematics, vol. 106, Springer, Dordrecht, 2009.
[22] P. Stevenhagen and H. W. Lenstra, Jr., *Chebotarëv and his density theorem*, Math. Intelligencer **18** (1996), no. 2, 26–37.
[23] I. Stewart, *Galois theory*, fourth ed., CRC Press, Boca Raton, FL, 2015.
[24] W. Trinks, *Ein Beispiel eines Zahlkörpers mit der Galoisgruppe* PSL(3, 2) *über* $\mathbb{Q}$, Manuscript Univ. Karlsruhe, 1968.
[25] A. Wiles, *Modular elliptic curves and Fermat's last theorem*, Ann. of Math. (2) **141** (1995), no. 3, 443–551.

Universitat de Barcelona, Gran Via de les Corts Catalanes, 585, 08007 Barcelona, España
*Email address*: ldieulefait@ub.edu

FaMAF-CIEM, Universidad Nacional de Córdoba. C.P:5000, Córdoba, Argentina.
*Email address*: apacetti@famaf.unc.edu.ar

ICTP Trieste.
*Email address*: villegas@ictp.it

Curso

# INTRODUCCIÓN A GRUPOS ARITMÉTICOS

EMILIO A. LAURET, ROBERTO J. MIATELLO, Y BENJAMIN LINOWITZ

Publicaciones Matemáticas del Uruguay
Volumen 18, Diciembre 2023, Páginas 165–201
S 0797-1443

# INTRODUCCIÓN A GRUPOS ARITMÉTICOS

EMILIO A. LAURET, ROBERTO J. MIATELLO, Y BENJAMIN LINOWITZ

## Índice



## Parte 1. Fundamentos de grupos aritméticos

### 1. Introducción

El objetivo de estas notas es introducir y describir las principales propiedades de los subgrupos aritméticos que es una clase muy importante de subgrupos discretos de grupos de Lie reductivos que a la vez son de gran utilidad en geometría diferencial y en teoría de números.

En estas notas $G$ denotará un grupo de Lie, es decir, un grupo que es a la vez una variedad diferenciable real donde la función $(x, y) \mapsto xy^{-1}$ es de clase $C^\infty$. Requeriremos además, que $G$ tenga un número finito de componentes conexas. El lector no familiarizado con grupos de Lie, puede pensar en ejemplos concretos de grupos de matrices tales como $\mathrm{GL}_n(\mathbb{R})$, $\mathrm{GL}_n(\mathbb{C})$, $\mathrm{SL}_n(\mathbb{R})$, $\mathrm{SL}_n(\mathbb{C})$, $\mathrm{SO}(n, m)$, $\mathrm{SU}(n, m)$, que iremos definiendo de manera precisa en el curso de estas notas.

Versión final: 6 de junio de 2019.

Estas notas corresponden al curso dictado por los autores en la escuela AGRA III, Aritmética, Grupos y Análisis, del 9 al 20 de Julio de 2018 en Córdoba, Argentina. Las primeras cinco secciones fueron redactadas por E. Lauret y R. Miatello de manera conjunta, mientras que las últimas cuatro fueron escritas por B. Linowitz.

El primer caso que surge naturalmente es el de un subgrupo discreto $L$ de $\mathbb{R}^n$, es decir $L = \sum_{j=1}^{m} \mathbb{Z} v_j$ donde $v_1, \ldots, v_m$ son vectores $\mathbb{R}$-linealmente independientes de $\mathbb{R}^n$. Un subgrupo $L$ discreto de rango máximo ($m = n$) es denominado *retículo*. En ese caso, el cociente $\mathbb{R}^n/L$ es isomorfo a un toro $n$-dimensional.

Otro ejemplo natural surge de la geometría hiperbólica. El plano hiperbólico se identifica al semiplano superior $H$ de $\mathbb{C}$ munido de la métrica hiperbólica y en $H$ el grupo $\mathrm{SL}_2(\mathbb{R})$ actúa por transformaciones de Möbius. Esto es, dado $g = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ sea $g \cdot z := \frac{az+b}{cz+d}$. El subgrupo de isotropía de $i$ es el subgrupo $\mathrm{SO}(2)$ y de este modo se tiene la identificación $\mathrm{SO}(2)\backslash \mathrm{SL}_2(\mathbb{R}) \simeq H$.

Los subgrupos discretos de $\mathrm{SL}_2(\mathbb{R})$ tales que el cociente $H/\Gamma$ tiene medida finita son llamados subgrupos fuchsianos de primera clase y el espacio cociente se puede representar por una región fundamental con identificaciones en el borde. Para estos grupos la región puede tomarse como un polígono hiperbólico con un número finito de lados. En consecuencia, el cociente tiene área hiperbólica finita.

Los casos más simples son los llamados subgrupos principales de congruencia $\Gamma(N)$, con $N \in \mathbb{N}$, donde

$$\Gamma(N) = \{g \in \mathrm{SL}_2(\mathbb{Z}) : g \equiv \mathrm{Id} \pmod{N}\}. \tag{1.1}$$

Se tiene la sucesión exacta

$$1 \to \Gamma(N) \to \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}_N) \to 1 \tag{1.2}$$

que muestra que $\Gamma(N)$ tiene índice finito en $\mathrm{SL}_2(\mathbb{Z})$.

Una región fundamental estándar para $\mathrm{SL}_2(\mathbb{Z})$ es la región triangular con área $\pi/3$ dada por

$$\mathcal{F} = \left\{ z \in H : |\operatorname{Re}(z)| \leq \tfrac{1}{2}, \quad |z| \geq 1 \right\}.$$

En el caso del grupo $\Gamma(2)$ que tiene índice 6 en $\mathrm{SL}_2(\mathbb{Z})$ se puede tomar

$$\mathcal{F} = \left\{ z \in H : |\operatorname{Re}(z)| \leq 1, \quad |z - \tfrac{1}{2}| \geq 1, \quad |z + \tfrac{1}{2}| \geq 1 \right\}.$$

Observemos que en estos casos $H/\Gamma$ no es compacto, luego tampoco lo es $\mathrm{SL}_2(\mathbb{R})/\Gamma$, pero sí es de área finita. En el caso de $\Gamma(2)$, el área es igual a $2\pi$.

**Definición 1.1.** Si $G$ es un grupo de Lie y $\Gamma \subset G$ es un subgrupo, $\Gamma$ es un *retículo* si $\Gamma$ es discreto y de covolumen finito; $\Gamma$ se dice *cocompacto* si $G/\Gamma$ es compacto.

**Ejemplo 1.2.** Algunos ejemplos típicos de retículos:

(i) Si $\{v_j\}_1^n$ es una base de $\mathbb{R}^n$, $L = \sum_1^n \mathbb{Z} v_j$ es un retículo en $\mathbb{R}^n$. Todo retículo en $\mathbb{R}^n$ es de este tipo.

(ii) $\Gamma(N)$ es un retículo en $\mathrm{SL}_2(\mathbb{R})$, para todo $N \in \mathbb{N}$.

(iii) $\mathrm{SL}_n(\mathbb{Z})$ es un retículo en $\mathrm{SL}_n(\mathbb{R})$ para todo $n$. Esto será probado más adelante.

(iv) Sea $\mathrm{O}(p,q) = \{g \in \mathrm{GL}_{p+q}(\mathbb{R}) : gJg^t = J\}$ donde $J = \left[\begin{smallmatrix} I_p & 0 \\ 0 & -I_q \end{smallmatrix}\right]$ es el grupo de transformaciones lineales en $\mathbb{R}^{p+q}$ que preservan la forma cuadrática $\sum_1^p x_j^2 - \sum_1^q x_i^2$. Este grupo contiene el retículo $\mathrm{O}(p,q)_{\mathbb{Z}}$ de matrices enteras en $\mathrm{O}(p,q)$.

Los ejemplos (ii), (iii) y (iv) son casos particulares de los llamados grupos aritméticos. Nos interesarán particularmente los subgrupos cocompactos, que son más difíciles de construir. Estos serán tratados en las secciones 6 a 9 de las notas.

**Agradecimientos.** Benjamin Linowitz agradece a Ángel Villanueva por la excelente traducción del inglés de la segunda parte. Asimismo los autores agradecen la detallada labor del referí por sus numerosas sugerencias que ayudaron a mejorar la presentación de estas notas.

## 2. Preliminares

En esta sección introduciremos varios conceptos necesarios para el desarrollo de los temas principales de las notas.

**2.1. Medidas y nociones topológicas.** Un *grupo topológico* $G$ es un grupo que es a la vez un espacio topológico donde la función $\phi(x,y) = xy^{-1}$, $\phi : G \times G \to G$ es continua. El grupo $G$ se dice *localmente compacto* si para cada punto existe un entorno compacto. Se asume además que tal grupo es Hausdorff, esto es dados dos puntos distintos $x$ e $y$, existen entornos disjuntos $U_x$ e $U_y$ de $x$ e $y$. Todo grupo topológico localmente compacto admite una medida invariante a izquierda positiva boreliana que es única salvo múltiplos positivos. Es la llamada *medida de Haar*.

Sea $G$ un grupo topológico localmente compacto. Se denota $C_c(G)$ al espacio de funciones continuas complejas sobre $G$ con soporte compacto. Fijamos $\mu_G$ una medida de Haar invariante a izquierda de $G$. Si $f \in C_c(G)$, la integral de $f$ con respecto a $\mu_G$ es

$$\int_G f(g) d\mu_G(g).$$

Para cada $x \in G$, la aplicación

$$f \mapsto \int_G f(gx) d\mu_G(g)$$

define una nueva medida invariante a izquierda en $G$. Por la unicidad de la medida de Haar, a menos de un múltiplo positivo, existe $\Delta_G : G \to \mathbb{R}^+$ ($\mathbb{R}^+ := \{t \in \mathbb{R} : t > 0\}$) en $G$, la *función modular*, definida por la ecuación

$$\Delta_G(x) \int_G f(g) d\mu_G(g) = \int_G f(gx) d\mu_G(g),$$

para cada $f \in C_c(G)$. Se puede ver que $\Delta_G$ es un morfismo continuo (Ejercicio 2.1).

Se dice que $G$ es *unimodular* si $\Delta_G \equiv 1$ esto es, si toda medida de Haar invariante a izquierda es también invariante a derecha. Se prueba que todos los grupos compactos son unimodulares, así como los grupos conmutativos y nilpotentes (ver Ejercicio 2.2). También lo son los grupos con conmutador denso (e.g. los grupos semisimples), esto es $\overline{[G,G]} = G$, ya que $\Delta_G([x,y]) = \Delta_G(xyx^{-1}y^{-1}) = 1$ para todo $x, y \in G$ pues $\mathbb{R}^+$ es abeliano.

**Ejemplo 2.1.** El siguiente grupo no es unimodular:

$$\left\{ \begin{pmatrix} y & x \\ 0 & y^{-1} \end{pmatrix} : x \in \mathbb{R},\, y \in \mathbb{R}^+ \right\} \simeq \mathbb{R}^+ \ltimes \mathbb{R}.$$

Sea $H$ un subgrupo cerrado de $G$. Denotamos por $C_c(G/H)$ al espacio de funciones complejas sobre $G/H$ con soporte compacto.

**Teorema 2.2.** *Si $G$ es unimodular y $H$ es un subgrupo cerrado unimodular, el espacio homogéneo $G/H$ admite una medida $G$-invariante. Esta medida es única salvo un múltiplo escalar positivo.*

La demostración del teorema se puede encontrar en [14, Lem. 1.4].

**2.2. Grupos de Lie.** Un grupo de Lie $G$ es un grupo topológico que a la vez es una variedad diferenciable real en el cual las operaciones $(x, y) \to xy$ y $x \to x^{-1}$ son de clase $C^\infty$. En particular, si $g \in G$, las traslaciones a izquierda y a derecha, $l_g : x \mapsto gx$ y $r_g : x \mapsto xg$ respectivamente, son difeomorfismos de $G$.

Un grupo de Lie mantiene todas las propiedades locales de $\mathbb{R}^n$, por ejemplo es localmente compacto y localmente conexo. Además, la componente conexa de la identidad $G^0$ de $G$ es un subgrupo abierto normal de $G$ y $\#G/G^0$ es igual al número de componentes conexas de $G$. En estas notas trabajaremos con grupos que poseen un número finito de componentes conexas.

Observamos que todo subgrupo abierto $H$ de un grupo de Lie $G$ es también cerrado pues $G$ es unión disjunta

$$G = \Big( \bigcup_{g \notin H} gH \Big) \cup H$$

y como $gH$ es abierto, para todo $g \in G$, vemos que el complemento de $H$ es también abierto.

Un grupo de Lie $G$ posee un cubrimiento universal simplemente conexo $\widehat{G}$ y existe una aplicación $\pi : \widehat{G} \to G$ que es un isomorfismo local suryectivo, en particular $\ker(\pi)$ es un subgrupo normal discreto que está contenido en el centro de $G$ (Ejercicio 2.6). Si $G$ es simplemente conexo, entonces $G$ coincide con $\hat{G}$.

Un grupo $G$ se dice *simple* si no posee subgrupos normales conexos no triviales y se dice *semisimple* si su cubrimiento universal es isomorfo a un producto de factores simples. Los grupos de Lie reales simples están clasificados (E. Cartan) y los grupos simples simplemente conexos están en correspondencia con las álgebras de Lie reales simples (ver [9] o [8]).

A modo de ejemplo, en estas notas trabajaremos con varios grupos semisimples, tales como los grupos compactos $\mathrm{SO}(n)$ y $\mathrm{SU}(n)$ de matrices ortogonales o unitarias de determinante 1, los grupos lineales especiales $\mathrm{SL}_n(\mathbb{R})$ y $\mathrm{SL}_n(\mathbb{C})$, y los grupos $\mathrm{O}(p, q)$ de transformaciones invertibles de $\mathbb{R}^n$, $n = p + q$, que preservan una forma cuadrática no degenerada de signatura $(p, q)$.

**2.3. Retículos.** Ahora consideremos el caso en que $H = \Gamma$ es un subgrupo discreto de $G$, es decir, para todo $\gamma \in \Gamma$, el conjunto $\{\gamma\}$ es abierto, con la topología relativa de $G$.

**Lema 2.3.** *Todo grupo discreto es unimodular.*

*Demostración.* Sea $\Gamma$ un grupo discreto. Como $\mu_\Gamma$ es una medida de Haar invariante a izquierda, si $\chi_x$ denota la función característica del conjunto $\{x\} \subseteq \Gamma$ para $x \in \Gamma$, tenemos que

$$\mu_\Gamma(\{x\}) = \int_\Gamma \chi_x(h) d\mu_\Gamma(h) = \int_\Gamma \chi_e(x^{-1}h) d\mu_\Gamma(h) = \int_\Gamma \chi_e(h) d\mu_\Gamma(h) = \mu_\Gamma(\{e\}).$$

Luego, todo conjunto puntual de $\Gamma$ tiene la misma medida (positiva).

Si $f \in C_c(\Gamma)$, entonces $f$ es idénticamente nula salvo en un conjunto finito, digamos $\{h_1, \dots, h_d\} \subset \Gamma$, entonces

$$\int_\Gamma f(h) d\mu_\Gamma(h) = \sum_{i=1}^{d} f(h_i) \mu_\Gamma(\{h_i\}) = \left( \sum_{i=1}^{d} f(h_i) \right) \mu_\Gamma(\{e\})$$

y por otro lado, para cada $x \in \Gamma$ se tiene que

$$\int_\Gamma f(hx)d\mu_\Gamma(h) = \sum_{i=1}^{d} f(h_i)\mu_\Gamma(\{h_i x^{-1}\}) = \left(\sum_{i=1}^{d} f(h_i)\right)\mu_\Gamma(\{e\}).$$

Luego $\Delta_\Gamma \equiv 1$ y $\Gamma$ es unimodular. □

**Definición 2.4.** Un subgrupo $\Gamma$ de $G$ es un *retículo* (o *lattice*) si $\Gamma$ es discreto y $G/\Gamma$ admite una medida $G$-invariante finita. Además, $\Gamma$ se dice *cocompacto* o *uniforme* si $G/\Gamma$ es compacto.

*Nota* 2.5. Todo subgrupo discreto y cocompacto de $G$ es un retículo.

**Ejemplo 2.6.** Sea $V$ un espacio vectorial real. Un subgrupo $\mathbb{Z}v_1 \oplus \ldots \oplus \mathbb{Z}v_m$ con $v_i \in V$ para todo $i$ es un retículo si y sólo si el conjunto $\{v_1, \ldots, v_m\}$ es una base de $V$. Más aún, en este caso, es un retículo cocompacto pues $V/\mathbb{Z}v_1 \oplus \ldots \oplus \mathbb{Z}v_m$ es homeomorfo a un toro de dimensión $m$.

**Teorema 2.7.** *Sea $G$ un grupo de Lie con un número finito de componentes conexas. Sea $\Gamma$ un retículo en $G$ y sea $\pi : G \to G/\Gamma$ la proyección canónica. Entonces, las siguientes condiciones son equivalentes:*

(i) *$G/\Gamma$ es compacto.*
(ii) *No existen $g_n \in G$ y $\gamma_n \in \Gamma$, tales que $g_n\gamma_n g_n^{-1} \neq e$ para todo $n$ y $g_n\gamma_n g_n^{-1} \to e$ cuando $n \to \infty$.*
(iii) *Sea $\pi_g(x) = \pi(xg)$. Entonces existe $U$ un entorno abierto de $e$ en $G$ tal que $\pi_g|_U$ es inyectiva para todo $g \in G$.*

*Demostración.* Mostremos primero que la condición (i) implica la (ii), razonando por el absurdo. Supongamos que $g_n\gamma_n g_n^{-1} \to e$ y $g_n\gamma_n g_n^{-1} \neq e$ para todo $n$. Como $G/\Gamma$ es compacto, existe una subsucesión convergente $g_{n_j}\Gamma \to g\Gamma$. Luego, existen $\beta_{n_j} \in \Gamma$ tal que $g_{n_j}\beta_{n_j} \to g$ en $G$, o sea $g_{n_j} = g\varepsilon_{n_j}\beta_{n_j}^{-1}$, con $\varepsilon_{n_j} \to e$. Por lo tanto, la sucesión $\alpha_j := \beta_{n_j}^{-1}\gamma_{n_j}\beta_{n_j}$ de $\Gamma$ satisface que $\alpha_j \neq e$ para todo $j$ y a su vez tiene límite $e$, un absurdo pues $\Gamma$ es discreto.

Ahora asumamos (ii) y supongamos que (iii) es falsa. Fijemos una base de entornos abiertos conexos $U_m$ de $e$ con $U_m = U_m^{-1}$, $U_{m+1}^2 \subset U_m$ para todo $m$ y tales que $\cap_m U_m = \{e\}$. Supongamos que para todo $m$ existe $g_m \in G$ tal que $\pi_{g_m}|_{U_m}$ no es inyectiva. Entonces, existen $u_m$, $v_m$ en $U_m$ y $\gamma_m \in \Gamma$ con $u_m \neq v_m$ y $u_m g_m = v_m g_m \gamma_m$. Luego, $v_m^{-1}u_m = g_m\gamma_m g_m^{-1} \to e$ y $v_m^{-1}u_m \neq e$ para todo $m$, lo cual contradice (ii).

Ahora asumiendo (iii), supongamos que $G/\Gamma$ no es compacto. Sea $U$ un entorno abierto conexo de $e$ con $\bar{U}$ compacto, $U = U^{-1}$ y sea $V = U^2$. Probaremos que $\pi_g|_U$ no es inyectiva para algún $g \in G$. Sea $C$ compacto en $G/\Gamma$. Entonces $G/\Gamma \setminus VC \neq \emptyset$. Sea $g_1 \in G$ con $g_1\Gamma \in G/\Gamma \setminus VC$. Entonces, $Ug_1\Gamma \subset G/\Gamma \setminus UC$. Similarmente existe $g_2 \in G$ con $g_2\Gamma \in G/\Gamma \setminus VC \cup Vg_1\Gamma$. Iterando este procedimiento se obtiene una sucesión de conjuntos disjuntos conexos $Ug_n\Gamma$ en $G/\Gamma$ para cada $n \in \mathbb{N}$. Si $\pi_g|_U$ es inyectiva para todo $n$ entonces $\pi_g|_U : U \to \pi_g(U)$ es un homeomorfismo y $\mu_{G/\Gamma}(\pi_g(U)) = \mu_G(U)$ para todo $n$. Luego $\mu(G/\Gamma) = \infty$, un absurdo pues $\Gamma$ es un retículo. □

**2.4. Conmensurabilidad.** La siguiente noción será muy útil en el resto de estas notas.

**Definición 2.8.** Dos subgrupos $A$ y $B$ en un grupo $G$ se dicen *conmensurables* si la intersección es de índice finito en ambos, es decir, si $|A/A\cap B| = [A : A\cap B] < \infty$ y $|B/A\cap B| = [B : A\cap B] < \infty$.

Se puede ver que la conmensurabilidad es una relación de equivalencia (Ejercicio 2.7)

**Proposición 2.9.** *Sean $\Gamma$ y $\Gamma'$ dos subgrupos conmensurables de un grupo topológico localmente compacto $G$. Entonces, si $\Gamma$ es discreto, un retículo, o un retículo cocompacto, $\Gamma'$ también lo es.*

*Demostración.* Si $\Gamma$ es discreto, entonces $\Gamma\cap\Gamma'$ es discreto. Para probar que $\Gamma'$ es discreto, es suficiente ver que toda sucesión convergente en $\Gamma'$ es estacionaria.

Sea $x_n \in \Gamma'$ una sucesión convergente a $x$. Sea $\{\gamma_1 = e, \gamma_2, \dots, \gamma_d\} \subset \Gamma'$ los distintos representantes de $\Gamma'/\Gamma\cap\Gamma'$. Para cada $n\in\mathbb{N}$, existe $\theta_n \in \Gamma\cap\Gamma'$ tal que $x_n = \gamma_{i_n}\theta_n$ con $i_n \in \{1,\dots,d\}$. Luego, existe al menos un $t$ tal que $i_n = t$ para una cantidad infinita de índices $n\in\mathbb{N}$. Definimos la subsucesión $n_j$, $j\in\mathbb{N}$ dada por tales índices.

Tenemos que $x_{n_j} = \gamma_t\theta_{n_j} \to x$ cuando $j\to\infty$, por lo tanto $\theta_{n_j}\to\gamma_t^{-1}x$. Como para cada $j\in\mathbb{N}$, $\theta_{n_j}$ vive en el conjunto discreto $\Gamma\cap\Gamma'$, la sucesión $\theta_{n_j}$ es estacionaria, es decir, $\theta_{n_j} = \theta \in \Gamma\cap\Gamma'$ para todo $j\geq J$ con $J\in\mathbb{N}$ suficientemente grande. Luego la subsucesión $x_{n_j}$ es estacionaria.

Si probamos que existe $N\in\mathbb{N}$ tal que $i_n = t$ para todo $n\geq N$, la prueba estará completa. Supongamos contrariamente que existen infinitos índices $m\in\mathbb{N}$ tal que $i_m \neq t$. En ese caso, existe $s\in\{1,\dots,d\}$ con $s\neq t$ e índices $m_k$ tales que $m_k = s$ para todo $k\in\mathbb{N}$ y así $x_{m_k} = \gamma_s\theta_{m_k}$. Entonces

$$\begin{cases} x_{n_j} = \gamma_t\theta_{n_j} \text{ con } \theta_{n_j} = \theta \text{ para todo } j\geq J, \\ x_{m_k} = \gamma_s\theta_{m_k} \text{ con } \theta_{m_k} = \theta_1 \text{ para todo } k\geq K. \end{cases}$$

Además, como $x_{n_j}$ y $x_{m_k}$ convergen a $x$, obtenemos que

$$\gamma_s^{-1}\gamma_t = \theta_{m_k}x_{m_k}^{-1}x_{n_j}\theta_{n_j}^{-1} \to \theta_1\theta^{-1},$$

por lo tanto $\gamma_s^{-1}\gamma_t = \theta_1\theta^{-1} \in \Gamma\cap\Gamma'$, lo que es una contradicción pues $\gamma_s$ y $\gamma_t$ son representantes distintos de $\Gamma/\Gamma\cap\Gamma'$. Luego $\Gamma'$ es discreto.

Veamos ahora que si $\Gamma$ es un retículo, entonces $\Gamma\cap\Gamma'$ es también un retículo. En efecto

$$\mathrm{vol}(G/\Gamma\cap\Gamma') = \mathrm{vol}(G/\Gamma)\cdot[\Gamma : \Gamma\cap\Gamma'] < \infty.$$

Por otro lado, $\Gamma'$ es un retículo pues

$$\mathrm{vol}(G/\Gamma') = \mathrm{vol}(G/\Gamma\cap\Gamma')[\Gamma' : \Gamma\cap\Gamma']^{-1} < \infty.$$

Supongamos finalmente que $G/\Gamma$ es compacto. Para ver que $G/\Gamma'$ es compacto basta ver que $G/\Gamma\cap\Gamma'$ es compacto. Sea $\pi' : G\to G/\Gamma\cap\Gamma'$ la proyección canónica. Como $G/\Gamma$ es compacto, existe $\Omega\subset G$ compacto tal que $\pi(\Omega) = G/\Gamma$ (Ejercicio 2.8). Ahora $\Gamma/\Gamma\cap\Gamma' = \bigcup_{i=1}^d \gamma_i(\Gamma\cap\Gamma')$ con $\gamma_i\in\Gamma$, $1\leq i\leq d$. Probaremos que

$$\pi'\left(\bigcup_{i=1}^{d}\Omega\gamma_i\right) = G/\Gamma\cap\Gamma'. \tag{2.1}$$

Si $x\in G$, entonces $k^{-1}x\in\Gamma$ para algún $k\in\Omega$, dado que $\pi(\Omega) = G/\Gamma$. Luego, $k^{-1}x = \gamma_i\theta$, $\theta\in\Gamma\cap\Gamma'$ para algún $i\in\{1,\dots,d\}$. Por lo tanto

$$\pi'(x) = \pi'(k\gamma_i\theta) = \pi'(k\gamma_i) \in \pi'(\Omega\gamma_i),$$

lo que verifica (2.1), luego $G/\Gamma \cap \Gamma'$ es compacto como se afirmó. Esto concluye la prueba de la proposición. □

### 2.5. Ejercicios.

**Ejercicio 2.1.** Demostrar que la función modular $\Delta_G : G \to \mathbb{R}^+$ de un grupo topológico localmente compacto $G$ es un morfismo continuo.

**Ejercicio 2.2.** Demostrar que las siguientes clases de grupos son unimodulares:

1. Grupos compactos.
2. Grupos conmutativos.
3. Grupos topológicos localmente compactos que admiten un retículo.

**Ejercicio 2.3.** Demostrar que el grupo en Ejemplo 2.1 no es unimodular. Además, encontrar el isomorfismo señalado en tal ejemplo.

**Ejercicio 2.4.** Probar que $\int_G f(x^{-1})dx = \int_G f(x)\Delta(x^{-1})dx$ para toda $f \in C_c(G)$.

**Ejercicio 2.5.** Sea $\Gamma$ un retículo en un grupo de Lie $G$. Probar que $\Gamma$ es finito si y sólo si $G$ es compacto.

**Ejercicio 2.6.** Sea $G$ un grupo de Lie con cubrimiento universal $\pi : \widehat{G} \to G$. Mostrar que $\ker(\pi)$ es un subgrupo normal discreto que está contenido en el centro de $G$.

**Ejercicio 2.7.** Demostrar que la conmensurabilidad es una relación de equivalencia.

**Ejercicio 2.8.** Sea $\Gamma$ un retículo cocompacto de un grupo de Lie $G$ y $\pi : G \to G/\Gamma$ la proyección canónica. Mostrar que existe un subespacio compacto $\Omega$ de $G$ tal que $\pi(\Omega) = G/\Gamma$.

**Ejercicio 2.9.** Un grupo se llama sin torsión si no tiene subgrupos finitos no triviales. Mostrar que $\mathrm{SL}_n(\mathbb{Z})$ tiene un subgrupo de índice finito sin torsión.

## 3. Grupos aritméticos

En esta sección introduciremos los subgrupos aritméticos de grupos algebraicos definidos sobre los números racionales. Los puntos reales de tales grupos son grupos de Lie semisimples.

**3.1. Grupos algebraicos.** Sean $V$ un $\mathbb{C}$-espacio vectorial de dimensión finita y $\mathbb{K}$ un subcuerpo de $\mathbb{C}$. Sea $V_{\mathbb{K}}$ una $\mathbb{K}$-forma de $V$, esto es, $V_{\mathbb{K}}$ es un $\mathbb{K}$-espacio vectorial tal que $V \simeq V_{\mathbb{K}} \otimes_{\mathbb{K}} \mathbb{C}$, o equivalentemente, $V_{\mathbb{K}}$ tiene una $\mathbb{K}$-base que también es $\mathbb{C}$-base de $V$.

Denotamos $\mathbb{C}[V]$ al espacio de funciones polinomiales en $V$ con coeficientes complejos. Esto es, $P \in \mathbb{C}[V]$ si $P : V \to \mathbb{C}$, y para una base de $V$, $\{v_1, \ldots, v_n\}$, se tiene que

$$P(a_1 v_1 + \ldots + a_n v_n) = \sum_{\alpha=(\alpha_1,\ldots,\alpha_n)\in\mathbb{N}_0^n} c_\alpha \, a_1^{\alpha_1} \ldots a_n^{\alpha_n},$$

con $c_\alpha \in \mathbb{C}$, todos nulos salvo una cantidad finita. Denotamos $\mathbb{K}[V]$ a las funciones polinomiales en $V$ que tienen coeficientes en $\mathbb{K}$ (ver Ejercicio 3.3).

Una *variedad algebraica* $Z$ es un subconjunto de $V$ que a su vez es el conjunto de ceros de una familia de funciones polinomiales en $V$. Sea $I(Z) \subseteq \mathbb{C}[V]$ el ideal de polinomios en $V$ que se anulan en $Z$ y $\mathbb{C}[Z] := \mathbb{C}[V]/I(Z)$ el anillo de funciones

regulares en $Z$. Diremos que $Z$ *está definida sobre* $\mathbb{K}$ (o que es una $\mathbb{K}$-*variedad algebraica*) si $I(Z)$ es generado como ideal por la intersección $I(Z) \cap \mathbb{K}[V]$. Esto es equivalente a que la variedad puede ser definida como los ceros de funciones polinomiales en $\mathbb{K}[V]$ (ver Ejercicio 3.4). Denotemos $\mathbb{K}[Z] = \mathbb{K}[V]/(I(Z) \cap \mathbb{K}[V])$ al anillo de *funciones* $\mathbb{K}$-*regulares de* $Z$. Un $\mathbb{K}$-*morfismo* de $\mathbb{K}$-variedades $\varphi : Z_1 \to Z_2$ es un mapa tal que $f \circ \varphi \in \mathbb{K}[Z_1]$ para todo $f \in \mathbb{K}[Z_2]$. En este caso diremos que $\varphi$ está definido sobre $\mathbb{K}$.

Notemos que el grupo de transformaciones lineales invertibles $\mathrm{GL}(V)$ hereda una estructura de variedad algebraica del espacio vectorial $\mathrm{End}_{\mathbb{C}}(V) \times \mathbb{C}$ de tal modo que $\mathrm{GL}(V)$ se identifica con la variedad algebraica $\{(T,t) \in \mathrm{End}_{\mathbb{C}}(V) \times \mathbb{C} : \det(T)t - 1 = 0\}$ por $T \mapsto (T, \det T^{-1})$. Más aún, $\mathrm{GL}(V)$ está definido sobre $\mathbb{Q}$, y por lo tanto sobre cualquier subcuerpo $\mathbb{K}$ de $\mathbb{C}$.

**Definición 3.1.** Un *grupo algebraico lineal definido sobre* $\mathbb{K}$ es un subgrupo $\mathbf{G} \subset \mathrm{GL}(V)$ que es una $\mathbb{K}$-subvariedad algebraica de $\mathrm{End}_{\mathbb{C}}(V)$ con respecto a la $\mathbb{K}$-forma $\mathrm{End}(V)_{\mathbb{K}}$.

*Nota* 3.2. La $\mathbb{K}$-forma $\mathrm{End}(V)_{\mathbb{K}}$ de $\mathrm{End}_{\mathbb{C}}(V)$ está definida en el Ejercicio 3.2.

En lo sucesivo, abreviaremos llamando $\mathbb{K}$-*grupo algebraico* a un grupo algebraico lineal definido sobre $\mathbb{K}$. En el caso $\mathbb{K} = \mathbb{C}$, diremos simplemente que $\mathbf{G}$ es un grupo algebraico.

*Nota* 3.3. Cada vez que tomamos un $\mathbb{K}$-grupo algebraico $\mathbf{G} \subset \mathrm{GL}(V)$, estamos asumiendo que $V$ posee una $\mathbb{K}$-forma $V_{\mathbb{K}}$, que a la vez induce la $\mathbb{K}$-forma $\mathrm{End}_{\mathbb{C}}(V)_{\mathbb{K}}$ que define la estructura de $\mathbb{K}$-variedad algebraica de $\mathbf{G}$. Fijada $\{v_1, \ldots, v_n\}$ una $\mathbb{K}$-base arbitraria de $V_{\mathbb{K}}$, se tiene la identificación de $\mathrm{GL}(V)$ con $\mathrm{GL}_n(\mathbb{C})$ dada por $T \mapsto (a_{i,j})_{i,j}$ donde $Tv_i = \sum_j a_{i,j} v_j$. También se tienen las identificaciones $V \equiv \mathbb{C}^n$ y $V_{\mathbb{K}} \equiv \mathbb{K}^n$.

Además, cuando escribamos $\mathrm{GL}_n(\mathbb{C})$, estaremos asumiendo que la $\mathbb{K}$-forma escogida en $V = \mathbb{C}^n$ es $V_{\mathbb{K}} = \mathbb{K}^n$. Además, si $\mathbf{G} \subset \mathrm{GL}_n(\mathbb{C})$ es un grupo algebraico y $A$ es un subanillo de $\mathbb{C}$, denotaremos $\mathbf{G}_A = \mathbf{G} \cap \mathrm{GL}_n(A)$.

Un $\mathbb{K}$-*morfismo* de $\mathbb{K}$-grupos $\varphi : \mathbf{G}_1 \to \mathbf{G}_2$ es un $\mathbb{K}$-morfismo de $\mathbb{K}$-variedades que es también un morfismo de grupos. Un $\mathbb{K}$-morfismo de la forma $\rho : \mathbf{G} \to \mathrm{GL}(V)$ es una $\mathbb{K}$-*representación* de $\mathbf{G}$.

*Nota* 3.4. Si $\mathbf{G} \subseteq \mathrm{GL}_n(\mathbb{C})$ es un $\mathbb{K}$-grupo, $\mathbf{G}_{\mathbb{R}} := \mathbf{G} \cap \mathrm{GL}_n(\mathbb{R})$ es un grupo de Lie con un número finito de componentes conexas. Diremos que un grupo algebraico $\mathbf{G}$ es *semisimple* si el álgebra de Lie del grupo de Lie $\mathbf{G}_{\mathbb{R}}$ es semisimple o equivalentemente, si el cubrimiento universal es un producto de grupos simples.

Terminaremos esta subsección con algunos ejemplos de grupos algebraicos definidos sobre $\mathbb{K}$. En lo que sigue, dada una matriz $A$, denotaremos con $a_{i,j}$ a sus entradas.

**Ejemplo 3.5.** Como ya mencionamos, $\mathrm{GL}(V)$ es un $\mathbb{Q}$-grupo algebraico. De manera similar, $\mathrm{SL}(V)$ también lo es ya que el polinomio que lo define, $\det(T) - 1 = 0$, tiene coeficientes racionales.

**Ejemplo 3.6.** Sea $J$ una matriz simétrica real, $n \times n$, no degenerada, con coeficientes en $\mathbb{K}$. Entonces,

$$\mathrm{O}(J) := \{A \in \mathcal{M}_d(\mathbb{C}) : A^t J A = J\}$$

es un grupo algebraico definido sobre $\mathbb{K}$. En efecto, el ideal de polinomios que lo define está generado por las coordenadas de la matriz $A^tJA - J$, las cuales pueden verse como polinomios en las entradas $a_{i,j}$ de $A$ con coeficientes en $\mathbb{K}$.

Sea $\mathrm{O}(J)_{\mathbb{R}}$, el grupo de transformaciones lineales de $\mathbb{R}^n$ que preservan la forma cuadrática $F(x) := x^tJx$ para todo $x = (x_1, \ldots, x_n)^t \in \mathbb{R}^n$, es decir, $\mathrm{O}(J)_{\mathbb{R}} = \{g \in \mathrm{GL}_n(\mathbb{R}) : F(gx) = F(x),\ \forall\, x \in \mathbb{R}^n\}$.

Claramente el subgrupo $\mathrm{SO}(J) := \mathrm{O}(J) \cap \mathrm{SL}_n(\mathbb{C})$ es también un $\mathbb{Q}$-grupo algebraico.

**Ejemplo 3.7.** Como caso particular del ejemplo anterior, tomemos $n = p + q$, y

$$J = \mathrm{diag}(r_1, \ldots, r_p, -s_1, \cdots - s_q),$$

con $r_i, s_j$ racionales positivos para todo $i$ y $j$. Entonces, $\mathrm{O}(J)_{\mathbb{R}}$ es el grupo transformaciones lineales de $\mathbb{R}^n$ que preservan la forma cuadrática

$$F(x) := x^tJx = r_1x_1^2 + \cdots + r_px_p^2 - s_1x_{p+1}^2 - \cdots - s_qx_{p+q}^2.$$

Entre los polinomios que definen $\mathrm{O}(J)$ tenemos por ejemplo al inducido por la ecuación determinada por la coordenada $(1,1)$: $P_{1,1}(A) = P_{1,1}(a_{1,1}, a_{1,2}, \ldots, a_{d,d}) = r_1a_{1,1}^2 + \cdots + r_pa_{p,1}^2 - s_1a_{p+1,1}^2 - \cdots - s_qa_{p+q,1}^2 - 1$.

*Nota* 3.8. Para

$$I_{p,q} := \mathrm{diag}(\underbrace{1, \ldots, 1}_{p\text{-veces}}, \underbrace{-1, \ldots, -1}_{q\text{-veces}}),$$

denotamos $\mathrm{O}(p,q) = \mathrm{O}(I_{p,q})$ y $\mathrm{SO}(p,q) = \mathrm{SO}(I_{p,q})$.

Se puede ver que para cualquier matriz simétrica real $J$ no degenerada $n \times n$, existen enteros $p$ y $q$, con $n = p+q$, tales que los grupos $\mathrm{O}(J)$ y $\mathrm{O}(p,q)$ son conjugados (ver Ejercicio 3.6). En particular, ellos son isomorfos como grupos algebraicos (sobre $\mathbb{C}$). Además $\mathrm{O}(J)_{\mathbb{R}}$ y $\mathrm{O}(p,q)_{\mathbb{R}}$ son conjugados, por lo que son isomorfos como grupos de Lie. A pesar de esto, $\mathrm{O}(J)$ y $\mathrm{O}(p,q)$ son en general diferentes como $\mathbb{Q}$-grupos algebraicos, y definen grupos aritméticos esencialmente distintos.

**3.2. Subgrupos aritméticos.** Sea $\mathbf{G}$ un subgrupo algebraico de $\mathrm{GL}(V)$ definido sobre $\mathbb{Q}$, donde $V$ es un $\mathbb{C}$-espacio vectorial munido de una $\mathbb{Q}$-estructura $V_{\mathbb{Q}}$. Sea $L$ un retículo contenido en $V_{\mathbb{Q}}$, i.e., un $\mathbb{Z}$-submódulo libre de $V_{\mathbb{Q}}$, generado por una $\mathbb{Q}$-base de $V$. [1]

Denotamos $\mathbf{G}_{\mathbb{Q}} = \{g \in \mathbf{G} : gV_{\mathbb{Q}} = V_{\mathbb{Q}}\}$, el grupo de puntos $\mathbb{Q}$-racionales de $\mathbf{G}$, y para $L$ un retículo en $V_{\mathbb{Q}}$, sea

$$\mathbf{G}_L := \{g \in \mathbf{G}_{\mathbb{Q}} : g(L) = L\}.$$

**Definición 3.9.** Sea $\mathbf{G} \subseteq \mathrm{GL}(V)$ un subgrupo algebraico definido sobre $\mathbb{Q}$. Un subgrupo *aritmético* de $\mathbf{G}$ es un subgrupo $\Gamma$ de $\mathbf{G}_{\mathbb{Q}}$ que es conmensurable con $\mathbf{G}_L$ para algún retículo $L$ en $V_{\mathbb{Q}}$.

**Ejemplo 3.10.** Por ejemplo, en el caso $\mathbf{G} \subset \mathrm{GL}_n(\mathbb{C})$ (i.e. $V = \mathbb{C}^n$ y $V_{\mathbb{Q}} = \mathbb{Q}^n$), tomando el retículo $L = \mathbb{Z}^n$ obtenemos $\mathbf{G}_L = \mathbf{G}_{\mathbb{Z}} = \mathbf{G} \cap \mathrm{GL}_n(\mathbb{Z})$.

Más aún, si $L$ es cualquier retículo en $\mathbb{Q}^n$ con $\mathbb{Z}$-base $\mathcal{B}$, y si $C \in \mathrm{GL}_n(\mathbb{Q})$ es la matriz de cambio de base de $\mathcal{B}$ a la base canónica $\mathcal{C}$, entonces se tiene que $\mathbf{G}_L = C^{-1}\mathbf{G}_{\mathbb{Z}}C$, y se puede verificar que $\mathbf{G}_L \subset \mathbf{G}_{\mathbb{Q}}$ es conmensurable con $\mathbf{G}_{\mathbb{Z}} \subset \mathrm{GL}_n(\mathbb{Z})$, pero $\mathbf{G}_L$ no es necesariamente un subgrupo de $\mathrm{GL}_n(\mathbb{Z})$.

[1] Usar la palabra *retículo* para $L$ en $V_{\mathbb{Q}}$ es un abuso del lenguaje muy útil. Asimismo, $L$ es un retículo en el sentido usual en el $\mathbb{R}$-espacio vectorial $V_{\mathbb{Q}} \otimes_{\mathbb{Q}} \mathbb{R}$.

*Nota* 3.11. Mencionamos en la Nota 3.3 que, tomando cualquier $\mathbb{Q}$-base de $V_{\mathbb{Q}}$, todo $\mathbb{Q}$-grupo algebraico $\mathbf{G} \subset \mathrm{GL}(V)$ puede realizarse dentro de $\mathrm{GL}_n(\mathbb{C})$ con respecto a la $\mathbb{Q}$-forma estándar $\mathbb{Q}^n$ en $\mathbb{C}^n$. Esto nos permitirá asumir (sin perder generalidad) en algunos de los siguientes enunciados que los grupos algebraicos sobre $\mathbb{Q}$ están contenidos en $\mathrm{GL}_n(\mathbb{C})$ (con respecto a la $\mathbb{Q}$-forma $\mathbb{Q}^n$ en $\mathbb{C}^n$).

**Definición 3.12.** Sea $\mathbf{G} \subset \mathrm{GL}_n(\mathbb{C})$ un $\mathbb{Q}$-subgrupo algebraico. Para $N \in \mathbb{N}$, definimos el subgrupo principal de congruencia de nivel $N$ como

$$\mathbf{G}_{\mathbb{Z}}(N) = \{g \in \mathbf{G}_{\mathbb{Z}} : g \equiv \mathrm{Id} \pmod{N}\}.$$

El siguiente resultado nos muestra que $\mathbf{G}_{\mathbb{Z}}$ no depende, salvo conmensurabilidad, de la realización de $\mathbf{G}$ como grupo algebraico definido sobre $\mathbb{Q}$.

**Proposición 3.13.** *Sea $\boldsymbol{G} \subseteq \mathrm{GL}_n(\mathbb{C})$ un subgrupo algebraico definido sobre $\mathbb{Q}$, y sea $\rho : \boldsymbol{G} \to \mathrm{GL}(V)$ un morfismo de $\mathbb{Q}$-grupos algebraicos.*

(i) *Si $L$ es un retículo en $V_{\mathbb{Q}}$, entonces existe $N \in \mathbb{N}$ tal que $\rho(\boldsymbol{G}_{\mathbb{Z}}(N)) \cdot L = L$.*
(ii) *$\rho(\boldsymbol{G}_{\mathbb{Z}})$ preserva algún retículo de $V_{\mathbb{Q}}$.*

*Demostración.* Sea $L$ un retículo en $V_{\mathbb{Q}}$, y fijemos una $\mathbb{Z}$-base de $L$. Sea $m$ la dimensión de $V$. Para cada $g \in \mathbf{G}$, sea $[\rho_{i,j}(g)]_{i,j}$ la matriz de $\rho(g) : V \to V$ con respecto a la base fijada. Se tiene que $\rho(g)_{i,j} \in \mathbb{Q}[g] = \mathbb{Q}[g_{1,1}, g_{1,2}, \ldots, g_{m,m}]$ por estar $\rho$ definida sobre $\mathbb{Q}$.

Para cada $1 \leq i, j \leq m$, consideremos

$$P_{i,j}(g - \mathrm{Id}) := \rho_{i,j}(g) - \delta_{i,j}.$$

Claramente $P_{i,j}$ es un polinomio con coeficientes racionales. La razón de tomar como variables las coordenadas de $g - \mathrm{Id} = (g_{k,l} - \delta_{k,l})_{k,l}$ en lugar de $g$ es para obtener un polinomio sin término constante. En efecto, $P_{i,j}(0) = \rho_{i,j}(\mathrm{Id}) - \delta_{i,j} = 0$ para todo $i, j$ ya que $\rho(\mathrm{Id}) = \mathrm{Id}$.

Sea $N \in \mathbb{N}$ elegido de modo que el polinomio $NP_{i,j}$ tenga coeficientes enteros para todo $i, j$. Si $g \in \mathbf{G}_{\mathbb{Z}}(N)$, entonces $g_{k,l} - \delta_{k,l} \equiv 0 \pmod{N}$ para todo $1 \leq k, l \leq m$. Como $P_{i,j}$ no tiene término constante, obtenemos que $P_{i,j}(g - \mathrm{Id}) \in \mathbb{Z}$ para todo $i, j$. Luego $\rho_{i,j}(g) \in \mathbb{Z}$ para todo $g \in \mathbf{G}_{\mathbb{Z}}(N)$, lo que significa que $\rho(g) \cdot L \subset L$. Como $\mathbf{G}_{\mathbb{Z}}(N)$ es un grupo, $\rho(g) \cdot L = L$ para todo $g \in \mathbf{G}_{\mathbb{Z}}(N)$. Esto prueba (i).

Para la segunda afirmación, tomamos cualquier retículo $L$ en $V_{\mathbb{Q}}$. Por (i), existe un subgrupo de índice finito $\Gamma$ de $\mathbf{G}_{\mathbb{Z}}$ que deja estable a $L$. Sean $\{\gamma_1, \ldots, \gamma_r\}$ representantes de $\mathbf{G}_{\mathbb{Z}}/\Gamma$, y definimos

$$L' = \sum_{k=1}^{r} \rho(\gamma_k)(L).$$

Se puede ver que $L'$ es un retículo en $V_{\mathbb{Q}}$ (Ejercicio 3.8) y además, claramente $L'$ es invariante por $\rho(\mathbf{G}_{\mathbb{Z}})$, lo que prueba (ii). □

**Proposición 3.14.** *Sea $\boldsymbol{G}$ un subgrupo algebraico de $\mathrm{GL}(V)$ definido sobre $\mathbb{Q}$. Si $L$ y $L'$ son dos retículos en $V_{\mathbb{Q}}$, entonces $\boldsymbol{G}_L$ y $\boldsymbol{G}_{L'}$ son conmensurables.*

*Demostración.* Por la Proposición 3.13 (i) con $\rho$ la identidad, existe un subgrupo $\Gamma$ de índice finito en $\mathbf{G}_{\mathbb{Z}}$ que deja invariante a $L'$. Por lo tanto $\Gamma \subset \mathbf{G}_{L'}$, luego $\Gamma \subset \mathbf{G}_L \cap \mathbf{G}_{L'} \subseteq \mathbf{G}_L$, lo que implica que $\mathbf{G}_L \cap \mathbf{G}_{L'}$ es de índice finito en $\mathbf{G}_L$. Intercambiando $L$ con $L'$ se ve que $\mathbf{G}_L \cap \mathbf{G}_{L'}$ es de índice finito también en $\mathbf{G}_{L'}$. □

*Nota* 3.15. A raíz de la proposición anterior, concluimos que todos los subgrupos aritméticos de un $\mathbb{Q}$-grupo algebraico fijo $\mathbf{G} \subset \mathrm{GL}(V)$ son conmensurables. En efecto, cada uno de ellos es conmensurable a $\mathbf{G}_L$ para cualquier $L$ retículo de $V_\mathbb{Q}$, y la conmensurabilidad es transitiva (ver Ejercicio 2.7). En particular, todos los subgrupos aritméticos serán cocompactos o no cocompactos simultáneamente.

**Corolario 3.16.** *Sea $\varphi : \boldsymbol{G} \to \boldsymbol{G}'$ un $\mathbb{Q}$-isomorfismo de $\mathbb{Q}$-grupos algebraicos. Si $\Gamma$ es un subgrupo aritmético de $\boldsymbol{G}$, entonces $\varphi(\Gamma)$ es un subgrupo aritmético de $\boldsymbol{G}'$.*

El siguiente teorema es uno de los resultados fundamentales sobre grupos aritméticos.

**Teorema 3.17.** *Sea $\boldsymbol{G} \subset \mathrm{GL}_n(\mathbb{C})$ un grupo algebraico semisimple definido sobre $\mathbb{Q}$. Entonces $\boldsymbol{G}_\mathbb{R}/\boldsymbol{G}_\mathbb{Z}$ tiene medida de Haar finita.*

En particular, este resultado asegura que todo subgrupo aritmético en $\mathbf{G}$ es un retículo en $\mathbf{G}_\mathbb{R}$. Su demostración, debida a Borel y Harish-Chandra [3], es muy técnica y extensa. En la siguiente sección, demostraremos el caso particular de $\mathrm{SL}_n(\mathbb{Z})$ como retículo de $\mathrm{SL}_n(\mathbb{R})$ usando los llamados dominios de Siegel. El caso general se basa en una generalización de estos dominios a grupos algebraicos arbitrarios.

**Ejemplo 3.18.** Sea $J = \mathrm{diag}(2, 3, -1)$ y definamos $B : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ la forma bilineal dada por $B(x, y) = 2x_1y_1 + 3x_2y_2 - x_3y_3$ para $x = (x_1, x_2, x_3)^t, y = (y_1, y_2, y_3)^t \in \mathbb{R}^3$. Claramente $B$ es no degenerada, aunque no es definida positiva sino que es indefinida de signatura $(2, 1)$. El grupo $\mathrm{O}(J)_\mathbb{Z}$ está dado por matrices enteras $g$ cuyas columnas $g_1, g_2, g_3 \in \mathbb{Z}^3$ cumplen $B(g_1, g_1) = 2$, $B(g_2, g_2) = 3$, $B(g_3, g_3) = -1$, y $B(g_i, g_j) = 0$ para $i \neq j$.

Observamos que no es fácil mostrar que $\mathrm{O}(J)_\mathbb{Z}$ tiene una cantidad infinita de elementos (cf. Ejercicio 3.10). El Teorema 3.17 nos dice que $\mathrm{O}(J)_\mathbb{Z}$ es un retículo en el grupo de Lie no compacto $\mathrm{O}(J)_\mathbb{R}$, lo cual implica en particular que $\mathrm{O}(J)_\mathbb{Z}$ debe tener una cantidad infinita de puntos (Ejercicio 2.5).

**3.3. Ejemplos de grupos aritméticos.** Terminamos esta sección con una serie de ejemplos de grupos aritméticos.

**Ejemplo 3.19.** Como ya hemos visto, $\mathrm{SL}_n(\mathbb{Z})$ (y cualquier subgrupo de $\mathrm{SL}_n(\mathbb{Q})$ conmensurable a él) es un subgrupo aritmético de $\mathbf{G} = \mathrm{SL}_n(\mathbb{C})$. Por lo tanto, $\mathrm{SL}_n(\mathbb{Z})$ es un retículo de $\mathbf{G}_\mathbb{R} = \mathrm{SL}_n(\mathbb{R})$. En la próxima sección mostraremos que no es cocompacto.

**Ejemplo 3.20.** De manera similar al ejemplo anterior, se tiene que $\mathrm{SL}_2(\mathbb{Z}[\sqrt{-1}])$ es un retículo en el grupo de Lie (real) $\mathrm{SL}_2(\mathbb{C})$. Para ello, es necesario exhibir un $\mathbb{Q}$-grupo algebraico $\mathbf{G}$ cuyo conjunto de puntos reales sea $\mathbf{G}_\mathbb{R} \simeq \mathrm{SL}_2(\mathbb{C})$. Lo mismo vale cuando reemplazamos los enteros de Gauss, $\mathbb{Z}[\sqrt{-1}]$, por los enteros algebraicos $\mathcal{O}_\mathbb{K}$ de cualquier extensión cuadrática imaginaria $\mathbb{K}$ de $\mathbb{Q}$. Estos grupos aritméticos son conocidos como *grupos de Bianchi*.

**Ejemplo 3.21.** Ahora demostraremos que $\mathrm{SL}_2(\mathbb{Z}[\sqrt{2}])$ es un grupo aritmético, es decir, construiremos un $\mathbb{Q}$-grupo algebraico $\mathbf{G} \subset \mathrm{GL}(V)$ y un retículo en $L$ en $V_\mathbb{Q}$ tales que $\mathbf{G}_L$ es isomorfo a $\mathrm{SL}_2([\sqrt{2}])$. Notar que $\mathbf{G}_\mathbb{R}$ no puede ser $\mathrm{SL}_2(\mathbb{R})$, ya que $\mathrm{SL}_2(\mathbb{Z}[\sqrt{2}])$ ni siquiera es discreto con la topología relativa (Ejercicio 3.11).

Sea $\mathbb{Z}[\sqrt{2}] = \mathbb{Z} + \sqrt{2}\mathbb{Z}$. Tomemos en $V := \mathbb{C}^2 \times \mathbb{C}^2$ la $\mathbb{Q}$-forma dada por

$$V_\mathbb{Q} = \{(v + \sqrt{2}w, v - \sqrt{2}w) : v, w \in \mathbb{Q}^2\}$$

y sea

$$\mathbf{G} = \left\{ g = \left(\begin{smallmatrix} g_1 & g_2 \\ g_3 & g_4 \end{smallmatrix}\right) \in \mathrm{GL}(\mathbb{C}^2 \times \mathbb{C}^2) : \det(g_1) = \det(g_4) = 1,\ g_2 = g_3 = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right) \right\},$$

el cual claramente es un grupo algebraico definido sobre $\mathbb{Q}$. Además, es isomorfo (como grupo abstracto) a $\mathrm{SL}_2(\mathbb{C}) \times \mathrm{SL}_2(\mathbb{C})$, y $\mathbf{G}_{\mathbb{R}} \simeq \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ como grupos de Lie.

Sea

$$L = \{(v + \sqrt{2}w, v - \sqrt{2}w) : v, w \in \mathbb{Z}^2\}.$$

Resta ver que $\mathrm{SL}_2(\mathbb{Z}[\sqrt{2}])$ puede ser identificado con $\mathbf{G}_L$. En efecto, si $\sigma : \mathbb{Q}[\sqrt{2}] \to \mathbb{C}$ denota la incrustación no trivial (i.e. $\sigma(a + \sqrt{2}b) = a - \sqrt{2}b$ para $a, b \in \mathbb{Q}$), y la extendemos a las matrices $2 \times 2$, entonces

$$g \mapsto \begin{pmatrix} g & 0 \\ 0 & \sigma(g) \end{pmatrix}$$

identifica $\mathrm{SL}_2(\mathbb{Z}[\sqrt{2}])$ con $\mathbf{G}_L$.

Los tres ejemplos anteriores pueden ser enmarcados dentro de un método general llamado *restricción de escalares* que ahora describiremos sin precisar mayores detalles (ver [17, §5E]). Sea $\mathbb{K}$ un cuerpo de números (i.e. una extensión finita de $\mathbb{Q}$) con $r$ incrustaciones reales y $2s$ incrustaciones complejas, y sea $\mathbf{G} \subset \mathrm{GL}_n(\mathbb{C})$ un grupo algebraico definido sobre $\mathbb{K}$. Entonces existe un $\mathbb{Q}$-grupo algebraico $\mathbf{G}' \subset \mathrm{GL}_m(\mathbb{C})$, con $m = n(r+2s)$, tal que $\mathbf{G}_{\mathcal{O}_{\mathbb{K}}} = \mathbf{G} \cap \mathrm{GL}_n(\mathcal{O}_{\mathbb{K}})$ se identifica de una manera natural con $\mathbf{G}'_{\mathbb{Z}}$ usando las incrustaciones de $\mathbb{K}$ a $\mathbb{C}$. Más aún, $\mathbf{G}'_{\mathbb{R}} \subset \mathrm{GL}_n(\mathbb{R})^r \times \mathrm{GL}_n(\mathbb{C})^s$.

Un buen ejercicio es intuir cómo definir $\mathbf{G}'$, en términos de una familia de polinomios con coeficientes en $\mathbb{K}$ que definen $\mathbf{G}$.

En la siguiente sección estudiaremos condiciones para que el subgrupo aritmético $\mathrm{O}(J)_{\mathbb{Z}}$ como en el Ejemplo 3.6 sea cocompacto en $\mathrm{O}(J)_{\mathbb{R}}$. Otro método muy usado para construir subgrupos aritméticos $\Gamma$ en $\mathrm{SL}_2(\mathbb{R})$ y $\mathrm{SL}_2(\mathbb{C})$ utiliza las álgebras de cuaterniones. Se darán muchos más detalles en las últimas secciones, donde propiedades aritméticas de estas álgebras se relacionarán con propiedades geométricas de los cocientes $H^2/\Gamma$ y $H^3/\Gamma$ respectivamente, donde $H^n$ denota el espacio hiperbólico de dimensión $n$.

### 3.4. Ejercicios.

**Ejercicio 3.1.** Sea $V$ un subespacio vectorial de $\mathbb{R}^n$. Probar que las siguientes afirmaciones son equivalentes:

1. $V \cap \mathbb{Z}^n$ es un retículo cocompacto en $V$.
2. $V$ es generado por $V \cap \mathbb{Z}^n$.
3. $V \cap \mathbb{Q}^n$ es denso en $V$.
4. $V$ puede ser definido por un conjunto de ecuaciones lineales con coeficientes racionales

**Ejercicio 3.2.** Sea $V_{\mathbb{K}}$ una $\mathbb{K}$-forma en el $\mathbb{C}$-espacio vectorial $V$. Mostrar que $\mathrm{End}(V)_{\mathbb{K}} := \{T \in \mathrm{End}_{\mathbb{C}}(V) : T(V_{\mathbb{K}}) \subset V_{\mathbb{K}}\}$ es una $\mathbb{K}$-forma de $\mathrm{End}_{\mathbb{C}}(V)$, y que $\mathrm{End}(V)_{\mathbb{K}} \simeq \mathrm{End}_{\mathbb{K}}(V_{\mathbb{K}})$ como $\mathbb{K}$-espacio vectorial.

**Ejercicio 3.3.** Sea $\mathbb{K}$ un subcuerpo de $\mathbb{C}$, sea $V_{\mathbb{K}}$ una $\mathbb{K}$-forma del espacio vectorial $V$ sobre $\mathbb{C}$ con base $\{v_1, \ldots, v_n\}$. Si $P : V \to \mathbb{C}$ es una función polinomial, entonces

$$P(a_1 v_1 + \ldots + a_n v_n) = \sum_{\alpha=(i_1,\ldots,i_n)} c_\alpha a_1^{i_1} \ldots a_n^{i_n},$$

con $c_\alpha \in \mathbb{C}$ todos nulos salvo una cantidad finita. Mostrar que $P(v) \in \mathbb{K}$ para todo $v \in V_{\mathbb{K}}$ si y sólo si $c_\alpha \in \mathbb{K}$ para todo $\alpha$.

**Ejercicio 3.4.** Sea $Z$ una variedad algebraica. Demostrar que $Z$ está definida sobre $\mathbb{K}$ si y sólo si $Z$ es el conjunto de ceros de una familia de polinomios en $\mathbb{K}[V]$.

**Ejercicio 3.5.** Sea $\mathbb{K}$ un subcuerpo de $\mathbb{C}$. Demostrar que si $\mathbf{G} \subset \mathrm{GL}_m(\mathbb{C})$ y $\mathbf{H} \subset \mathrm{GL}_m(\mathbb{C})$ son grupos algebraicos sobre $\mathbb{K}$, entonces $\mathbf{G} \times \mathbf{H}$ también lo es.

**Ejercicio 3.6.** Probar que los grupos $\mathrm{O}(J)$ y $\mathrm{O}(p,q)$, como en la Nota 3.8, son conjugados. Para $J$ como en el Ejemplo 3.7, dar explícitamente la matriz que conjuga $\mathrm{O}(J)$ en $\mathrm{O}(p,q)$. Notar que en general sus coeficientes no son racionales, por lo que la conjugación por ella no es un $\mathbb{Q}$-morfismo.

**Ejercicio 3.7.** Demostrar que el subgrupo principal de congruencia $\mathbf{G}_{\mathbb{Z}}(N)$ de nivel $N$ (ver Definición 3.12) tiene índice finito en $\mathbf{G}_{\mathbb{Z}}$.

**Ejercicio 3.8.** Probar que $L'$ es un retículo en la demostración de la Proposición 3.13. Ayuda: Es suficiente demostrar que es discreto, para lo cual basta encontrar $m \in \mathbb{Z}$ tal que $mL' \subset L$.

**Ejercicio 3.9.** Demostrar el Corolario 3.16.

**Ejercicio 3.10.** Demostrar (sin usar el Teorema 3.17) que $\mathrm{O}(J)_{\mathbb{Z}}$, como en el Ejemplo 3.18, tiene infinitos elementos.

**Ejercicio 3.11.** Demostrar que $\mathrm{SL}_2(\mathbb{Z}[\sqrt{2}])$ no es discreto en $\mathrm{SL}_2(\mathbb{R})$.

**Ejercicio 3.12.** Mostrar que una matriz real triangular superior con entradas diagonales positivas que a su vez es ortogonal, es necesariamente la matriz identidad.

**Ejercicio 3.13.** Sea $L$ un retículo en $\mathbb{R}^n$ y sean $\{v_1, \ldots, v_n\}$ y $\{w_1, \ldots, w_n\}$ dos $\mathbb{Z}$-bases de $L$. Demostrar que $\det(v_1, \ldots, v_n) = \pm \det(w_1, \ldots, w_n)$.

## 4. Dominios de Siegel

El objetivo de esta sección es construir ciertos dominios fundamentales aproximados para la acción de $\mathrm{GL}_n(\mathbb{Z})$ en $\mathrm{GL}_n(\mathbb{R})$ llamados conjuntos de Siegel.

Borel y Harish-Chandra [3] generalizaron esta construcción a todo subgrupo aritmético de un grupo algebraico semisimple $\mathbf{G}$ definido sobre los números racionales, demostrando que $\mathbf{G}_{\mathbb{Z}}$ tiene covolumen finito en el grupo de Lie $\mathbf{G}_{\mathbb{R}}$ (Teorema 3.17), es decir, es un retículo en $\mathbf{G}_{\mathbb{R}}$.

Los conjuntos de Siegel tiene aplicaciones a la teoría de reducción de formas cuadráticas, tal como veremos en la Subsección 4.2.

**4.1. Descomposición de Iwasawa.** El resultado siguiente describe la llamada *descomposición de Iwasawa* $G = KAN$ del grupo $G = \mathrm{GL}_n(\mathbb{R})$. La misma descomposición es válida para todo grupo de Lie semisimple real, pero daremos la prueba sólo en el presente caso por simplicidad.

**Proposición 4.1.** *Sean $G = \mathrm{GL}_n(\mathbb{R})$, $K = \mathrm{O}(n)$ el subgrupo de matrices ortogonales, $A$ el subgrupo de matrices diagonales con entradas positivas, y $N$ el subgrupo de matrices unipotentes triangulares superiores. Entonces la función $\Phi : K \times A \times N \to G$, dada por $\Phi(k,a,n) = kan$ es un difeomorfismo.*

*Demostración.* Veamos primero la inyectividad de $\Phi$. Si $g = kan = k'a'n'$ con $k, k' \in \mathrm{O}(n)$, $a, a' \in A$ y $n, n' \in N$, entonces $k'^{-1}k = a'n'n^{-1}a^{-1}$ es a la vez ortogonal y triangular superior con entradas diagonales positivas. Esto implica que $k'^{-1}k = a'n'n^{-1}a^{-1} = \mathrm{Id}$. Luego $an = a'n'$, o sea $a'^{-1}a = n'n^{-1}$. Como $n'n^{-1}$ es diagonal con 1's en la diagonal se deduce que $a' = a$, $n' = n$ y ya habíamos visto que también $k' = k$.

Ahora definiremos la inversa de $\Phi$. Sea $g \in \mathrm{GL}_n(\mathbb{R})$ y sean $v_i = g^{-1}e_i$ donde $e_1, \ldots, e_n$ es la base canónica de $\mathbb{R}^n$. Aplicamos el método de Gram-Schmidt a $v_1, \ldots, v_n$. Definamos $u_1 = \|v_1\|^{-1}v_1$ e inductivamente, sean

$$u_{k+1} = \left\| v_{k+1} - \sum_{j=1}^{k} \langle v_{k+1}, u_j \rangle u_j \right\|^{-1} \left( v_{k+1} - \sum_{j=1}^{k} \langle v_{k+1}, u_j \rangle u_j \right).$$

Entonces $u_i = \sum_{j \le i} a_{ji} v_j$, donde la matriz $a_{ij}$ es triangular superior y $a_{ii} > 0$ para todo $i$. Esto define $a(g) \in A$ y $n(g) \in N$ tales que $(a_{ij}) = a(g)n(g)$ y $k(g) \in \mathrm{O}(n)$ de modo que para todo $i$, $a(g)n(g)g^{-1}e_i = k(g)^{-1}e_i$.

Se tiene así que $g = k(g)a(g)n(g)$, luego la función $\Phi$ es sobre y $\Psi : g \mapsto (k(g), a(g), n(g))$ es la inversa de $\Phi$. Claramente $\Phi$ y $\Psi$ son continuas, luego $\Phi$ es un homeomorfismo. Además puede comprobarse que $\Phi$ y $\Psi$ son diferenciables, pero omitiremos esta última verificación. □

**Definición 4.2.** Sean $t, u > 0$. Un conjunto de Siegel en $\mathrm{GL}_n(\mathbb{R})$ es un conjunto de la forma $\mathcal{S}_{t,u} = KA_tN_u$, donde

$$A_t = \{a \in A : a_{i,i} \le t a_{i+1,i+1} \quad i = 1, \ldots, n-1\},$$
$$N_u = \{n \in N : |n_{i,j}| \le u \quad 1 \le i < j \le n\}.$$

*Nota* 4.3. Recordemos que $N$ es difeomorfo a $\mathbb{R}^{n(n-1)/2}$ y $A \cong (\mathbb{R}^*)^n$. Observemos también que si $\omega \subset N$ es relativamente compacto, entonces el conjunto

$$\bigcup_{a \in A_t} a\omega a^{-1}$$

es también relativamente compacto. En efecto, si $n = (n_{i,j}) \in \omega$, entonces obtenemos $(ana^{-1})_{i,j} = \frac{a_{i,i}}{a_{j,j}} n_{i,j}$. Luego

$$|(ana^{-1})_{i,j}| \le t^{j-i} |n_{i,j}|$$

para todo $i < j$, $a \in A_t$, $n \in N$.

Para cada $a \in A$, sea $\sigma_a$ el automorfismo de $N$ dado por $\sigma_a(n) = ana^{-1}$. Entonces se verifica que $|\det(\sigma_{a|}\mathfrak{n})| = \prod_{i<j} a_{i,i}/a_{j,j}$.

El siguiente es uno de los principales resultados de la sección.

**Teorema 4.4.** *Sea $G = \mathrm{GL}_n(\mathbb{R})$ y $\Gamma = \mathrm{GL}_n(\mathbb{Z})$. Entonces $G = \mathcal{S}_{t,u}\Gamma$ para todo $t \ge 2/\sqrt{3}$ y $\|u\| \le 1/2$.*

*Demostración.* En primer lugar veamos que

$$N = N_{1/2}N_{\mathbb{Z}}, \tag{4.1}$$

donde $N_{\mathbb{Z}} = N \cap \mathrm{GL}_n(\mathbb{Z})$. En efecto, si $u \in N, z \in N_{\mathbb{Z}}$, entonces

$$(uz)_{i,j} = z_{i,j} + u_{i,i+1}z_{i+1,j} + \ldots + u_{i,j}$$

para $1 \le i < j \le n$, lo que permite definir $z_{i,j}$ por recurrencia desde $z_{n-1,n}$.

Definamos $\Phi : G \to \mathbb{R}^+$ dada por $\Phi(g) = \|ge_1\|$. Claramente $\Phi(kan) = \|ae_1\| = a_{1,1} = \Phi(a)$. Si $g \in G$ fijo, la función $z \mapsto \Phi(gz)$ con $z \in \Gamma$ tiene un mínimo positivo en $\Gamma$ pues $\{g\gamma e_1 : \gamma \in \Gamma\}$ es un subconjunto de elementos no nulos del retículo $g(\mathbb{Z}^n)$. Notar también que para $u \in N_{\mathbb{Z}}$ se verifica que $\Phi(gu) = \Phi(g)$ y $a_{gu} = a_g$.

**Afirmación 1.** *Si $\Phi(g) \leq \Phi(g\gamma)$ para todo $\gamma \in \Gamma$, entonces $a_{1,1} \leq (2/\sqrt{3})a_{2,2}$.*

*Demostración.* En efecto, sea $\sigma \in \Gamma$ que permuta $e_1$ y $e_2$ y fija los otros $e_i$. Entonces

$$g\sigma(e_1) = g(e_2) = ka(e_2 + n_{1,2}e_1) = k(a_{2,2}e_2 + a_{1,1}n_{1,2}e_1).$$

Luego $\|g\sigma(e_1)\|^2 = a_{2,2}^2 + a_{1,1}^2 n_{1,2}^2 \leq a_{1,1}^2/4 + a_{2,2}^2$. Entonces

$$\Phi(g)^2 = a_{1,1}^2 \leq a_{1,1}^2/4 + a_{2,2}^2,$$

lo que implica la afirmación. ■

**Afirmación 2.** *Si $g \in G$, el mínimo de $\Phi$ en $g\Gamma$ se alcanza en $g\Gamma \cap \mathcal{S}_{2/\sqrt{3},1/2}$.*

*Demostración.* Probaremos la afirmación por inducción. Si $n = 1$ no hay nada que probar.

Si $x \in G$, existe $y \in x\Gamma$ tal que $\Phi(y) \leq \Phi(x\gamma)$ para todo $\gamma \in \Gamma$. Fijemos tal $y$. Escribamos $k_y^{-1}y = \left[\begin{smallmatrix} a_{1,1} & * \\ 0 & b \end{smallmatrix}\right]$, con $b \in \mathrm{GL}_{n-1}(\mathbb{R})$. Entonces, por hipótesis inductiva, existe $z' \in \mathrm{GL}_{n-1}(\mathbb{Z})$ tal que $bz' \in \mathcal{S}^{n-1}_{2/\sqrt{3},1/2}$.

Escribamos $bz' = k'a'n'$. Luego

$$k_y{}^{-1}yz = \begin{bmatrix} a_{1,1} & * \\ 0 & k'a'n' \end{bmatrix} = k''a''n'',$$

con $z = \left[\begin{smallmatrix} 1 & 0 \\ 0 & z' \end{smallmatrix}\right]$, $k'' \in K$, $a'' = \left[\begin{smallmatrix} a_{1,1} & 0 \\ 0 & a' \end{smallmatrix}\right]$ y $n'' = \left[\begin{smallmatrix} 1 & * \\ 0 & n' \end{smallmatrix}\right] \in N$ y donde $a'_{i,i} \leq (2/\sqrt{3})a'_{i+1,i+1}$ para todo $i$. Además $\Phi(yz) = \Phi(y)$ y $\Phi(yz) \leq \Phi(yz\gamma)$ para todo $\gamma$.

Como además por la Afirmación 1 se tiene que $a''_{1,1} \leq (2/\sqrt{3})a''_{2,2}$, resulta que $yz \in KA_{2/\sqrt{3}}N$, luego usando (4.1), $x \in y\Gamma \subset KA_{2/\sqrt{3}}N_{1/2}\Gamma = \mathcal{S}_{2/\sqrt{3},1/2}\Gamma$, lo que prueba la afirmación. ■

Claramente, esto completa la demostración del teorema. □

En el caso de $G = \mathrm{SL}_n(\mathbb{R})$, la descomposición de Iwasawa es la misma que para $\mathrm{GL}_n(\mathbb{R})$, es decir $\mathrm{SL}_n(\mathbb{R}) = \mathrm{SO}(n)A^*N$, con $A^* = \mathrm{SL}_n(\mathbb{R}) \cap A$. El conjunto de Siegel $\mathcal{S}^*_{t,u}$ para $\mathrm{SL}_n(\mathbb{R})$ se define igual que en el caso de $\mathrm{GL}_n(\mathbb{R})$ y se tiene $S^*_{t,u} = \mathrm{SL}_n(\mathbb{R}) \cap S_{t,u}$.

El siguiente resultado nos dice que una medida de Haar invariante a izquierda en $G$ está dada por $\rho(a)dkdadn$ donde $\rho(a) = \prod_{i<j} a_{i,i}/a_{j,j}$.

**Proposición 4.5.** *Sean $G = \mathrm{GL}_n(\mathbb{R})$, $K$, $A$ y $N$ como en la Proposición 4.1, y sean $dk$, $da$, $dn$ medidas de Haar en $K$, $A$ y $N$ respectivamente. Entonces existe $C > 0$ tal que para toda $f \in C_c(G)$,*

$$\int_G f(g)\,dg = C \int_{K\times A\times N} f(kan) \left(\prod_{i<j} a_{i,i}/a_{j,j}\right) dk\,da\,dn.$$

*Demostración.* Por la fórmula del cambio de variables, existe una función suave $h(k, a, n)$ tal que

$$\int_G f(g)\,dg = \int_{K\times A\times N} f(kan)h(k, a, n)\,dk\,da\,dn.$$

Como $dg$ es invariante a izquierda y a derecha, $h$ es independiente de $k$ y $n$. Entonces podemos escribir $h(k,a,n) = h(a)$. Luego

$$\int_G f(g)\,dg = \int_{K\times A\times N} f(kan)h(a)\,dk\,da\,dn. \tag{4.2}$$

Ahora, si $a_0 \in A$,

$$\begin{aligned}\int_G f(g)\,dg &= \int_G f(ga_0)\,dg = \int_{K\times A\times N} f(kana_0)h(a)\,dk\,da\,dn \\ &= \int_{K\times A\times N} f(kan)h(aa_0^{-1})|\det(jac(n\to a_0 n a_0^{-1})\,dk\,da\,dn \\ &= \int_{K\times A\times N} f(kan)h(aa_0^{-1})\prod_{i<j} a_{0i,i}/a_{0j,j}\,dk\,da\,dn.\end{aligned} \tag{4.3}$$

Por lo tanto, de (4.2) y (4.3) resulta que

$$h(a) = h(aa_0^{-1})\prod_{i<j} a_{0i,i}/a_{0j,j}$$

para todo $a, a_0 \in A$. Tomando $a = a_0$ se tiene que $h(a_0) = h(e)\prod_{i<j} a_{0i,i}/a_{0j,j}$, lo que prueba la proposición. □

**Proposición 4.6.** *El volumen de un conjunto de Siegel $\mathcal{S}^*_{t,u}$ en $\mathrm{SL}_n(\mathbb{R})$ con respecto a la medida de Haar es finito.*

*Demostración.* Si $K^* = \mathrm{SO}(n)$ y $B^* = A^*N$, sean $dk$, $da$, $dn$ medidas de Haar en $K^*$, $A^*$ y $N$, todas biinvariantes ya que estos grupos son unimodulares.

De manera análoga a la Proposición 4.5, $dg = \rho(a)dkdadn$ es una medida de Haar en $G = \mathrm{SL}_n(\mathbb{R})$, con $\rho(a) = \prod_{i<j} a_{i,i}/a_{j,j}$. Además $\rho(a) = \prod_{1\le i\le n}(a_{i,i}/a_{i+1,i+1})^{r_i}$, con $r_i \in \mathbb{N}$.

Se tiene entonces que existe $C_u > 0$ tal que

$$\begin{aligned}(C_u)^{-1}\int_{\mathcal{S}_{t,u}} dg \le \int_{A_t}\rho(a)da &= \int_{A_t}\prod (a_{i,i}/a_{i+1,i+1})^{r_i}da \\ &= \prod_{1\le i<n}\int_{-\infty}^{\log(t)} (\exp\, r_i y_i)\,dy_i, = \prod_{1\le i<n} t^{r_i} < \infty,\end{aligned} \tag{4.4}$$

lo que completa la demostración. □

**4.2. Teoría de reducción de formas cuadráticas.** El siguiente resultado es consecuencia de Teorema 4.4.

**Corolario 4.7** (Hermite)**.** *Si $g \in G$, entonces*

$$\min_{\lambda\in\mathbb{Z}^n\smallsetminus\{0\}} \|g(\lambda)\| \le (2/\sqrt{3})^{(n-1)/2}\,|\det(g)|^{1/n}.$$

*Demostración.* Sea $g' \in g\Gamma \cap \mathcal{S}_{2/\sqrt{3},1/2}$. Se tiene

$$\min_{\lambda\in\mathbb{Z}^n\smallsetminus\{0\}} \|g(\lambda)\| \le \min_{\gamma\in\Gamma} \|g\gamma(e_1)\| = \|g'(e_1)\| = a'_{1,1},$$

donde $g' = k'a'n'$ es la descomposición de Iwasawa de $g'$. Luego

$$(a'_{1,1})^n \le (a'_{1,1})^{n-1}(2/\sqrt{3})\,a'_{2,2} \le (2/\sqrt{3})^{n(n-1)/2}\,a'_{1,1}a'_{2,2}\cdot\ldots\cdot a'_{n,n},$$

lo cual finaliza la prueba pues $\det(g) = a'_{1,1}a'_{2,2}\cdot\ldots\cdot a'_{n,n}$. □

*Nota* 4.8. El lector interesado en conjuntos de Siegel podrá encontrar en [12, 13, 18] investigaciones recientes sobre ellos, como así también algunas de sus aplicaciones en otras áreas de la matemática.

### 4.3. Ejercicios.

**Ejercicio 4.1.** Describir una descomposición de Iwasawa para $\mathrm{SO}(n,1)$. Más precisamente, encontrar un subgrupo compacto $K$, un subgrupo abeliano $A$, y un subgrupo nilpotente $N$ tal que $\Phi : K \times A \times N \to \mathrm{SO}(n,1)$, $\Phi(k,a,n) = kan$ es un difeomorfismo.

**Ejercicio 4.2.** Describir el conjunto de Siegel para el retículo $\mathrm{SL}_2(\mathbb{Z})$ de $\mathrm{SL}_2(\mathbb{R})$.

## 5. Criterio de compacidad

En esta sección daremos algunos criterios de compacidad de $G/\Gamma$, donde $\Gamma$ es un retículo en un grupo de Lie semisimple $G$.

**5.1. Espacio de retículos.** Denotaremos $\mathcal{L}$ el espacio de retículos de $\mathbb{R}^n$. Sea la aplicación $\mathrm{GL}_n(\mathbb{R}) \to \mathcal{L}$ dada por $g \mapsto g(\mathbb{Z}^n)$. Se ve fácilmente que esta función es suryectiva. Además, si $g(\mathbb{Z}^n) = h(\mathbb{Z}^n)$ para $g, h \in \mathrm{GL}_n(\mathbb{R})$, entonces $h^{-1}g(\mathbb{Z}^n) = \mathbb{Z}^n$ y por lo tanto $h^{-1}g \in \mathrm{GL}_n(\mathbb{Z})$ y en consecuencia tenemos la identificación $\mathrm{GL}_n(\mathbb{R})/\mathrm{GL}_n(\mathbb{Z}) \simeq \mathcal{L}$. Damos a $\mathcal{L}$ la topología inducida por esta identificación.

La siguiente proposición ayuda a comprender la topología de $\mathcal{L}$.

**Proposición 5.1.** *Una sucesión $\{L_m\}_{m=1}^{\infty}$ en $\mathcal{L}$ converge a $L_0$ si y sólo si existe una base $\mathcal{B}_m := \{v_1^{(m)}, \dots, v_n^{(m)}\}$ de $L_m$ para cada $m \in \mathbb{Z}_{\geq 0}$ tal que $v_i^{(m)} \to v_i^{(0)}$ para todo $1 \leq i \leq n$.*

*Demostración.* Denotamos $\pi : \mathrm{GL}_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})/\mathrm{GL}_n(\mathbb{Z})$ la proyección canónica; $\pi$ es abierta pues si $U$ es abierto, $\pi^{-1}\pi(U) = \bigcup_{A \in \mathrm{GL}_n(\mathbb{Z})} A \cdot U$ es abierto.

Sea $g_m \in \mathrm{GL}_n(\mathbb{Z})$ tal que $g_m(\mathbb{Z}^n) = L_m$ para todo $m \in \mathbb{Z}_{\geq 0}$. Entonces, $\pi(g_m) \to \pi(g_0)$. Si $U$ es un abierto de $\mathrm{GL}_n(\mathbb{R})$ que contiene a $g_0$, entonces existe $N \in \mathbb{N}$ tal que $\pi(g_m) \in \pi(U)$ para todo $m \geq N$. Entonces $g_m \in \pi^{-1}\pi(U)$, por lo tanto existen $h_m \in \mathrm{GL}_n(\mathbb{Z})$ tales que $h_m^{-1}g_m \in U$ para todo $m \geq N$, o equivalentemente, $h_m^{-1}g_m \to g_0$. Si $m \geq 0$, definimos $v_i^{(m)}$ como la $i$-ésima columna de la matriz $h_m^{-1}g_m$, con $h_0 = \mathrm{Id}$. Se tiene claramente que $v_i^{(m)} \to v_i^{(0)}$ para $1 \leq i \leq n$, luego las bases $\mathcal{B}_m = \{v_1^{(m)}, \dots, v_n^{(m)}\}$ tienen la propiedades requeridas.

Ahora consideremos la recíproca. Para cada $m \geq 0$ sea

$$g_m = \begin{pmatrix} | & & | \\ v_1^{(m)} & \cdots & v_n^{(m)} \\ | & & | \end{pmatrix}.$$

Como $v_i^{(m)} \to v_i^{(0)}$, se tiene que $g_m \to g_0$ en $\mathrm{GL}_n(\mathbb{R})$ y $\pi(g_m) \to \pi(g_0)$ en $\mathrm{GL}_n(\mathbb{R})/\mathrm{GL}_n(\mathbb{Z})$. Como $L_m = g_m\mathbb{Z}^n$, resulta que $L_m \to L_0$. □

Para cada $L \in \mathcal{L}$, definamos $\Delta(L) = |\det(v_1, \dots, v_n)|$ para cualquier base $\{v_1, \dots, v_n\}$ de $L$. Como dos bases difieren en un elemento $g$ de $\mathrm{GL}_n(\mathbb{Z})$ con determinante $\pm 1$, $\Delta(L)$ está bien definido (Ejercicio 3.13).

**Teorema 5.2** (criterio de Mahler)**.** *Dado un subconjunto $\mathcal{M} \subset \mathcal{L}$, las siguientes afirmaciones son equivalentes:*

(i) $\mathcal{M}$ *es relativamente compacto,*
(ii) $\Delta|_{\mathcal{M}}$ *está acotada y existe* $U$ *un entorno abierto de* $0$ *en* $\mathbb{R}^n$ *tal que* $L \cap U = \{0\}$ *para todo* $L \in \mathcal{M}$.

*Demostración.* Sea $\mathcal{S}_{t,u}$ un conjunto de Siegel mapeado sobre $\mathcal{L}$ por la aplicación $g \mapsto g(\mathbb{Z}^n)$. Es claro, por el Teorema 4.4, que (i) equivale a la existencia de $M' \subset S_{t,u}$ relativamente compacto tal que $M'(\mathbb{Z}^n) = \mathcal{M}$.

Por otra parte, por la definición de $S_{t,u}$, $M'$ es relativamente compacto si y sólo si las componentes $a_x$, $x \in M'$ forman un subconjunto relativamente compacto de $A$. Esto es, si y sólo si existen constantes $\alpha, \beta > 0$ tales que

$$\alpha \leq (a_g)_{i,i} \leq \beta, \quad g \in M' \tag{5.1}$$

para todo $1 \leq i \leq n$.

A su vez, (ii) equivale a decir que $g \mapsto |\det(g)|$ está acotado en $M'$ y existe $c > 0$ tal que $\|g(\lambda)\| \geq c$ para todo $g \in M'$ y $\lambda \in \mathbb{Z}^n$. Teniendo en cuenta estas consideraciones, veremos ahora que las condiciones (i) y (ii) son equivalentes.

Supongamos que (i) es cierta. Claramente $|\det(g)| < C$ para todo $g \in M'$.

Sea $\lambda \in \mathbb{Z}^n \smallsetminus \{0\}$, $\lambda = \sum_1^k m_i e_i$ con $m_k \neq 0$, $k \leq n$. Entonces $\|g(\lambda)\| = \|a_g n_g(\lambda)\|$ y la $k$-ésima coordenada de $a_g n_g(\lambda)$ es $(a_g)_{k,k} m_k$ lo cual implica que $\|g(\lambda)\| \geq \alpha$. Por lo tanto, (ii) es válida.

Asumiendo (ii), tenemos que $\|g(e_1)\| = (a_g)_{1,1} \geq c$. Como $a_g \in A_t$, esto implica que existe $\alpha > 0$ tal que $(a_g)_{i,i} \geq \alpha$ para todo $i$. Como el producto de los $(a_g)_{i,i}$ está acotado, se obtiene (5.1). □

**5.2. Criterios de compacidad.** En esta subsección daremos criterios de cocompacidad para subgrupos aritméticos. Para probar los resultados principales, serán de suma utilidad los siguientes lemas.

**Lema 5.3** (Jacobson-Morozov)**.** *Sea* $G$ *un grupo de Lie real conexo semisimple con centro finito. Para todo elemento unipotente* $u \in G$*, existe un homomorfismo continuo* $\varphi : \mathrm{SL}_2(\mathbb{R}) \to G$ *tal que*

$$\varphi\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\right) = u.$$

**Lema 5.4.** *Sea* $\boldsymbol{G} \subset \mathrm{GL}_n(\mathbb{C})$ *un subgrupo algebraico definido sobre* $\mathbb{Q}$ *sin* $\mathbb{Q}$*-caracteres no triviales (i.e. todo* $\mathbb{Q}$*-morfismo* $\chi$ *de* $\boldsymbol{G}$ *a* $\mathrm{GL}_1(\mathbb{C}) \simeq \mathbb{C}^\times$ *es trivial). Supongamos que existe un polinomio* $\boldsymbol{G}$*-invariante* $P \in \mathbb{Q}[x_1, \ldots, x_n]$ *tal que*

$$P(v) = 0, \;\; v \in \mathbb{Q}^n \iff v = 0.$$

*Entonces* $\boldsymbol{G}_{\mathbb{R}}/\boldsymbol{G}_{\mathbb{Z}}$ *es compacto.*

*Demostración.* Escribimos

$$P(x_1, \ldots, x_n) = \sum_{\alpha=(\alpha_1,\ldots,\alpha_d)} a_\alpha \, x_1^{\alpha_1} \ldots x_d^{\alpha_d},$$

con $a_\alpha \in \mathbb{Q}$ para todo $\alpha$. Como $P(0) = 0$, $a_{(0,\ldots,0)} = 0$, es decir, $P$ no tiene término constante. Sea $m \in \mathbb{N}$ tal que $m a_\alpha \in \mathbb{Z}$ para todo $\alpha$. Entonces $P(mx_1, \ldots, mx_n) \in \mathbb{Z}$ si $x_i \in \mathbb{Z}$ para todo $i$, o equivalentemente $P(m\mathbb{Z}^n) \subseteq \mathbb{Z}$. Para probar el lema es suficiente ver que $\mathbf{G}_{\mathbb{R}}/\mathbf{G}_{m\mathbb{Z}}$ es compacto por la conmensurabilidad de $\mathbf{G}_{\mathbb{Z}}$ y $\mathbf{G}_{m\mathbb{Z}}$ (Proposición 3.14).

Dado que $\mathcal{L} \simeq \mathrm{GL}_d(\mathbb{R})/\mathrm{GL}_d(\mathbb{Z})$, el cociente $\mathbf{G}_{\mathbb{R}}/\mathbf{G}_{\mathbb{Z}}$ se identifica con el conjunto $\mathcal{M} := \{g(\mathbb{Z}^n) : g \in \mathbf{G}_{\mathbb{R}}\}$. Veamos que $\mathcal{M}$ es compacto usando el Teorema 5.2.

Para esto debemos probar que existen $\alpha, \beta > 0$ tales que

$$(i)\ \Delta(g(\mathbb{Z}^n)) \leq \beta, \qquad (ii)\ \inf_{\lambda \in \mathbb{Z}^n - \{0\}} ||g(\lambda)|| \geq \alpha,\ \forall g \in \mathbf{G}_Z.$$

El $\mathbb{Q}$-carácter $\det : \mathbf{G} \to \mathbb{C}^\times$ es trivial por hipótesis, por lo que $\mathbf{G} \subset \mathrm{SL}_n(\mathbb{C})$. Luego, $(i)$ vale pues $\Delta(g(\mathbb{Z}^n)) = |\det g|$.

Para ver $(ii)$, supongamos que es falsa, entonces existen sucesiones $g_j \in \mathbf{G}_\mathbb{R}$ y $\lambda_j \in \mathbb{Z}^n \setminus \{0\}$ tales que $g_j(\lambda_j)$ converge a cero. Como $P$ es $\mathbf{G}$-invariante, tenemos que $|P(g_j(\lambda_j))| = |P(\lambda_j)| \to 0$. A la vez $P(\mathbb{Z}^n) \subseteq \mathbb{Z}$ y $P(\lambda) \neq 0$ para todo $\lambda \neq 0$, lo que implica que $|P(\lambda_j)| \geq 1$, que es una contradicción. □

**Teorema 5.5** (Criterio de compacidad de Godement)**.** *Sea $\boldsymbol{G} \subset \mathrm{GL}_n(\mathbb{C})$ un $\mathbb{Q}$-grupo algebraico semisimple. Entonces, $\boldsymbol{G}_\mathbb{R}/\boldsymbol{G}_\mathbb{Z}$ es compacto si y sólo si $\boldsymbol{G}_\mathbb{Z}$ no tiene elementos unipotentes distintos de la matriz identidad.*

*Demostración.* Supongamos que $\mathbf{G}_\mathbb{R}/\mathbf{G}_\mathbb{Z}$ es compacto. Sea $u \in \mathbf{G}_\mathbb{Q}$ un elemento unipotente. Por el Lema 5.3, existe un homomorfismo continuo $\varphi : \mathrm{SL}_2(\mathbb{R}) \to \mathbf{G}_\mathbb{R}$ con

$$\varphi\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\right) = u$$

Sea $a = \varphi\left(\begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}\right) \in \mathbf{G}_\mathbb{R}$. Entonces

$$a^{-k}ua^k = \varphi\left(\begin{bmatrix} 2^{-k} & 0 \\ 0 & 2^k \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2^k & 0 \\ 0 & 2^{-k} \end{bmatrix}\right)$$

$$= \varphi\left(\begin{bmatrix} 1 & 2^{-2k} \\ 0 & 1 \end{bmatrix}\right) \longrightarrow \varphi\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = e$$

si $k \to \infty$. Si $u \neq e$, entonces $a^{-k}ua^k \neq e$ para todo $k$, luego $\mathbf{G}_\mathbb{R}/\mathbf{G}_\mathbb{Z}$ no es compacto por el Teorema 2.7, lo cual es una contradicción. Por lo tanto $u = e$.

Ahora veamos la recíproca, razonando por el absurdo, es decir, supongamos que $\mathbf{G}_\mathbb{R}/\mathbf{G}_\mathbb{Z}$ no es compacto. Como $\mathbf{G}_\mathbb{Z}$ es un retículo en $\mathbf{G}_\mathbb{R}$ por el Teorema 3.17, el Teorema 2.7 implica que existen sucesiones $g_k \in \mathbf{G}_\mathbb{R}$ y $\gamma_k \in \Gamma$ tales que $g_k\gamma_k g_k^{-1} \to e$ si $k \to \infty$ y $g_k\gamma_k g_k^{-1} \neq e$ para todo $k \in \mathbb{N}$.

Ahora bien, esto implica que el polinomio característico de $g_k\gamma_k g_k^{-1}$, $\chi_{g_k\gamma_k g_k^{-1}}(x) = \chi_{\gamma_k}(x)$, tiende al polinomio característico de $e$, que es igual a $(x-1)^k$. Como $\gamma_k \in \mathbf{G}_\mathbb{Z} \subset \mathrm{GL}_n(\mathbb{Z})$, su polinomio característico $\chi_{\gamma_k}(x)$ tiene coeficientes enteros. Luego, si $\chi_{\gamma_k}(x) \to (x-1)^k$, existe $k_0$ tal que $\chi_{\gamma_k}(x) = (x-1)^k$, para todo $k \geq k_0$, luego $\mathbf{G}_\mathbb{Z}$ tiene elementos unipotentes no triviales ya que $\gamma_k \neq e$. □

**5.3. Ejemplos de subgrupos aritméticos cocompactos.** Como ya vimos en el Ejemplo 3.6, si $J$ es una matriz simétrica racional no degenerada, entonces $\mathrm{O}(J)_\mathbb{Z}$ es un retículo en el grupo de Lie $\mathrm{O}(J)_\mathbb{R}$. Aplicaremos el Teorema 5.5 para obtener un criterio de compacidad para $\mathrm{O}(J)_\mathbb{R}/\mathrm{O}(J)_\mathbb{Z}$. Denotemos $J[x] := x^tJx \in \mathbb{Q}$ si $x \in \mathbb{Q}^n$.

**Teorema 5.6.** *El cociente $\mathrm{O}(J)_\mathbb{R}/\mathrm{O}(J)_\mathbb{Z}$ es compacto si y sólo si $J[x] \neq 0$ para cualquier $x \in \mathbb{Q}^n$, $x \neq 0$ (i.e. la forma cuadrática asociada a $J$ no representa a cero en $\mathbb{Q}^n$ de manera no trivial).*

*Demostración.* Supongamos que $\mathrm{O}(J)_{\mathbb{R}}/\mathrm{O}(J)_{\mathbb{Z}}$ no es compacto. Por el Teorema 5.5, existe un elemento unipotente $U \neq \mathrm{Id}$ en $\mathrm{O}(J)_{\mathbb{Q}}$. Entonces $N := U - \mathrm{Id}$ es nilpotente, por lo que podemos elegir $v_1 \in \mathbb{Q}^d$ tal que $Nv_1 \neq 0$ y $N^2 v_1 = 0$. Mostraremos que $J[Nv_1] = 0$. Denotemos $B(x,y) = x^t J y$, la forma bilineal asociada a $J$, la cual es no degenerada. Para $x, y \in W := \mathbb{C}v_1 \oplus \mathbb{C}Nv_1$, usando que $U \in \mathrm{O}(J)_{\mathbb{Q}}$ y $N^2|_W \equiv 0$, tenemos

$$\begin{aligned} B(Nx,y) &= B((U-\mathrm{Id})x, y) = B(Ux,y) - B(x,y) \\ &= B(x, U^{-1}y) - B(x,y) = B(x, (\mathrm{Id} - N)y) - B(x,y). \end{aligned}$$

Por lo tanto $B(Nx,y) = -B(x,Ny)$ para todo $x, y \in W$. En particular

$$J[Nv_1] = B(Nv_1, Nv_1) = -B(v_1, N^2 v_1) = 0,$$

lo que prueba el resultado en una dirección.

Ahora supongamos que existe $v_0 \in V_{\mathbb{Q}} \smallsetminus \{0\}$ tal que $J[v_0] = B(v_0, v_0) = 0$. Como la forma es no degenerada, existe $\widetilde{v}_1 \in V_{\mathbb{Q}}$ tal que $B(v_0, \widetilde{v}_1) \neq 0$. Sea $q \in \mathbb{Q}$ la solución a la ecuación

$$0 = B(\widetilde{v}_1 + qv_0, \widetilde{v}_1 + qv_0) = B(\widetilde{v}_1, \widetilde{v}_1) + 2qB(\widetilde{v}_1, v_0).$$

Sea $v_1 = \widetilde{v}_1 + qv_0$. Tenemos que $B(v_0, v_1) = c \neq 0$ y $B(v_i, v_i) = 0$, $i = 0, 1$.

Sea $W = \mathbb{C}v_0 \oplus \mathbb{C}v_1$, y $W^\perp$ el subespacio de $V$ ortogonal a $W$ con respecto a $B(\cdot,\cdot)$. Veamos que $V = W \oplus W^\perp$. Supongamos que $w \in W \cap W^\perp$, luego $w = k_0 v_0 + k_1 v_1$ ya que $w \in W$, pero como $w \in W^\perp$ tenemos $0 = B(w, v_0) = k_1 c$, lo que implica que $k_1 = 0$. Análogamente $0 = B(w, v_1) = B(k_0 v_0, v_1) = k_0 c$, por lo tanto $k_0 = 0$. Entonces $W \cap W^\perp = \{0\}$. Resta ver que $V = W + W^\perp$. Veamos que la sucesión de espacios vectoriales

$$0 \to W^\perp \to V \xrightarrow{\eta} W^* \to 0,$$

es exacta, donde $\eta(v)(\cdot) = B(v, \cdot)$.

Es claro que el núcleo de $\eta$ coincide con $W^\perp$, por lo tanto sólo resta mostrar que $\eta$ es sobre. Sea $\gamma \in W^\perp$ y si denotamos $\gamma' \in V^*$ a su extensión a $V$, la aplicación $V \to V^*$ es sobre pues tienen la misma dimensión, por lo tanto existe $v_0 \in V$ que cumple $\gamma(w) = \gamma'(w) = B(v_0, w) = \eta(v_0)(w)$ para todo $w \in W$. Finalmente concluimos que $V = W \oplus W^\perp$ pues la exactitud de la sucesión implica que $\dim W + \dim W^\perp = \dim V$.

Veamos que existe $v_2 \in W^\perp_{\mathbb{Q}}$ tal que $B(v_2, v_2) = c \neq 0$. Supongamos que $B(v,v) = 0$ para todo $v \in W^\perp$. Sea $w \in W^\perp$, luego existe $w' \in W^\perp$ tal que $B(w, w') \neq 0$. Entonces $0 \neq 4B(w,w') = B(w+w', w+w') - B(w-w', w-w') = 0 - 0 = 0$ lo cual es un absurdo.

Sea ahora la transformación lineal $N : V \to \mathbb{C}v_0 \oplus \mathbb{C}v_2$ dada por $N(x) = -B(x, v_0)v_2 + B(x, v_2)v_0$, es decir,

$$N \sim \begin{pmatrix} 0 & 0 & c & \\ 0 & 0 & 0 & * \\ 0 & -c & 0 & \\ & & & \\ & 0 & & 0 \\ & & & \end{pmatrix}$$

con respecto a una base con primeros elementos $v_0$, $v_1$ y $v_2$. Claramente $N^2 = 0$, pues $N^2$ es triangular superior con ceros en la diagonal. De manera similar, tenemos $W^\perp = \mathbb{C}v_2 \oplus \{\mathbb{C}v_2\}^\perp$, donde $\{\mathbb{C}v_2\}^\perp$ denota el subespacio de $W^\perp$ ortogonal a $v_2$.

Luego si $x = k_0v_0 + k_1v_1 + k_2v_2 + w, y = k_0'v_0 + k_1'v_1 + k_2'v_2 + w' \in V$ con $k_i, k_i' \in \mathbb{C}$ y $w, w' \in \{\mathbb{C}v_2\}^\perp$, tenemos

$$\begin{aligned} B(Nx, y) &= B(N(k_0v_0 + k_1v_1 + k_2v_2 + w), k_0'v_0 + k_1'v_1 + k_2'v_2 + w') \\ &= cB(-k_1v_2 + k_2v_0, k_1'v_1 + k_2'v_2) \\ &= c^2(-k_1k_2' + k_2k_1'). \end{aligned}$$

Similarmente se calcula que $B(x, Ny) = c^2(-k_1'k_2 + k_2'k_1)$, por lo tanto

$$B(x, Ny) = -B(Nx, y).$$

Sea $U = e^N = \sum_{i=1}^{n} \frac{1}{i!}N^i$, luego

$$\begin{aligned} B(Ux, Uy) &= \sum_{i,j=0}^{n} \tfrac{1}{i!}\tfrac{1}{j!}B(N^ix, N^jy) = \sum_{i,j=0}^{n} \tfrac{(-1)^j}{i!j!}B(N^{i+j}x, y) \\ &= B(\left(\textstyle\sum_{i,j=0}^{n} \frac{(-1)^j}{i!j!}N^{i+j}\right)x, y). \end{aligned}$$

Más aún,

$$\begin{aligned} \sum_{i,j=0}^{n} \tfrac{(-1)^j}{i!j!}N^{i+j} &= \sum_{k=0}^{n} \left( \sum_{j=0}^{k} \tfrac{1}{(k-j)!j!}(-1)^j \right) N^k \\ &= \sum_{k=0}^{n} \frac{1}{k!} \left( \cdot \sum_{j=0}^{k} \binom{k}{j}(-1)^j \right) N^k = \mathrm{Id}, \end{aligned}$$

lo que implica que $U \in \mathrm{O}(J)$. Para finalizar, es claro que $U = e^N \neq \mathrm{Id}$ pues $N \neq 0$, y además $U \in \mathrm{O}(J)_{\mathbb{Q}}$, pues $\{v_0, v_1, v_2\} \subset V_{\mathbb{Q}}$, y si a este conjunto lo completamos a una base de $V_{\mathbb{Q}}$, se ve que $N(V_{\mathbb{Q}}) \subseteq V_{\mathbb{Q}}$, por lo tanto $U(V_{\mathbb{Q}}) \subseteq V_{\mathbb{Q}}$. □

### 5.4. Ejercicios.

**Ejercicio 5.1.** Usar Teorema 5.5 para mostrar que el retículo $\mathrm{SL}_2(\mathbb{Z})$ de $\mathrm{SL}_2(\mathbb{R})$ no es cocompacto. De manera similar, mostrar que los grupos de Bianchi (ver Ejemplo 8.5) no son cocompactos en $\mathrm{SL}_2(\mathbb{C})$.

**Ejercicio 5.2.** Dar un ejemplo concreto de una forma cuadrática $J[x]$ que no represente a cero en $\mathbb{Q}^n$, y así obtener un retículo cocompacto en $\mathrm{O}(J)$ por Teorema 5.6.

## Parte 2. Geometría espectral de 3-variedades hiperbólicas aritméticas

En la primera parte de estas notas se dio la noción general de un grupo aritmético. Uno de los primeros ejemplos dados fue el subgrupo $\mathrm{SL}_2(\mathbb{Z}[\sqrt{-1}])$ de $\mathrm{SL}_2(\mathbb{C})$ en Ejemplo 3.20. Este grupo actúa por isometrías en el espacio hiperbólico de dimensión 3 y por lo tanto define una 3-orbifold hiperbólica. En esta segunda parte del curso exploraremos la geometría de los subgrupos aritméticos de $\mathrm{SL}_2(\mathbb{C})$ (en realidad de $\mathrm{PSL}_2(\mathbb{C})$) en gran detalle. Por ejemplo, mostraremos cómo construir una cantidad

infinita de tales grupos (cocompactos como así también no cocompactos) y daremos una fórmula para sus covolúmenes. Como objetivo final, usaremos estos grupos para construir 3-variedades hiperbólicas isospectrales pero no isométricas.

## 6. Álgebras de cuaterniones

Uno de los temas a tratar en estas notas es la idea de que muchos aspectos de la geometría de una 3-variedad hiperbólica pueden ser caracterizados de una manera algebraica y estudiados usando técnicas provenientes del álgebra no conmutativa y la teoría de números. Es crucial para esta caracterización la noción de álgebra de cuaterniones, ya que veremos que todo grupo Kleiniano aritmético de covolumen finito es un álgebra de cuaterniones definido sobre un cuerpo de números. En esta sección veremos algunas de las propiedades básicas de las álgebras de cuaterniones. Muchos resultados de esta sección serán mencionados sin demostración. Para las pruebas, ver [16] o [11].

**6.1. Álgebras de cuaterniones: generalidades.** Sea $k$ un cuerpo de característica distinta que 2.

**Definición 6.1.** Un *álgebra de cuaterniones* sobre $k$ es un álgebra simple central de dimensión 4 sobre $k$ con base $\{1, i, j, ij\}$ que satisface las relaciones

$$i^2 = a, \qquad j^2 = b, \qquad ij = -ji,$$

para algún $a, b \in k^* = k \setminus \{0\}$.

Denotaremos el álgebra de cuaterniones en la Definición 6.1 por su *símbolo de Hilbert* $\left(\frac{a,b}{k}\right)$. Notar que la terminología "álgebra de cuaterniones" está motivada por el hecho de que las álgebras definidas arriba generalizan la construcción de Hamilton de $\mathbb{H}$, que con nuestra notación se corresponde al álgebra $\left(\frac{-1,-1}{\mathbb{R}}\right)$. Es llamado por el *álgebra de división* sobre $\mathbb{R}$.

**Teorema 6.2.** *El álgebra de cuaterniones* $\left(\frac{a,b}{k}\right)$ *es un álgebra simple central de dimensión cuatro. Recíprocamente, si $A$ es un álgebra simple central de dimensión cuatro sobre $k$ entonces existen $a, b \in k^*$ tales que* $A \cong \left(\frac{a,b}{k}\right)$.

En su total generalidad, el teorema de estructura de Wedderburn implica que toda álgebra central simple es isomorfa a un álgebra de matrices sobre un álgebra de división central.

**Teorema 6.3** (teorema de estructura de Wedderburn para álgebras de cuaterniones)**.** *Sea $A$ un álgebra de cuaterniones sobre un cuerpo $k$. Si $A$ no es un álgebra de división entonces* $A \cong \mathrm{M}_2(k)$.

Notar que si bien el Teorema 6.2 asegura que si $A$ es un álgebra central simple de dimensión cuatro sobre $k$ entonces existen $a, b \in k^*$ tales que $A \cong \left(\frac{a,b}{k}\right)$, esto no implica que $a, b$ determinen unívocamente la clase de isomorfismo de $A$. Esto se explica en el siguiente resultado, que dice que la clase de isomorfismo de $\left(\frac{a,b}{k}\right)$ no cambia si multiplicamos a $a$ o $b$ por cuadrados.

**Proposición 6.4.** *Si $a, b, x, y \in k^*$ entonces*

$$\left(\frac{a,b}{k}\right) \cong \left(\frac{ax^2, by^2}{k}\right).$$

*Demostración.* Sean $\{1, i, j, ij\}$ y $\{1, i', j', i'j'\}$ bases para $\left(\frac{a,b}{k}\right)$ y $\left(\frac{ax^2,by^2}{k}\right)$ respectivamente, y sea

$$\phi : \left(\frac{ax^2, by^2}{k}\right) \to \left(\frac{a, b}{k}\right)$$

el homomorfismo obtenido al definir $\phi(1) = 1$, $\phi(i') = xi$, $\phi(j') = yj$, $\phi(i'j') = xyij$ y extenderlo linealmente. La imagen de $\phi$ es la $k$-subálgebra de $\left(\frac{a,b}{k}\right)$ con base $\{1, xi, yj, xyij\}$. Como esta subálgebra tiene dimensión cuatro sobre $k$, debe coincidir con $\left(\frac{a,b}{k}\right)$. En otras palabras, $\phi$ es suryectiva. Cualquier homomorfismo suryectivo entre $k$-álgebras de la misma dimensión es un isomorfismo, por lo que la proposición sigue. □

Naturalmente, es muy útil saber cuándo $\left(\frac{a,b}{k}\right)$ es isomorfo a $\mathrm{M}_2(k)$ y cuándo lo es a un álgebra de división (que por el teorema de Wedderburn son las únicas dos posibilidades). La siguiente proposición será muy útil en este aspecto.

**Proposición 6.5.** *Las álgebras de cuaterniones* $\left(\frac{1,b}{k}\right)$ *y* $\mathrm{M}_2(k)$ *son isomorfas para cualquier* $b \in k^*$.

*Demostración.* El isomorfismo buscado está dado por

$$\psi : \left(\frac{1, b}{k}\right) \longrightarrow \mathrm{M}_2(k),$$

donde

$$\psi(x + yi + zj + wij) = \begin{pmatrix} x + y & z + w \\ b(z - t) & x - y \end{pmatrix}.$$

La inversa está dada por

$$\psi^{-1}\left(\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}\right) = \frac{1}{2}(\alpha + \delta + (\alpha - \delta)i + (\beta + b^{-1}\gamma)j + (\beta - b^{-1}\gamma)ij).$$

□

**6.2. Álgebra de cuaterniones sobre los números complejos.** Las álgebras de cuaterniones sobre $\mathbb{C}$ son muy fáciles de entender. Hay una única álgebra de cuaterniones sobre $\mathbb{C}$: $\mathrm{M}_2(\mathbb{C})$.

**Teorema 6.6.** *Si* $A$ *es un álgebra de cuaterniones sobre* $\mathbb{C}$ *entonces* $A \cong \mathrm{M}_2(\mathbb{C})$.

*Demostración.* El teorema fundamental del álgebra implica que todo elemento de $\mathbb{C}^*$ es un cuadrado, entonces $A \cong \left(\frac{1,1}{\mathbb{C}}\right)$ por la Proposición 6.4. Esta última álgebra de cuaterniones es isomorfa a $\mathrm{M}_2(\mathbb{C})$ por la Proposición 6.5. □

**6.3. Álgebras de cuaterniones sobre los números reales.** La estructura de las álgebras de cuaterniones sobre $\mathbb{R}$ es más complicada que sobre $\mathbb{C}$. Hay sólo dos álgebras de cuaterniones sobre $\mathbb{R}$ salvo isomorfismos: $\mathrm{M}_2(\mathbb{R})$ y $\mathbb{H}$.

**Teorema 6.7.** *Si* $A$ *es un álgebra de cuaterniones sobre* $\mathbb{R}$ *entonces es* $A \cong \mathrm{M}_2(\mathbb{R})$ *o* $A \cong \mathbb{H}$.

*Demostración.* La Proposición 6.4 implica que $A$ es isomorfa a una de las siguientes tres álgebras de cuaterniones: $\left(\frac{-1,-1}{\mathbb{R}}\right)$, $\left(\frac{1,-1}{\mathbb{R}}\right)$ o $\left(\frac{1,1}{\mathbb{R}}\right)$. La primera de estas álgebras es isomorfa a $\mathbb{H}$ por definición, mientras que la segunda y la tercera son isomorfas a $\mathrm{M}_2(\mathbb{R})$ por Proposición 6.5. □

**6.4. Álgebras de cuaterniones sobre cuerpos p-ádicos.** Sea $K$ un cuerpo $p$-ádico con uniformizador fijo $\pi$. Como ocurrió en el caso sobre $\mathbb{R}$, hay precisamente dos clases de isomorfismo de álgebras de cuaterniones sobre $k$. Además, otra vez tenemos una descripción explícita de la única álgebra de cuaterniones de división sobre $k$. Como la demostración nos llevaría demasiado lejos, sólo citamos el siguiente resultado y remitimos al lector a [16] para ver más detalles.

**Teorema 6.8.** *La $k$-álgebra $\left(\frac{u,\pi}{k}\right)$ es la única álgebra de cuaterniones de división sobre $k$, donde $k(\sqrt{u})$ es la única extensión cuadrática no ramificada de $k$.*

**6.5. Álgebras de cuaterniones sobre cuerpos de números.** Sea $k$ un cuerpo de números, $a, b \in k^*$ y consideramos el álgebra de cuaterniones $\left(\frac{a,b}{k}\right)$. Si $K$ es un cuerpo que contiene a $k$ entonces podemos obtener una $K$-álgebra de cuaterniones de $\left(\frac{a,b}{k}\right)$ por extensión de escalares: $\left(\frac{a,b}{k}\right) \otimes_k K \cong \left(\frac{a,b}{K}\right)$. En el estudio de la estructura de las álgebras de cuaterniones sobre cuerpos de números, a veces se elige como $K$ a la completación de $k$ (i.e., $\mathbb{C}, \mathbb{R}$ o un cuerpo $p$-ádico $k_{\mathfrak{p}}$ para algún primo $\mathfrak{p}$ de $k$) y se estudia el álgebra sobre $K$ obtenida por extensión de escalares. Por supuesto, uno espera poder obtener información acerca de la estructura del álgebra original sobre $k$.

Para hacer todo esto más preciso, sea $\{1, i, j, ij\}$ la base estándar de $\left(\frac{a,b}{k}\right)$ y $\sigma : k \hookrightarrow K$ una incrustación fija.

**Lema 6.9.** *Tenemos el siguiente isomorfismo*

$$\left(\frac{a,b}{k}\right) \otimes_\sigma K \cong \left(\frac{\sigma(a),\sigma(b)}{K}\right).$$

*Demostración.* Sea $\{1, i', j', i'j'\}$ la base estándar para $\left(\frac{\sigma(a),\sigma(b)}{K}\right)$. El isomorfismo es el que asigna

$$(a_0 + a_1 i + a_2 j + a_3 ij) \otimes_\sigma \alpha \mapsto \alpha(\sigma(a_0) + \sigma(a_1)i' + \sigma(a_2)j' + \sigma(a_3)i'j').$$

□

Como una aplicación de esto, consideramos el álgebra de cuaterniones $\left(\frac{-1,-1}{\mathbb{Q}}\right)$. Si $\sigma : \mathbb{Q} \to \mathbb{R}$ es la inclusión estándar, entonces el Lema 6.9 implica que $\left(\frac{-1,-1}{\mathbb{Q}}\right)$ es un álgebra de división (ya que $\sigma(-1) = -1$ y $\left(\frac{-1,-1}{\mathbb{R}}\right)$ es un álgebra de división). Un hecho interesante en la teoría de las álgebras de cuaterniones sobre cuerpos de números es que a diferencia de lo que sucede sobre $\mathbb{R}$, no hay una única álgebra de división sobre un cuerpo de números. De hecho, sobre todo cuerpo de números ¡hay infinitas clases de isomorfismos de álgebras de cuaterniones!

**Definición 6.10.** Sea $k$ un cuerpo de números, $v$ un lugar de $k$ correspondiente a $\sigma$, y sea $k_v$ la correspondiente completación de $k$. Decimos que un álgebra de cuaterniones $A$ sobre $k$ es *ramificada en $\sigma$ y $v$* si $A \otimes_\sigma k_v$ es un álgebra de división. Si no, decimos que $A$ *se parte en $\sigma$ y $v$*.

*Nota* 6.11. Por conveniencia, usualmente diremos que un álgebra de cuaterniones $A$ sobre $k$ es ramificada en un lugar $v$ de $k$ y omitiremos mencionar la incrustación $\sigma$ asociada.

*Nota* 6.12. Notar que si $A = \mathrm{M}_2(k)$ entonces $A \otimes_\sigma k_v \cong \mathrm{M}_2(k_v)$ para todo $\sigma, v$. En particular, todo lugar de $k$ se parte en $\mathrm{M}_2(k)$.

Recordar que por el Teorema 6.6, toda álgebra de cuaterniones sobre $k$ se parte en todos los lugares complejos. Entonces sólo los lugares reales o $p$-ádicos podrían ramificar.

Supongamos ahora que $k$ tiene $r_1$ lugares reales y $r_2$ lugares complejos. Denotamos por $S_\infty$ el conjunto de lugares arquimedianos de $k$, que es la unión de los lugares reales y complejos. Tenemos entonces los isomorfismos

$$\begin{aligned} A \otimes_{\mathbb{Q}} \mathbb{R} &\cong \bigoplus_{v \in S_\infty} A \otimes_k k_v \\ &\cong \mathrm{M}_2(\mathbb{C})^{r_2} \times \bigoplus_{\sigma : k \hookrightarrow \mathbb{R}} A \otimes_\sigma k_v \\ &\cong \mathrm{M}_2(\mathbb{C})^{r_2} \times \mathrm{M}_2(\mathbb{R})^s \times \mathbb{H}^{r_1 - s}, \end{aligned}$$

donde $s$ es el número de lugares reales de $k$ en los que $A$ se parte. En la Sección 8 veremos que los grupos Kleinianos aritméticos se construyen a partir de las álgebras de cuaterniones en que $r_2 = 1$ y $s = 0$. Una manera simple de ver esto es tomando como $k$ un cuerpo cuadrático imaginario, en cuyo caso $s = 0$ ya que allí no hay lugares reales.

Sea $\mathrm{Ram}(A)$ el conjunto de lugares de $k$ (pueden ser finitos o infinitos) en que $A$ es ramificado. El siguiente teorema clasifica las álgebras de cuaterniones sobre cuerpos de números e implica, como fue dicho arriba, que hay infinitas clases de isomorfismo de álgebras de cuaterniones de división sobre todo cuerpo de números. Para la demostración de esto, ver [16, Chapitre III.3].

**Teorema 6.13** (Clasificación de álgebras de cuaterniones sobre cuerpos de números)**.** *Sea $k$ un cuerpo de números. Si $A$ es un álgebra de cuaterniones sobre $k$ entonces* $\mathrm{Ram}(A)$ *es finito y de cardinalidad par. Recíprocamente, dado cualquier conjunto finito $S$ de lugares (finitos o infinitos) de $k$ con cardinalidad par, existe una única álgebra de cuaterniones $A$ sobre $k$ tal que* $\mathrm{Ram}(A) = S$.

El siguiente es un corolario inmediato del Teorema 6.13.

**Corolario 6.14.** *Si $k$ es un cuerpo de números y $A, A'$ son álgebras de cuaterniones sobre $k$ entonces $A \cong A'$ si y sólo si* $\mathrm{Ram}(A) = \mathrm{Ram}(A')$.

## 7. Órdenes en álgebras de cuaterniones: un primer vistazo

En esta sección introducimos la noción de orden en álgebras de cuaterniones y exploramos algunas de sus propiedades básicas. Nuestro objetivo es proveer los conocimientos necesarios para describir la construcción de subgrupos discretos de $\mathrm{PSL}_2(\mathbb{C})$ a partir de órdenes en álgebras de cuaterniones definidas sobre ciertos cuerpos de números. Esto nos permitirá dar la definición de una 3-variedad hiperbólica.

**7.1. Definiendo órdenes.** Sea $R$ un dominio de Dedekind con cuerpo cociente $K$. En la práctica siempre tomaremos $K$ un cuerpo de números o su completación en un primo finito y $R$ denotará su anillo de enteros. Sea $A$ un álgebra de cuaterniones sobre $K$.

**Definición 7.1.** Un elemento $\alpha \in A$ es *integral* con respecto a $R$ si su polinomio característico (reducido) $x^2 - \mathrm{tr}(\alpha)x + n(\alpha)$ tiene coeficientes en $R$. Llamamos a $\mathrm{tr}(\alpha)$ la *traza* (reducida) de $\alpha$ y a $n(\alpha)$ la *norma* (reducida) de $\alpha$.

Recordar que el conjunto de todos los elementos integrales de un cuerpo de números forman un anillo (y muy importante para lo que sigue, un $\mathbb{Z}$-módulo finitamente generado). Sin embargo, esto no es cierto en el caso de álgebra de cuaterniones. Considerar los siguientes dos elementos de $\mathrm{M}_2(\mathbb{Q})$:

$$A = \begin{pmatrix} \frac{5}{4} & -\frac{1}{8} \\ \frac{1}{2} & \frac{3}{4} \end{pmatrix}, \qquad B = \begin{pmatrix} \frac{11}{6} & \frac{1}{2} \\ \frac{5}{18} & \frac{7}{6} \end{pmatrix}.$$

Los polinomios característicos de $A$ y $B$ son $p_A(x) = x^2 - 2x + 1$ y $p_B(x) = x^2 - 3x + 2$. Entonces $A$ y $B$ son integrales (con respecto a $\mathbb{Z}$). Sin embargo, ni $A + B$ ni $AB$ son integrales; sus polinomios característicos son $p_{A+B}(x) = x^2 - 5x + \frac{809}{144}$ y $p_{AB}(x) = x^2 - \frac{487}{144}x + 2$. Veremos que el hecho de que el conjunto de elementos integrales en un álgebra de cuaterniones no es un anillo hace que la teoría de órdenes en álgebras de cuaterniones sea significativamente más complicada que el estudio de órdenes en cuerpos de números. Por otro lado, esto también hace que la teoría sea mucho más rica. En efecto, este hecho hace posible la construcción de Vignéras de 3-variedades hiperbólicas isospectrales.

**Definición 7.2.** Sea $V$ un espacio vectorial sobre $K$. Un *$R$-retículo* en $V$ es un $R$-módulo finitamente generado contenido en $V$. Un $R$-retículo $L$ se dice *completo* si $L \otimes_R K \cong V$.

El siguiente es un resultado básico en álgebra conmutativa.

**Proposición 7.3** ([1, Prop. 5.1])**.** *Un elemento $\alpha \in A$ es integral si y sólo si $R[\alpha]$ es un $R$-retículo en $A$.*

Ahora somos capaces de dar nuestra primera definición de órdenes en álgebras de cuaterniones.

**Definición 7.4.** Un *orden* $\mathcal{O}$ en $A$ es un $R$-retículo completo en $A$ que es también un subanillo de $A$. Un *orden maximal* es un orden en $A$ que es maximal con respecto a la inclusión.

**Ejemplo 7.5.** Damos algunos ejemplos de órdenes.

1. El anillo $\mathrm{M}_2(R)$ es siempre un orden de $\mathrm{M}_2(K)$. (Ver el Lema 7.7 abajo.)
2. Supongamos que $A = \left(\frac{a,b}{K}\right)$, donde $a, b$ son elementos integrales de $K$. (Notar que $A$ puede escribirse siempre de esta forma ya que el símbolo de Hilbert está definido salvo cuadrados, por lo que podemos 'limpiar denominadores' multiplicando a $a$ y $b$ por un cuadrado de $K$.) Entonces $R[1, i, j, ij]$ es un orden de $A$.

**Proposición 7.6.** *Un anillo $\mathcal{O}$ es un orden en $A$ si y sólo si $\mathcal{O}$ es un anillo de elementos integrales en $A$ que contiene a $R$ y satisface $\mathcal{O} \otimes_R K = A$. Además, todo orden está contenido en un orden maximal.*

*Demostración.* Sea $\mathcal{O}$ un orden de $A$ y $\alpha \in \mathcal{O}$. Como $\mathcal{O}$ es un $R$-retículo, también lo es $R[\alpha]$. Sigue entonces de la Proposición 7.3 que $\alpha$ es integral. Que $\mathcal{O}$ satisface las otras propiedades sigue de nuestra definición de orden.

Veamos ahora la recíproca. Como $\mathcal{O} \otimes_R K = A$, podemos elegir una base $\{x_1, x_2, x_3, x_4\}$ de $A$ en la que todos los $x_i$ están en $\mathcal{O}$. Como la traza reducida determina una forma bilineal simétrica no singular sobre $A$, $d = \det(\mathrm{tr}(x_i x_j)) \neq 0$. Sea $L = \{\sum a_i x_i : a_i \in R\}$. Entonces $L \subset \mathcal{O}$ pues $R \subset \mathcal{O}$ y $x_i \in \mathcal{O}$ para todo

$i$. Supongamos que $\alpha \in \mathcal{O}$ con $\alpha = \sum b_i x_i$ y $b_i \in K$. Para cada $j$ tenemos que $\alpha x_j \in \mathcal{O}$, entonces $\operatorname{tr}(\alpha x_j) = \sum b_i \operatorname{tr}(x_i x_j) \in R$. Por lo tanto $b_i \in \frac{1}{d}R$ y $\mathcal{O} \subset \frac{1}{d}L$. Sigue entonces que $\mathcal{O}$ es finitamente generado, lo que prueba la primera afirmación. La segunda afirmación se demuestra utilizando el lema de Zorn. □

**Lema 7.7.** *El orden* $\mathrm{M}_2(R)$ *es un orden maximal de* $\mathrm{M}_2(K)$.

*Demostración.* Si $\mathrm{M}_2(R)$ no es maximal, entonces sea $\mathcal{O}$ un orden maximal que contiene a $\mathrm{M}_2(R)$ y algún elemento $\begin{pmatrix} x & y \\ z & w \end{pmatrix}$ con al menos uno de los $x, y, z, w$ que no pertenezca a $R$. Sumando y multiplicando elementos de $R$ podemos conseguir un elemento $\alpha \in \mathcal{O}$ de la forma $\alpha = \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}$ con $a \notin R$. Tal elemento claramente no es integral, lo cual es una contradicción. □

Hemos demostrado que si $\mathcal{O}$ es un orden en $A$ entonces todo elemento de $\mathcal{O}$ es integral. La siguiente proposición nos provee una recíproca de este enunciado.

**Proposición 7.8.** *Si* $\alpha \in A$ *es integral entonces* $\alpha$ *está contenido en un orden maximal de* $A$.

*Demostración.* Si $\alpha \in R$ entonces $\alpha$ está en *todo* orden $A$ por la Proposición 7.6. Podemos asumir entonces que $\alpha \notin R$. En tal caso, $K(\alpha)$ es una extensión cuadrática de $K$ que está contenida en $A$. Sea $\beta \in A^*$ tal que $\beta\alpha\beta^{-1} = \overline{\alpha}$. La existencia de tal elemento es debida al teorema de Skolem–Noether y podemos tomar $\beta$ integral simplemente limpiando denominadores. El $R$-módulo generado por $\alpha$ y $\beta$ es $R + R\alpha + R\beta + R\alpha\beta$ y es claramente un orden de $A$. Este orden podría no ser maximal, pero hemos visto que todo orden está contenido en un orden maximal. □

**7.2. Números de tipo.** Supongamos que $\mathcal{O}_1$ y $\mathcal{O}_2$ son órdenes en $A$ que son isomorfos vía algún isomorfismo $f : \mathcal{O}_1 \to \mathcal{O}_2$. Por extensión de escalares, la función $f$ induce un isomorfismo $\hat{f} : \mathcal{O}_1 \otimes_R K \to \mathcal{O}_2 \otimes_R K$ tal que $\hat{f}(x) = f(x)$ para todo $x \in \mathcal{O}_1$. Como $\mathcal{O}_1$ y $\mathcal{O}_2$ son órdenes en $A$, $\mathcal{O}_1 \otimes_R K \cong A \cong \mathcal{O}_2 \otimes_R K$. Por lo tanto $\hat{f}$ es un automorfismo de $A$ y está dado por conjugación por un elemento $a \in A^*$ por el teorema de Skolem–Noether. En particular, $\mathcal{O}_2 = a\mathcal{O}_1 a^{-1}$. Concluimos que en un álgebra de cuaterniones, dos órdenes son isomorfos si y sólo si son conjugados.

**Definición 7.9.** El *número de tipo* de un álgebra de cuaterniones es el número de clases de conjugación de órdenes maximales.

El número de tipo de un álgebra de cuaterniones sobre un cuerpo de números es de algún modo una reminiscencia del número de clases de un cuerpo de números. Si bien el número de tipo es siempre finito (lo cual a priori no es obvio), puede ser arbitrariamente grande. Además, cuando $A$ es no ramificado en un primo arquimediano de $K$ veremos que el número de tipo es siempre una potencia de 2.

El número de tipo de un álgebra de cuaterniones sobre un cuerpo de números juega un rol crucial en la construcción de Vignéras de 3-variedades hiperbólicas isospectrales, como también en aplicaciones a otros temas como las formas modulares.

Sea $k$ un cuerpo de números y $A/k$ un álgebra de cuaterniones que cumple que $A \otimes_{\mathbb{Q}} \mathbb{R} \ncong \mathbb{H}^{[k:\mathbb{Q}]}$. Sea $k_A$ la extensión abeliana maximal de $k$ que tiene exponente 2, que es no ramificada fuera de los lugares reales en $\operatorname{Ram}(A)$, y en la que todo primo finito de $\operatorname{Ram}(A)$ se parte completamente. El siguiente teorema sigue de los resultados probados en [10, Section 3].

**Teorema 7.10.** *Las clases de conjugación de órdenes maximales en $A$ están en correspondencia uno a uno con los elementos de* $\mathrm{Gal}(k_A/k)$.

Los números de tipo son muy fáciles de computar usando Magma. A continuación, mostraremos cuántos cálculos se pueden realizar fácilmente usando Magma. Los lectores pueden realizar de manera gratuita éstos o otros cálculos similares en línea en `http://magma.maths.usyd.edu.au/calc/`.

**Ejemplo 7.11.** Sea $k = \mathbb{Q}(\sqrt{-10})$. Consideramos el álgebra de cuaterniones $A = \left(\frac{-1,-3}{k}\right)$. Veremos que el número de tipo de $A$ es 2 y calcularemos los conjuntos generadores para los representantes de las dos clases de conjugación de órdenes maximales de $A$ (considerados como módulos sobre $\mathcal{O}_k$).

```
> k<t> := QuadraticField(-10);
> t^2;
-10
> A<i, j, ij> := QuaternionAlgebra<k | -1, -3>;
> C := ConjugacyClasses(MaximalOrder(A));
> #C;
2
> IsConjugate(C[1], C[2]);
false
> Generators(C[1]);
[ 1, i, 1/2*i + 1/2*j, 1/2 + 1/2*t*i + 1/6*t*j + 1/6*ij ]
> Generators(C[2]);
[ 1, 2*i, t*i, 1 + 1/2*i + 1/2*j, 1/2*t + 1/4*t*i + 1/4*t*j,
    1/2*(t + 1) + 1/4*(t + 4)*i - 1/12*t*j + 1/6*ij ]
```

## 8. Grupos Kleinianos aritméticos y 3-variedades hiperbólicas

**8.1. Espacio hiperbólico tridimensional.** Comenzamos definiendo el espacio hiperbólico de dimensión 3, representado en el modelo del semiespacio superior

$$\mathbf{H}^3 = \{(z,t) : z \in \mathbb{C}, t > 0\}$$

con la métrica

$$ds^2 = \frac{|dz|^2 + dt}{t^2}.$$

De esta forma, $\mathbf{H}^3$ es la única variedad riemanniana conexa, simplemente conexa de dimensión 3 con curvatura seccional constante $-1$. Veremos a la esfera de Riemann $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ como la *esfera al infinito* correspondiente a $t = 0$. Las geodésicas en $\mathbf{H}^3$ son rectas euclidianas verticales o semicírculos ortogonales a $\mathbb{C}$.

**8.2. Grupos Kleinianos.** Un *grupo Kleiniano* es un subgrupo discreto de isometrías del espacio hiperbólico $\mathbf{H}^3$ que preserva orientación. Ya era conocido por Poincaré que el grupo $\mathrm{Isom}^+(\mathbf{H}^3)$ de isometrías del espacio hiperbólico de dimensión 3 es isomorfo a $\mathrm{PSL}_2(\mathbb{C})$, luego un grupo Kleiniano es simplemente un subgrupo discreto de $\mathrm{PSL}_2(\mathbb{C})$ de covolumen finito.

Los elementos de $\mathrm{PSL}_2(\mathbb{C})$ inducen una función biholomorfa de $\hat{\mathbb{C}}$ dada por transformaciones de Möbius:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto \left( z \mapsto \frac{az+b}{cz+d} \right).$$

Estas transformaciones fraccionarias lineales de $\hat{\mathbb{C}}$ se extienden a $\mathbf{H}^3$ vía la extensión de Poincaré. La extensión de Poincaré puede ser descripta geométricamente como sigue. Toda transformación fraccionaria lineal de $\hat{\mathbb{C}}$ puede escribirse como una composición de inversiones en círculos y rectas de $\hat{\mathbb{C}}$. Dado un círculo o una recta, hay un único hemisferio o plano en $\mathbf{H}^3$ que es ortogonal a $\hat{\mathbb{C}}$ y que corta a $\hat{\mathbb{C}}$ precisamente en dicho círculo o recta. La extensión de Poincaré es simplemente la composición de las inversiones en $\mathbf{H}^3$. Más concretamente, la extensión está dada por la fórmula:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto \left( (z,t) \mapsto \left( \frac{(az+b)\overline{(cz+d)} + a\bar{c}t^2}{|cz+d|^2 + |c|^2t^2}, \frac{t}{|cz+d|^2 + |c|^2t^2} \right) \right).$$

Por ejemplo, la traslación $z \mapsto z+1$ se extiende a $(z,t) \mapsto (z+1,t)$.

Si bien hay varias razones excelentes para estudiar grupos Kleinianos, nuestro interés en ellos se debe principalmente a lo siguiente:

**Teorema 8.1.** *Si $M$ es una 3-variedad hiperbólica orientable entonces $M$ es isométrica a $\mathbf{H}^3/\Gamma$ donde $\Gamma$ es un grupo Kleiniano libre de torsión.*

**8.3. Conmensurabilidad.** Introducimos ahora la noción de *conmensurabilidad*, la cual estará muy presente en lo que sigue. La conmensurabilidad fue ya definida en la primera parte de las notas en Definición 2.8. A continuación, daremos una noción ligeramente más refinada de conmensurabilidad en el caso de los subgrupos de $\mathrm{PSL}_2(\mathbb{C})$ y definiremos una noción relacionada para las 3-variedades hiperbólicas. En efecto, uno de los objetivos principales será asociar a una 3-variedad hiperbólica de volumen finito invariantes que sólo dependan de la clase de conmensurabilidad de la variedad. La noción de conmensurabilidad es natural del punto de vista de las 3-variedades hiperbólicas aritméticas, ya que veremos que la clase de conmensurabilidad de tales variedades corresponden a una cierta *álgebra de cuaterniones*, es decir un objeto de la teoría de números con una teoría de estructura muy rica.

**Definición 8.2.** Sean $\Gamma_1, \Gamma_2$ subgrupos de $\mathrm{PSL}_2(\mathbb{C})$.

- Decimos que $\Gamma_1$ y $\Gamma_2$ son *directamente conmensurables* si $\Gamma_1 \cap \Gamma_2$ tiene índice finito en $\Gamma_1$ y $\Gamma_2$. Decimos que $\Gamma_1$ y $\Gamma_2$ son *conmensurables en un sentido amplio* si $\Gamma_1$ y un conjugado de $\Gamma_2$ son directamente conmensurables.
- Sean $M_1, M_2$ 3-variedades hiperbólicas. Decimos que $M_1$ y $M_2$ son *conmensurables* si tienen un cubrimiento finito hiperbólico en común.

Notar que en la definición de conmensurabilidad, el cubrimiento común será considerado único salvo isometrías. En este caso, las dos variedades serán conmensurables si y sólo si sus grupos fundamentales son conmensurables en el sentido amplio. Es por esta razón que estaremos interesados en la conmensurabilidad en el sentido amplio.

**8.4. Grupos aritméticos Kleinianos.** En esta sección construiremos subgrupos discretos de $\mathrm{PSL}_2(\mathbb{C})$ a partir de órdenes en álgebras de cuaterniones y relacionaremos las propiedades geométricas de los grupos Kleinianos resultantes con propiedades algebraicas de las álgebras de cuaterniones asociadas. Esto nos permitirá definir lo que significa que un grupo Kleiniano sea aritmético.

A lo largo de esta sección emplearemos la siguiente notación. Sea $k$ un cuerpo de números con anillo de enteros $\mathcal{O}_k$ y sea $A$ un álgebra de cuaterniones sobre $k$. Un

*orden $\mathcal{O}$ de $A$* significará siempre que $\mathcal{O}$ es un $\mathcal{O}_k$-orden de $A$. Si $B$ es un subanillo de $A$ entonces denotaremos por $B^1$ el subgrupo multiplicativo de $B^*$ generado por elementos que tienen norma reducida igual a 1.

Finalmente, recordamos que $\mathrm{Ram}(A)$ (respectivamente $\mathrm{Ram}_f(A)$ o $\mathrm{Ram}_\infty(A)$) denota el conjunto de todos los lugares de $k$ (respectivamente finitos o infinitos) que ramifican en $A$.

**8.5. Grupos discretos de órdenes en álgebras de cuaterniones.** Sea $k$ un cuerpo de números de grado $n$ con un único lugar complejo $\nu$. Recordar que esto significa que de las $n$ incrustaciones $\sigma : k \hookrightarrow \mathbb{C}$, la imagen $\sigma(k)$ de $k$ estará contenida en $\mathbb{R}$ para precisamente $n-2$ de ellas. Las otras dos incrustaciones estarán dadas por $\nu$ y $\overline{\nu}$, el conjugado complejo de $\nu$. Denotaremos $S_\infty$ al conjunto de lugares arquimedianos de $k$.

Suponemos ahora que $A$ es un álgebra de cuaterniones sobre $k$ que es ramificado en todos los lugares reales de $k$. Al recordar que siempre hay un isomorfismo

$$A \otimes_{\mathbb{Q}} \mathbb{R} \cong \bigoplus_{v \in S_\infty} A \otimes_k k_v,$$

deducimos que

$$A \otimes_{\mathbb{Q}} \mathbb{R} \cong \mathbb{H}^{n-2} \times \mathrm{M}_2(\mathbb{C}). \tag{8.1}$$

Sea $\psi : A \hookrightarrow \mathrm{M}_2(\mathbb{C})$. Denotamos por $\psi : A \hookrightarrow \mathrm{M}_2(\mathbb{C})$ la composición de la incrustación natural $A \hookrightarrow A \otimes_{\mathbb{Q}} \mathbb{R}$ con el isomorfismo en (8.1) y la proyección de $\mathbb{H}^{n-2} \times \mathrm{M}_2(\mathbb{C})$ sobre $\mathrm{M}_2(\mathbb{C})$.

**Teorema 8.3.** *Sea $\mathcal{O}$ un orden maximal de $A$ y $\Gamma_\mathcal{O} = P\psi(\mathcal{O}^1) \subset \mathrm{PSL}_2(\mathbb{C})$.*

1. *$\Gamma_\mathcal{O}$ es un subgrupo discreto de $\mathrm{PSL}_2(\mathbb{C})$.*
2. *El volumen de $\mathbf{H}^3/\Gamma_\mathcal{O}$ está dado por*

$$\mathrm{Vol}(\mathbf{H}^3/\Gamma_\mathcal{O}) = \frac{d_k^{3/2}\zeta_k(2)}{(4\pi^2)^{[k:\mathbb{Q}]-1}} \cdot \left( \prod_{\mathfrak{p} \in \mathrm{Ram}_f(A)} (N(\mathfrak{p}) - 1) \right),$$

*donde $d_k$ es el valor absoluto del discriminante de $k$ y $\zeta_k(2)$ es la función zeta de Dedekind de $k$ evaluada en $s = 2$.*

*Demostración.* Para la prueba de que $\Gamma_\mathcal{O}$ es discreto, ver [16, Chapitre IV, Theoreme 1.1]. La fórmula para el covolumen de $\Gamma_\mathcal{O}$ se debe a Borel [2, Section 7.3]. □

Ahora podemos definir lo que significa que un grupo Kleiniano sea aritmético.

**Definición 8.4.** Sea $k$ un cuerpo de números con un único lugar complejo, sea $A$ un álgebra de cuaterniones sobre $k$ que ramifica en todos los lugares reales de $k$ y sea $\mathcal{O}$ un orden maximal de $A$. Un subgrupo de $\mathrm{PSL}_2(\mathbb{C})$ es un *grupo Kleiniano aritmético* si es conmensurable con $\Gamma_\mathcal{O}$ para algún triple $(k, A, \mathcal{O})$.

Los grupos $\Gamma_\mathcal{O}$ serán lo suficientemente importantes en nuestra discusión de grupos Kleinianos por lo que será muy útil denotarlos de una manera simple. De ahora en adelante, nos referiremos a los grupos de la forma $\Gamma_\mathcal{O}$ como *grupos Kleinianos aritméticos principales*. Así, un subgrupo de $\mathrm{PSL}_2(\mathbb{C})$ es un grupo Kleiniano aritmético si y sólo si es conmensurable a un grupo Kleiniano aritmético principal.

Recordar que el teorema de Wedderburn nos decía que si un álgebra de cuaterniones sobre $k$ no es isomorfo a $\mathrm{M}_2(k)$ entonces es un álgebra de división. En el contexto de las álgebras de cuaterniones $A$ que ramifican en todos los lugares reales de $k$, esto significa que $A$ no es un álgebra de división si y sólo si $k$ no tiene lugares reales (i.e., $k = \mathbb{Q}(\sqrt{-d})$ para algún entero positivo $d$ libre de cuadrados) y $A \cong \mathrm{M}_2(\mathbb{Q}(\sqrt{-d}))$.

**Ejemplo 8.5.** Consideremos un cuerpo cuadrático imaginario $\mathbb{Q}(\sqrt{-d})$ con anillo de enteros $\mathcal{O}_d$. En Lema 7.7 vimos que $\mathrm{M}_2(\mathcal{O}_d)$ es un orden maximal en el álgebra de cuaterniones $\mathrm{M}_2(\mathbb{Q}(\sqrt{-d}))$. El grupo $\mathrm{PSL}_2(\mathcal{O}_d)$ es llamado un *grupo de Bianchi* (ver también Ejemplo 3.20). Todo grupo de Bianchi contiene la isometría $z \mapsto z + 1$ de $\hat{\mathbb{C}}$ y por lo tanto es no compacto. De hecho, es un resultado conocido que el número de cúspides del grupo de Bianchi asociado a $\mathbb{Q}(\sqrt{-d})$ es igual al número de clases de ideales de $\mathbb{Q}(\sqrt{-d})$. De acuerdo al Teorema 8.3, $\mathrm{PSL}_2(\mathcal{O}_3)$ es el grupo de Bianchi de covolumen más chico. Usando Magma es fácil calcular los covolúmenes de los grupos de Bianchi.

Cuadro 1. Volúmenes de pequeñas orbifolds de Bianchi

| **d** | $\mathrm{Vol}(\mathbf{H}^3/\mathrm{PSL}_2(\mathcal{O}_d))$ |
|---|---|
| 1 | 0.30532186472574... |
| 2 | 1.00384100334120... |
| 3 | 0.16915693440160... |
| 5 | 4.20396925947605... |
| 6 | 5.18217289781959... |
| 7 | 0.88891492781635... |
| 10 | 9.81811844389802... |
| 11 | 1.38260830790264... |
| 13 | 13.9979614019778... |
| 14 | 20.3513407500735... |

Dado un entero positivo libre de cuadrados $d$, el volumen $\mathrm{Vol}(\mathbf{H}^3/\mathrm{PSL}_2(\mathcal{O}_d))$ puede ser computado en Magma con los siguientes comandos:

```
> d := 1;
> RR := RealField();
> pi := Pi(RR);
> R<x> := PolynomialRing(Rationals());
> k := NumberField(x^2 + d);
> Dk := Abs(Discriminant(Integers(k)));
> Zeta := Evaluate(LSeries(k),2);
> Dk^(3/2) * Zeta / (4*pi^2);
0.30532186472573967168486783831
```

Si bien no es necesario usar un programa computacional para calcular el discriminante de un cuerpo cuadrático, el código anterior puede modificarse fácilmente para calcular el volumen de grupos Kleinianos aritméticos de la forma $\Gamma_{\mathcal{O}}$. Uno

simplemente necesita reemplazar $x^2+d$ por el polinomio que define $k$ y calcular el término

$$\left(\prod_{\mathfrak{p}\in \mathrm{Ram}_f(A)} (N(\mathfrak{p})-1)\right)$$

que aparece en el Teorema 8.3, ya que este término es trivial en el caso en que $A \cong \mathrm{M}_2(k)$.

El siguiente teorema relaciona la topología de una 3-variedad hiperbólica aritmética $\mathbf{H}^3/\Gamma$ con la estructura de un álgebra de cuaterniones asociada.

**Teorema 8.6.** *Sea $M = \mathbf{H}^3/\Gamma$ una 3-variedad hiperbólica aritmética y supongamos que $\Gamma$ es conmensurable con $\Gamma_{\mathcal{O}}$, donde $\mathcal{O}$ es un orden maximal en un álgebra de cuaterniones $A$ sobre un cuerpo de números $k$. Son equivalentes:*

1. *$M$ es no compacta.*
2. *$k$ es un cuerpo cuadrático imaginario y $A \cong \mathrm{M}_2(k)$.*
3. *$\Gamma$ es conmensurable en el sentido amplio con un grupo de Bianchi.*

*Demostración.* Si $\Gamma$ no es cocompacto entonces tampoco es $\Gamma_{\mathcal{O}}$. Luego $\Gamma_{\mathcal{O}}$ contiene un elemento parabólico $\gamma$. Como el elemento $\gamma - \mathrm{Id}$ no es invertible podríamos concluir que $A$ no es un álgebra de división. Por el teorema de Wedderburn, $A \cong M_2(k)$. Ya vimos que un orden maximal de $\mathrm{M}_2(k)$ será un grupo Kleiniano aritmético sólo si $k$ es cuadrático imaginario. Luego (1) implica (2). Que (2) implica (3) sigue de la definición de un grupo de Bianchi y del hecho que los órdenes maximales en la misma álgebra de cuaterniones siempre nos darán grupos Kleinianos aritméticos que son conmensurables. Para probar que (3) implica (1), notar que todos los grupos de Bianchi contienen elementos parabólicos, por lo tanto $\Gamma$ también lo hará si es conmensurable en un sentido amplio a un grupo de Bianchi. □

Como consecuencia del Teorema 8.6 obtenemos lo siguiente.

**Corolario 8.7.** *Sea $M = \mathbf{H}^3/\Gamma$ una 3-variedad hiperbólica aritmética y supongamos que $\Gamma$ es conmensurable con $\Gamma_{\mathcal{O}}$, donde $\mathcal{O}$ es un orden maximal en un álgebra de cuaterniones $A$ sobre un cuerpo de números $k$. La variedad $M$ es compacta si y sólo si $A$ es un álgebra de división.*

## 9. Una construcción de Vignéras: ejemplos de 3-variedades hiperbólicas isospectrales

Sea $M$ una 3-variedad hiperbólica compacta y $\mathcal{E}(M)$ el multiconjunto de autovalores del operador de Laplace-Beltrami actuando en $L^2(M)$. Llamamos $\mathcal{E}(M)$ al *espectro de autovalores de Laplace* de $M$. Es conocido que $\mathcal{E}(M)$ es un subconjunto discreto e infinito de los números reales positivos. Si $M$ y $N$ son 3-variedades hiperbólicas compactas para las que $\mathcal{E}(M) = \mathcal{E}(N)$, entonces decimos que $M$ y $N$ son *isospectrales*.

En esta sección construiremos pares de 3-variedades hiperbólicas aritméticas compactas con el mismo espectro de autovalores del operador de Laplace. El método es debido originalmente a Vignéras [15]. De hecho, las 3-variedades hiperbólicas que construiremos serán siempre *fuertemente isospectrales*; esto es, tendrán el mismo espectro de autovalores con respecto a cualquier operador diferencial elíptico autoadjunto natural, e.g., el Laplaciano actuando sobre $p$-formas.

**9.1. Generalidades sobre isospectralidad.** Sea $G$ un grupo de Lie semisimple y $\Gamma$ un subgrupo discreto cocompacto de $G$. Denotamos por $L^2(\Gamma\backslash G)$ el espacio de funciones complejas de cuadrado integrable sobre $\Gamma\backslash G$ con respecto a la medida de Haar inducida de $G$ y por $C_c(G)$ el espacio de funciones a valores complejos con soporte compacto e infinitamente diferenciables sobre $G$. Definimos un operador unitario $R_\Gamma$ de $G$ en $L^2(\Gamma\backslash G)$ por

$$(R_\Gamma(g)f)(x) = f(xg)$$

donde $f \in L^2(\Gamma\backslash G), x \in \Gamma\backslash G$, y $g \in G$. Si $\Gamma'$ es otro subgrupo discreto y cocompacto de $G$ decimos que $\Gamma$ y $\Gamma'$ son *equivalentes en representaciones* si existe un isomorfismo unitario $T : L^2(\Gamma\backslash G) \to L^2(\Gamma'\backslash G)$ tal que

$$T(R_\Gamma(g)f) = R_{\Gamma'}(g)T(f)$$

para todo $g \in G$ y $f \in L^2(\Gamma\backslash G)$.

Es un hecho bien conocido que la equivalencia en representaciones implica isospectralidad con respecto al espectro de Laplace. De hecho, es un teorema de DeTurck y Gordon [5, Teorema 1.16] que equivalencia en representaciones implica isospectralidad fuerte.

**Teorema 9.1** (DeTurck y Gordon)**.** *Sea $G$ un grupo de Lie que actúa sobre una variedad riemanniana $M$ por isometrías. Supongamos que $\Gamma, \Gamma' \leq G$ actúan propia y discontinuamente sobre $M$. Si $\Gamma$ y $\Gamma'$ son equivalentes en representaciones entonces $\Gamma\backslash M$ y $\Gamma'\backslash M$ son fuertemente isospectrales.*

Sea $\phi \in C_c(G)$ y definimos el operador $R_\Gamma(\phi)$ sobre $L^2(\Gamma\backslash G)$ por

$$(R_\Gamma(\phi)f)(x) = \int_G \phi(g)f(xg)dg.$$

Este operador satisface la fórmula de la Traza de Selberg.

**Teorema 9.2** (Fórmula de la Traza de Selberg)**.** *Tenemos*

$$\operatorname{tr} R_\Gamma(\phi) = \sum_{[\gamma]\in A_\Gamma} \int_{C(\gamma,\Gamma)\backslash G} \phi(g^{-1}\gamma g)dg,$$

*donde $A_\Gamma$ denota el conjunto de clases de conjugación de elementos en $\Gamma$ y $C(\gamma, \Gamma)$ es el centralizador en $\Gamma$ de $\gamma$.*

Notar que $R_\Gamma$ está determinado por su traza. Esto es esencialmente debido a Dixmier [6] y usa el hecho que $R_\Gamma$ se descompone como suma discreta de representaciones unitarias irreducibles de $G$ con multiplicidades finitas. La idea es como sigue. Sea $(\pi_i)$ una colección de representaciones unitarias irreducibles de $G$ tal que para toda $\Phi \in C_c(G)$ tenemos

$$\sum m_i \operatorname{tr} \pi_i(\Phi) = \sum n_i \operatorname{tr} \pi_i(\Phi).$$

Supongamos que hay algún $i$ para el que $m_i \neq n_i$. Sin pérdida de generalidad podemos suponer que $m_i > 0$ y $n_i = 0$. Por Dixmier [6, Propositions 5.3.1 and 6.6.5], las representaciones $\sum m_i \operatorname{tr} \pi_i$ y $\sum n_i \operatorname{tr} \pi_i$ son *cuasi-equivalentes*, una condición que implica que $n_i \neq 0$.

Definimos el *peso* de una clase de conjugación $[\gamma]$ en $\Gamma$, por una medida sobre $C(\gamma, \Gamma)$, como el volumen $\operatorname{vol}(C(\gamma, \Gamma)\backslash C(\gamma, G))$. Uno entonces deduce lo siguiente de la fórmula de la Traza de Selberg.

**Teorema 9.3.** *Si dos subgrupos discretos cocompactos $\Gamma, \Gamma' \leq G$ tienen el mismo número de clases de conjugación con un mismo peso y clase en $G$, entonces $\Gamma$ y $\Gamma'$ son equivalentes en representaciones.*

**9.2. Espectro de grupos Kleinianos aritméticos principales.** Sea $k$ un cuerpo de números que tiene un único lugar complejo y $A$ un álgebra de cuaterniones de división sobre $k$. Sean $\mathcal{O}, \mathcal{O}'$ órdenes maximales de $A$ y $\Gamma_{\mathcal{O}}, \Gamma_{\mathcal{O}'}$ los grupos Kleinianos aritméticos asociados. El Corolario 8.7 dice que $\Gamma_{\mathcal{O}}, \Gamma_{\mathcal{O}'}$ son cocompactos.

Como estamos interesados en construir 3-variedades hiperbólicas, necesitamos asegurarnos que $P\psi(A^1)$ contiene elementos no triviales de orden finito. Supongamos que $P\psi(A^1)$ contiene un elemento de orden $n$. Entonces $\cos(\pi/n) \in k$ y $k(e^{\pi i/n})$ es una extensión cuadrática de $k$ que se incrusta en $A$. Hay finitos $n \geq 4$ para los que $[k(e^{\pi i/n}) : \mathbb{Q}] = 2[k : \mathbb{Q}]$, entonces usando apropiadamente el teorema de Albert-Brauer-Hasse-Noether (que implica que para un álgebra de cuaterniones $A$ sobre $k$ y una extensión cuadrática $L$ de $k$, existe una incrustación de $L$ en $A$ si y sólo si ningún primo que se parte en $L/k$ ramifica en $A$) cuando construimos $A$ vía el conjunto $\mathrm{Ram}(A)$ de primos que ramifican en $A$ podemos asumir que $P\psi(A^1)$ es libre de torsión. Sigue que $\Gamma_{\mathcal{O}}, \Gamma_{\mathcal{O}'}$ son libres de torsión.

Dado un grupo $U$ y un elemento $x \in U$, denotamos por $[x]_U$ la clase de conjugación de $x$ en $U$. El siguiente lema se hace claro.

**Lema 9.4.** *La incrustación $\psi$ de $A^1$ en $G = \mathrm{SL}_2(\mathbb{C})$ induce una biyección entre elementos $\mathcal{O}^1 \setminus \{\pm 1\}$ y $\Gamma_{\mathcal{O}} \setminus \{\pm 1\}$. Sea $x \in \mathcal{O}^1 \setminus \{\pm 1\}$ tal que $\gamma = \psi(x)$ es el elemento correspondiente de $\Gamma_{\mathcal{O}} \setminus \{\pm 1\}$. El centralizador $C(\gamma, \Gamma)$ corresponde a $k(x) \cap \mathcal{O}^1$ y la clase de conjugación $[\gamma]_G \cap \Gamma_{\mathcal{O}}$ corresponde a $[x]_A \cap \mathcal{O}^1$.*

Notar que el cuerpo $k(x)$ es una extensión cuadrática de $k$ que se incrusta en $A$ y $\Omega := k(x) \cap \mathcal{O}$ es un $\mathcal{O}_k$-orden cuadrático de $k(x)$ que es independiente de la selección de $x$ en $[x]_{\mathcal{O}^1}$. Llamaremos $B$ al *orden de la clase de conjugación de* $x$. Esta discusión, junto con los Teoremas 9.1 y 9.3 y un resultado de Eichler [7, Theorem 2], nos permiten deducir lo siguiente.

**Teorema 9.5.** *Supongamos que $\mathcal{O}^1$ y $\mathcal{O}'^1$ tienen el mismo número de clases de conjugación de elementos con una traza reducida fija y orden fijos, entonces $\Gamma_{\mathcal{O}} \backslash \mathbf{H}^3$ y $\Gamma_{\mathcal{O}'} \backslash \mathbf{H}^3$ son fuertemente isospectrales.*

Hemos reducido nuestra construcción de 3-variedades hiperbólicas isospectrales al estudio del número de clases de conjugación de elementos en un álgebra de cuaterniones con traza reducida fija. Para simplificar aún más este problema haremos uso del siguiente hecho, probado en [11, Theorem 12.4.5].

**Teorema 9.6.** *Sea $\mathcal{O}$ como antes y asumimos que $\Gamma_{\mathcal{O}}$ contiene un elemento de traza $t_0$. Entonces el número de clases de conjugación en $\Gamma_{\mathcal{O}}$ de elementos de $\Gamma_{\mathcal{O}}$ con traza $t_0$ es independiente de la elección del orden maximal $\mathcal{O}$.*

A la luz de los teoremas 9.5 y 9.6 es suficiente mostrar que si $\Omega$ es un $\mathcal{O}_k$-orden cuadrático que se incrusta en $\mathcal{O}$ entonces $\Omega$ se incrusta en $\mathcal{O}'$ también. En efecto, si $\Omega = \mathcal{O}_k[x]$ entonces toda incrustación de $\Omega$ en $\mathcal{O}$ determina (y es determinada por) un elemento de $\mathcal{O}$ con el mismo polinomio característico que $x$, la imagen en $\mathcal{O}$ de $x$.

Recordar de la Sección 7.2 que el número de clases de isomorfismos de órdenes maximales de $A$ es llamado el *número de tipo* de $A$. Resulta que cuando $A \otimes_{\mathbb{Q}} \mathbb{R} \ncong \mathbb{H}^{[k:\mathbb{Q}]}$, es siempre el caso que el número de tipo es una potencia de dos (para una

prueba, ver [4]). En particular, en el caso que estamos considerando tiene sentido hablar de $\Omega$ incrustado en $\frac{1}{2}$ de las clases de isomorfismos de órdenes maximales de $A$. (Esto es por supuesto un abuso de lenguaje. Sería más correcto decir que $\Omega$ se incrusta en representantes de la mitad de las clases de isomorfismos de órdenes maximales de $A$.)

La pregunta de si todo orden maximal de $A$ admite una incrustación de un orden cuadrático fijo $\Omega$ tiene una larga historia que retrocede al trabajo de Chevalley en los 1930's. En 1999, Chinburg y Friedman [4] resolvieron completamente este problema y mostraron que la proporción de clases de isomorfismos de órdenes maximales de $A$ que admite una incrustación de $\Omega$ es igual a $0, \frac{1}{2}$ o 1. De hecho, su teorema principal da condiciones necesarias y suficientes para que cada una de esas proporciones ocurran. Uno de los resultados de su trabajo, que será suficiente para nuestro propósito, es el siguiente.

**Teorema 9.7** (Chinburg-Friedman)**.** *Sea $k$ un cuerpo de números y $A$ un álgebra de cuaterniones sobre $k$ para el que $A \otimes_{\mathbb{Q}} \mathbb{R} \ncong \mathbb{H}^{[k:\mathbb{Q}]}$. Si $A$ es ramificado en un primo finito de $k$ y $\Omega$ es un $\mathcal{O}_k$-orden cuadrático que se incrusta en un orden maximal de $A$ entonces todo orden maximal de $A$ admite una incrustación de $\Omega$.*

Del Teorema 9.7 y la discusión anterior concluimos lo siguiente.

**Teorema 9.8.** *Sea $k$ un cuerpo de número con un único lugar complejo y $A$ un álgebra de cuaterniones de división sobre $k$ que ramifica en todos los lugares reales de $k$. Sean $\mathcal{O}, \mathcal{O}'$ órdenes maximales de $A$ para los que $\Gamma_{\mathcal{O}}$ y $\Gamma_{\mathcal{O}'}$ son libres de torsión. Si $A$ ramifica en un primo finito de $k$ entonces las variedades $\Gamma_{\mathcal{O}}\backslash\mathbf{H}^3$ y $\Gamma_{\mathcal{O}'}\backslash\mathbf{H}^3$ son fuertemente isospectrales.*

Para estar seguros de que las 3-variedades hiperbólicas que construimos no son isométricas primero vemos que si $\Gamma_{\mathcal{O}}\backslash\mathbf{H}^3$ y $\Gamma_{\mathcal{O}'}\backslash\mathbf{H}^3$ fueran isométricas entonces habría un elemento $\gamma$ en $\mathrm{PGL}_2(\mathbb{C})$ para el que $\Gamma_{\mathcal{O}} = \gamma\Gamma_{\mathcal{O}'}\gamma^{-1}$. La siguiente proposición demuestra que esto a su vez prueba que $\mathcal{O}$ y $\mathcal{O}'$ son conjugados en $A^*$. Para obtener variedades que no son isométricas es entonces suficiente elegir órdenes maximales que tengan diferentes tipos; esto es, que no sean conjugados en $A^*$.

**Proposición 9.9.** *Bajo la misma notación de antes y suponiendo que $\Gamma_{\mathcal{O}} = \gamma\Gamma_{\mathcal{O}'}\gamma^{-1}$ para algún $\gamma \in \mathrm{PGL}_2(\mathbb{C})$, se tiene que $\mathcal{O}$ y $\mathcal{O}'$ son conjugados en $A^*$.*

*Demostración.* Sea $\gamma = P(c)$ donde $c \in \mathrm{GL}_2(\mathbb{C})$. Entonces $\psi(A) = A\Gamma_{\mathcal{O}} = A\Gamma_{\mathcal{O}'}$, luego conjugar por $c$ induce un $k$-automorfismo de $A$ vía

$$\sum a_i\gamma_i \mapsto \sum a_i c\gamma_i c^{-1}$$

para $a_i \in k$ y $\gamma_i \in \psi(\mathcal{O}^1)$. Por el teorema de Skolem–Noether este es un automorfismo interno y existe un elemento $a \in A^*$ tal que $a\mathcal{O}^1a^{-1} = \mathcal{O}'^1$. Ahora consideramos el orden $\mathcal{O}\psi(\mathcal{O}^1)$ de $\psi(A)$ definido por

$$\mathcal{O}\psi(\mathcal{O}^1) := \left\{\sum a_i\gamma_i : a_i \in \mathcal{O}_k, \gamma_i \in \psi(\mathcal{O}^1)\right\}.$$

Supongamos que $\mathcal{D}$ es un orden maximal de $A$ para el que $\psi(D)$ contiene $\mathcal{O}\psi(\mathcal{O}^1)$. Si $\mathcal{D} \neq \mathcal{O}$ entonces $[\Gamma_{\mathcal{O}} : P\psi(\mathcal{D} \cap \mathcal{O})^1] > 1$. Pero $\psi(\mathcal{D} \cap \mathcal{O})^1 \supset (\mathcal{O}\psi(\mathcal{O}^1))^1 \supset \psi(\mathcal{O}^1)$. Luego $\mathcal{D} = \mathcal{O}$ y similarmente, $\mathcal{O}'$ es el único orden maximal de $A$ para el que $\mathcal{O}\psi(\mathcal{O}'^1)$ está contenido en $\psi(\mathcal{O}')$. Como $\psi(a)$ conjuga $\mathcal{O}\psi(\mathcal{O}^1)$ en $\mathcal{O}\psi(\mathcal{O}'^1)$, $a$ debe conjugar $\mathcal{O}$ en $\mathcal{O}'$. □

**9.3. Un ejemplo.** Sea $k = \mathbb{Q}(\sqrt{-5})$ y consideramos los ideales $\mathfrak{p}_1 = (11)$ y $\mathfrak{p}_2 = (3 + 2\sqrt{-5})$ de $\mathbb{Q}(\sqrt{-5})$. Ambos son ideales primos y tienen norma 121 y 29 respectivamente. Sea $A$ el álgebra de cuaterniones de división sobre $k$ definida por $\mathrm{Ram}(A) = \{\mathfrak{p}_1, \mathfrak{p}_2\}$. En términos de símbolos de Hilbert, $A$ está dada por $\left(\frac{44-11\sqrt{-5},\ -38-6\sqrt{-5}}{\mathbb{Q}(\sqrt{-5})}\right)$. Todo esto puede verificarse con el siguiente código de Magma.

```
> k<t> := QuadraticField(-5);
> Zk := Integers(k);
> p1 := 11*Zk;
> IsPrime(p1);
true
> p2 := (3 + 2*t)*Zk;
> IsPrime(p2);
true
> Norm(p1);
121
> Norm(p2);
29
> A := QuaternionAlgebra(p1 * p2);
> Basis(A);
[ 1, i, j, k ]
> i := Basis(A)[2];
> j := Basis(A)[3];
> i^2;
(44 - 11*t)
> j^2;
(-38 - 6*t)
```

El ideal primo $\mathfrak{p}_1 = (11)$ se parte completamente en $k(\sqrt{-1})$ y $k(\sqrt{-3})$, entonces el teorema de Albert-Brauer-Hasse-Noether implica que ninguna de estas extensiones se incrusta en $A$. Ninguna otra extensión ciclotómica de $k$ es cuadrática, entonces $A$ no contiene otras raíces de la unidad aparte de $\pm 1$.

El número de tipo de $A$ es dos, luego existen órdenes maximales $\mathcal{O}$ y $\mathcal{O}'$ de $A$ que no son conjugados.

```
> #ConjugacyClasses(MaximalOrder(A));
2
```

Hemos mostrado que $\Gamma_{\mathcal{O}}$ y $\Gamma_{\mathcal{O}'}$ son libres de torsión. Sigue del Teorema 9.8 y la Proposición 9.9 que las 3-variedades hiperbólicas aritméticas $\Gamma_{\mathcal{O}}\backslash\mathbf{H}^3$ y $\Gamma_{\mathcal{O}'}\backslash\mathbf{H}^3$ son fuertemente isospectrales pero no isométricas.

Ahora usamos el Teorema 8.3 para calcular el volumen de nuestras 3-variedades hiperbólicas isospectrales. (La ley de Weyl implica que las variedades riemannianas compactas isospectrales tienen siempre el mismo volumen). En este caso tenemos

$$d_k = 20$$

y

$$\zeta_k(2) = 1{,}855556893747120634762713411 65\ldots$$

entonces

$$\mathrm{Vol}(\Gamma_{\mathcal{O}}\backslash\mathbf{H}^3) = \mathrm{Vol}(\Gamma_{\mathcal{O}'}\backslash\mathbf{H}^3) = \frac{20^{3/2}\cdot(1{,}8555\ldots)\cdot 120\cdot 28}{4\pi^2} = 14125{,}336712\ldots$$

*Nota* 9.10. Notar que el teorema de rigidez de Mostow implica que cualquier isomorfismo de $\Gamma_{\mathcal{O}}$ y $\Gamma_{\mathcal{O}'}$ debería ser inducido por una isometría de $\Gamma_{\mathcal{O}}\backslash\mathbf{H}^3$ y $\Gamma_{\mathcal{O}'}\backslash\mathbf{H}^3$. Sigue que nuestras 3-variedades fuertemente isospectrales y no isométricas tienen grupos fundamentales no isomorfos.

## Referencias


[1] M.F. Atiyah, I.G. Macdonald, *Introduction to commutative algebra*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1969.

[2] A. Borel, *Commensurability classes and volumes of hyperbolic* 3*-manifolds*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **8**:1, 1–33 (1981).

[3] A. Borel, Harish-Chandra, *Arithmetic subgroups of algebraic groups*, Ann. of Math. (2) **75**, 485–535 (1962).

[4] T. Chinburg, E. Friedman, *An embedding theorem for quaternion algebras*, J. London Math. Soc. (2) **60**:1, 33–44 (1999). DOI: 10.1112/S0024610799007607.

[5] D.M. DeTurck, C. Gordon, *Isospectral deformations. II. Trace formulas, metrics, and potentials*, Comm. Pure Appl. Math. **42**, 1067–1095 (1989). DOI: 10.1002/cpa.3160420803.

[6] J. Dixmier, *Les $C^*$-algèbres et leurs représentations*, Cahiers Scientifiques, Fasc. XXIX. Gauthier-Villars & Cie, Éditeur-Imprimeur, Paris, 1964.

[7] M. Eichler, *Allgemeine Kongruenzklasseneinteilungen der Ideale einfacher Algebren über algebraischen Zahlkörper und ihre L-Reihen*, J. Reine Angew. Math. **179**, 227–251 (1938). DOI: 10.1515/crll.1938.179.227.

[8] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, *Grad. Stud. Math.* **34**. Amer. Math. Soc., Providence, 2001.

[9] A.W. Knapp, *Lie groups beyond an introduction*, *Progr. Math.* **140**. Birkhäuser Boston Inc., 2002.

[10] B. Linowitz, *Selectivity in quaternion algebras*, J. Number Theory **132**, 1425–1437 (2012). DOI: 10.1016/j.jnt.2012.01.012.

[11] C. Maclachlan, A.W. Reid, *The arithmetic of hyperbolic 3-manifolds*, Graduate Texts in Mathematics 219, Springer-Verlag, 2003.

[12] M. Orr, *Height bounds and the Siegel property*, Algebra Number Theory **12**:2, 455–478 (2018). DOI: 10.2140/ant.2018.12.455.

[13] G.T. Paula, *Comparison of volumes of Siegel sets and fundamental domains for* $\mathrm{SL}_n(\mathbb{Z})$, Geom. Dedicata **199**:1, 291–306 (2019). DOI: 10.1007/s10711-018-0350-5.

[14] M.S. Raghunathan, *Discrete Subgroups of Lie Groups*, Springer-Verlag, New York-Heidelberg, 1972.

[15] M.-F. Vignéras, *Variétés riemanniennes isospectrales et non isométriques*, Ann. of Math. (2) **112**:1, 21–32 (1980). DOI: 10.2307/1971319.

[16] M.-F. Vignéras, *Arithmétique des algèbres de quaternions*, Lecture Notes in Math. **800**, Springer-Verlag Berlin Heidelberg, 1980. DOI: 10.2307/1971319.

[17] D. Witte, *Introduction to Arithmetic Groups*, Deductive Press, 2015.

[18] R. Young, *The Dehn function of* $\mathrm{SL}(n,\mathbb{Z})$, Ann. of Math. (2) **177**, 969–1027 (2013). DOI: 10.4007/annals.2013.177.3.4.



Instituto de Matemática (INMABB), Departamento de Matemática, Universidad Nacional del Sur (UNS)-CONICET, Bahía Blanca, Argentina.
*Email address*: emilio.lauret@uns.edu.ar

CIEM-FaMAF, Universidad Nacional de Córdoba, Argentina.
*Email address*: miatello@famaf.unc.edu.ar

Department of Mathematics, Oberlin College, Oberlin, OH, 44074.
*Email address*: benjamin.linowitz@oberlin.edu

Curso

# PRIMOS, PARIDAD Y ANÁLISIS

HARALD HELFGOTT Y ADRIÁN UBIS

Publicaciones Matemáticas del Uruguay
Volumen 18, Diciembre 2023, Páginas 205–283
S 0797-1443

# PRIMOS, PARIDAD Y ANÁLISIS

HARALD HELFGOTT Y ADRIÁN UBIS

Resumen. Sea $\lambda(n) = 1$ cuando $n$ tiene un número par de divisores primos (contados con multiplicidad) y $\lambda(n) = -1$ de lo contrario. En general, distinguir entre números con un número par o impar de divisores primos es una de las tareas más difíciles en la teoría analítica de números. Un trabajo reciente de Matomäki y Radziwiłł muestra que, en promedio, ambos existen con la misma frecuencia aún en intervalos muy cortos. Este avance ya ha tenido varias aplicaciones importantes en las manos de Matomäki, Radziwiłł, Tao y Teräväinen. Explicaremos en detalle una prueba completa del resultado original de Matomäki y Radziwiłł, así como de varias aplicaciones.

## Índice



Versión final: 6 de junio de 2019.
Estas notas corresponden al curso dictado por los autores en la escuela AGRA III, Aritmética, Grupos y Análisis, del 9 al 20 de Julio de 2018 en Córdoba, Argentina.

## 1. Introducción

**1.1. Primalidad y paridad.** Los números primos son uno de los objetos de estudio principales de la teoría de números. Los problemas clásicos acerca de los primos son clásicos en parte por su dificultad: nuestras técnicas nos permiten acercarnos a la meta, pero, una vez que se trata de encontrar todos los casos primos y sólo ellos, nos vemos a menudo frustrados.

Consideremos un ejemplo entre muchos. La *conjetura fuerte de Goldbach* dice que todo número par $n \geq 4$ puede escribirse como la suma de dos primos. Rényi [22] mostró que todo número par suficientemente grande puede escribirse como la suma de un primo y un número que es el producto de a lo más $K$ primos, donde $K$ es una constante. Luego hubo una sucesión de trabajos que mejoraron la constante. Finalmente, en [4], Jing-Run Chen logró mostrar que todo par suficientemente grande puede escribirse como la suma de un primo y un número que es ya sea un primo o el producto de dos primos. Empero, la conjetura fuerte de Goldbach sigue sin resolver.

La misma situación ocurre con la *conjetura de los primos gemelos*, la cual plantea que hay un número infinito de primos $p$ tales que $p+2$ también es primo. Se puede dar una cota superior al número de tales $p$ con $p \leq N$ mediante métodos de *criba* (cómo veremos en los ejercicios en la sección 2.2), pero no tenemos una cota inferior. Como en el caso de Goldbach, hubo una sucesión de trabajos llevando a un resultado similar al de Chen; también hay otras aproximaciones a las dos conjeturas (*Goldbach débil*, *distancias acotadas entre primos*). Empero, las conjeturas en sí siguen abiertas.

En general, distinguir entre un primo y el producto de dos primos es muy difícil. Si se usa un procedimiento de *criba* de tipo tradicional, hacer tal distinción es no sólo difícil sino imposible. En general, las cribas, por sí solas, no llegan a distinguir entre números con un número par o impar de factores primos; se trata del *problema de paridad* (§2.2). Claro está, hay otros procedimientos, generalmente analíticos, que nos permiten probar algunos enunciados sobre los números primos.

Un resultado de base de este tipo es el *teorema de los números primos*. Recordamos que nos dice que el número $\pi(x)$ de primos entre 1 y $x$ es aproximadamente $x/\log x$. Para ser más precisos, en la versión de Hadamard y de la Vallée Poussin (1896), nos dice que

$$\pi(x) = Li(x) + O\left(xe^{-c\sqrt{\log x}}\right), \tag{1.1}$$

donde $Li(x) = \int_2^x \frac{dt}{\log t}$, $c > 0$ es una constante y $O(f(x))$ quiere decir un término cuyo valor absoluto está acotado por $f(x)$ por una constante.

La prueba se basa sobre las propiedades de la función zeta de Riemann $\zeta(s)$, y, en particular sobre el hecho que sabemos que no toma el valor 0 cuando $s$ está dentro de una cierta región. Por los mismos métodos, o usando (1.1), podemos mostrar que el número de enteros $n \leq x$ con un número par o impar de factores primos es asintóticamente el mismo:

$$\sum_{n \leq x} \lambda(n) = O\left(e^{-c\sqrt{\log x}}x\right), \tag{1.2}$$

donde $\lambda(n)$ es la *función de Liouville*, la cual es la función completamente multiplicativa tal que $f(p) = -1$ para todo número primo. El enunciado (1.2) también es cierto si $\lambda(n)$ se reemplaza por la más familiar *función de Möbius* $\mu(n)$, definida

como $\mu(n) = \lambda(n)$ para $n$ sin divisores cuadrados (aparte de 1), y como $\mu(n) = 0$ si $d^2|n$ para algún $d > 1$.

De hecho lo más importante es que la cota en (1.2) dividida por la cota trivial $x$ va a cero:

$$\sum_{n \le x} \lambda(n) = o(x),$$

donde $o(f(x))$ denota cualquier función $g(x)$ tal que $\lim_{x\to\infty} g(x)/f(x) = 0$. (Aquí, $f(x) = x$.) En otras palabras, el promedio de $\lambda$ en el intervalo $1 \le n \le x$ es $o(1)$, es decir, tiende a cero.

En general, es muy interesante saber que $\lambda$ tiene promedio $o(1)$ en un intervalo o conjunto de números, pues esto quiere decir precisamente que sabemos que en ese conjunto el número de enteros con un número par o impar de factores primos (contados con multiplicidad) es asintóticamente el mismo. Sabíamos, por ejemplo, que

$$\frac{1}{H} \sum_{n=N+1}^{N+H} \mu(n) = o(1)$$

para $H \ge N^\alpha$ y $\alpha$ dentro de un cierto rango. El primer resultado con $\alpha < 1$ fue probado por Hoheisel (1930); el mejor valor de $\alpha$ conocido en nuestros días es $7/12 + \epsilon$, $\epsilon > 0$ arbitrario ([20] y [21], basados en parte en [8]). También se tenían resultados "en promedio": por ejemplo, se sabía que

$$\int_X^{2X} \left| \sum_{x < n \le x+h} \lambda(n) \right|^2 dx = o(Xh^2) \tag{1.3}$$

para $h \ge X^{1/6+\epsilon}$, $\epsilon > 0$ arbitrario. La desigualdad (1.3) es equivalente al enunciado siguiente: dado $h = h(x)$ tal que $h(x) \ge x^{1/6+\epsilon}$,

$$\sum_{N < n \le N+h(x)} \lambda(n) = o(h(x))$$

para todo entero $N \in [x, 2x]$ fuera de un conjunto de $o(x)$ excepciones. (La equivalencia es inmediata; el lado izquierdo de (1.3) es una varianza.)

La *conjetura de Chowla* nos dice que, para $h_1, h_2, \dots, h_k$ enteros distintos,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n \le N} \lambda(n + h_1) \cdots \lambda(n + h_k) = 0. \tag{1.4}$$

Cuando $k > 1$, la conjetura está abierta y se considera muy difícil. En general, se cree que, para todo polinomio $P \in \mathbb{Z}[x]$ no constante,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n \le N} \mu(P(x)) = 0.$$

(Las conjeturas de *Hardy-Littlewood* y de *Schinzel* son enunciados relacionados para los números primos.) Nuevamente, no hay caso probado con $\deg P > 1$, si bien se tienen resultados para polinomios en dos variables.

Una puerta fue abierta recientemente por Kaisa Matomäki y Maksym Radziwiłł [11]. Ha conducido ya a numerosos resultados.

**1.2. Resultados de Matomäki, Radziwiłł y Tao.** El siguiente resultado fue un gran avance, puesto que, como hemos visto, la conclusión se sabía sólo bajo la condición $h(x) \geq x^{1/6+\epsilon}$, mucho más fuerte que $h(x) \to \infty$.

**Teorema 1.1.** *Sea $h = h(x)$ tal que $h(x) \to \infty$ cuando $x \to \infty$. Entonces, cuando $X \to \infty$,*

$$\left| \sum_{N < n \leq N + h(X)} \lambda(n) \right| = o(h(X)) \tag{1.5}$$

*para todo entero $N \in [X, 2X]$ fuera de un conjunto de $o(X)$ excepciones.*

En verdad, con algo de trabajo adicional, Matomäki y Radziwiłł demuestran un análogo del Teorema 1.1 para todo $f : \mathbb{Z} \to \mathbb{R}$ multiplicativo que satisfaga $|f(n)| \leq 1$ para todo $n$ [11, Thm. 1]:

$$\left| \frac{1}{h(X)} \sum_{N < n \leq N + h(X)} f(n) - \frac{1}{X} \sum_{X < n \leq 2X} f(n) \right| = o(1)$$

para todo $N \in [X, 2X]$ fuera de $o(X)$ excepciones. (Hay generalizaciones también para $f$ con valores complejos [12, Thm. A.1].)

(Por cierto, es plausible que (1.5) sea cierto para *todo* $N \in [X, 2X]$, con tal que, digamos, $h(x) \geq C(\log x)^2$. Empero, es fácil construir contraejemplos a tal enunciado con $f$ multiplicativa, $f \neq \lambda$.)

Así como el Teorema 1.1 es un resultado sobre la cancelación en intervalos cortos en promedio, se puede probar la conjetura de Chowla en promedio.

**Teorema 1.2.** ([12, Thm. 1.1]) *Sea $k \geq 1$ arbitrario. Sea $h = h(x)$ tal que $h(x) \to \infty$ cuando $x \to \infty$. Entonces, cuando $N \to \infty$,*

$$\sum_{n \leq N} \lambda(n + h_1) \cdots \lambda(n + h_k) = o(N)$$

*para enteros $(h_1, \dots, h_k)$, $1 \leq h_i \leq h(N)$, cualesquiera, fuera de un conjunto de $o\left(h(N)^k\right)$ excepciones.*

La prueba se basa en la del Teorema 1.1, vía el método del círculo.

Los Teoremas 1.1 y 1.2 admiten variantes cuantitativas. Aún para $h(X)$ pequeño, los métodos de Matomäki, Radziwiłł y Tao pueden dar una cota de la forma

$$O((\log \log h(X))/ \log h(X)).$$

Los términos de error de esta forma son típicos de pruebas en las cuales el caso de los primos se ve como una excepción. Las pruebas que veremos no son una excepción; de hecho, ponen a los enteros que no tienen una factorización típica en el término de error.

(En verdad, nosotros probaremos cotas del tipo $O(1/(\log h(x))^\alpha)$, $\alpha > 0$. Nuestro énfasis será en dar una exposición clara y concisa, y no en conseguir necesariamente las mejores cotas que los métodos permiten.)

Los métodos y resultados de Matomäki y Radziwiłł han tenido varias otras aplicaciones [13], [25], [26], [15], [14]. Discutiremos una de ellas: la prueba de una versión logarítmica de la conjetura de Chowla con $k = 2$.

**Teorema 1.3.** ([26, Thm. 1.1]) *Sean $a_1, a_2, b_1, b_2$ enteros tales que (a) $a_1, a_2 \geq 1$ y (b) $(a_1, b_1)$, $(a_2, b_2)$ no son múltiplos el uno del otro. Sea $\omega = \omega(x)$ tal que $\omega(x) \to \infty$ cuando $x \to \infty$. Entonces, cuando $x \to \infty$,*

$$\sum_{x/\omega(x)<n\leq x} \frac{\lambda(a_1 n + b_1)\lambda(a_2 n + b_2)}{n} = o(\log \omega(x)).$$

También este teorema se generaliza a una clase de funciones multiplicativas, incluyendo algunas para las cuales la conjetura de Chowla "estándar" (1.4) no sería cierta.

**1.3. Notación.** Cuando escribimos $f(n) \ll g(n)$, $g(n) \gg f(n)$ o $f(n) = O(g(n))$, queremos decir la misma cosa, esto es, que hay $N > 0$ y $C > 0$ tales que $|f(n)| \leq C \cdot g(n)$ para todo $n \geq N$. Escribimos $\ll_a$, $\gg_a$, $O_a$ si $N$ y $C$ dependen de $a$ (digamos). Como de costumbre, $f(n) = o(g(n))$ significa que $|f(n)|/g(n)$ tiende a 0 cuando $n \to \infty$.

Dado un subconjunto $A \subset X$, $1_A : X \to \mathbb{C}$ es la función característica de $A$:

$$1_A(x) = \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{si no.} \end{cases}$$

Denotamos por $|A|$ el número de elementos de un conjunto finito $A$.

Escribiremos $(a, b)$ por el máximo comun divisor de dos enteros $a, b \neq 0$, siempre y cuando no haya posibilidad de confusión con el par ordenado $(a, b)$.

La distancia $d(\beta_1, \beta_2)$ entre dos elementos $\beta_1, \beta_2 \in \mathbb{R}/\mathbb{Z}$ se define como $\min_{a\in\mathbb{Z}} |a + \beta_1 - \beta_2|$. Por ejemplo, la distancia entre 0,01 y 0,99 es 0,02. Podemos también definir, para $\beta \in \mathbb{R}$, la distancia $d(\beta, \mathbb{Z})$ como la distancia entre $\beta$ y el entero más cercano. Está claro que $d(\beta, \mathbb{Z})$ depende sólo de $\beta \bmod \mathbb{Z}$, y que $d(\beta_1, \beta_2) = d(\beta, \mathbb{Z})$ para cualquier $\beta$ tal que $\beta \equiv \beta_1 - \beta_2 \bmod \mathbb{Z}$.

**1.4. Agradecimientos.** El viaje y estadía de los autores fueron financiados en parte por la fundación Humboldt. Se deben las gracias a Boris Bukh y Nikos Frantzikinakis por sus valiosos comentarios.

## 2. Breve repaso y resultados preliminares

**2.1. Función zeta. Fórmula de Perron.** Sea $f : \mathbb{N} \to \mathbb{C}$. Muchas veces estamos interesados en sumas finitas de $f$:

$$\sum_{n\leq x} f(n).$$

En general, podemos tratar de obtener resultados sobre una función $f$ estudiando una funcion generatriz de $f$, como, por ejemplo, $\sum_{n=1}^{\infty} f(n) z^n$. Si $f$ está asociada a un problema multiplicativo, tiene sentido estudiar la siguiente función generatriz, llamada *función zeta* de $f$:

$$Z_f(s) = \sum_{n=1}^{\infty} f(n) n^{-s} \qquad s \in \mathbb{C}.$$

La idea de usar los factores $n^{-s}$ es que son multiplicativos: $(ab)^{-s} = a^{-s} b^{-s}$. La idea es que, si obtenemos suficiente información sobre $Z_f(s)$, podremos deducir información sobre $f(n)$.

Si $f$ es multiplicativa, entonces, como lo notó ya Euler, podemos factorizar $Z_f(s)$ como un producto sobre los primos:

$$Z_f(s) = \prod_p (1 + f(p)p^{-s} + f(p^2)p^{-2s} + \ldots) \qquad \Re s > 1 \tag{2.1}$$

Por ejemplo, en el caso $f = 1$ tenemos que su función zeta es

$$Z_1(s) = \zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \prod_p \frac{1}{1 - 1/p^s}, \tag{2.2}$$

de donde, por ejemplo, obtenemos que hay infinitos primos, puesto que $\lim_{s\to 1^+} \sum_n n^{-s}$ diverge, y $\lim_{s\to 1^+} 1/(1 - 1/p^s)$ puede divergir sólo si es un producto infinito. Escogiendo $s$ de manera más precisa, podemos obtener cotas útiles sobre los primos (ejercicio 2.4).

Veamos cómo usar la función zeta de manera elemental para acotar la suma de una función aritmética, $\sum_{n<x} \tau(n^2)n^{-1}$, con $\tau$ la función divisor. Para $s > 1$,

$$\begin{aligned} \sum_{n\le x} \frac{\tau(n^2)}{n} &\le \sum_n \frac{\tau(n^2)}{n} \left(\frac{x}{n}\right)^{s-1} = x^{s-1} \sum_n \frac{\tau(n^2)}{n^s} \\ &= x^{s-1} \prod_p \left(1 + \frac{3}{p^s} + O\left(\frac{1}{p^{2s}}\right)\right) \ll x^{s-1} \prod_p \left(1 + \frac{3}{p^s}\right) \\ &\le x^{s-1} \prod_p \left(1 + \frac{1}{p^s}\right)^3 \le x^{s-1} (\sum_n n^{-s})^3 \ll \frac{x^{s-1}}{(s-1)^3}, \end{aligned} \tag{2.3}$$

donde hemos usado el hecho que $\sum_n n^{-s} - \int_1^\infty t^{-s} dt \ll 1$. Tomando $s = 1 + \frac{1}{\log x}$, obtenemos

$$\sum_{n\le x} \frac{\tau(n^2)}{n} \ll (\log x)^3. \tag{2.4}$$

Éste es el orden de magnitud correcto: en verdad, $\sum_{n\le x} \tau(n^2)n^{-1}$ es asintótica a una constante por $(\log x)^3$.

Si queremos asintóticas para sumas necesitamos algo más preciso para relacionar las sumas con la función zeta. La herramienta que lo permite es la integral de Perron, que es capaz de expresar que un número $y$ sea mayor o menor que 1 en términos analíticos, y en particular en términos de los «armónicos» $y^s$, $s$ complejo: para $\sigma > 0$,

$$\frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} y^{-s} \frac{ds}{s} = \begin{cases} 1 & \text{si } 0 < y < 1, \\ 1/2 & \text{si } y = 1, \\ 0 & \text{si } y > 1. \end{cases} \tag{2.5}$$

Esta fórmula puede probarse usando el teorema de los residuos, o el teorema de inversión de Fourier para la función $1_{(0,\infty)}(x)e^{-\sigma x}$ en $x = \log y$, ya que $y^{-it} = e^{-it \log x}$. La integral debe comprenderse como el límite $\lim_{T\to\infty} \int_{\sigma - iT}^{\sigma + iT} y^{-s} ds/s$.

Así, usando Perron, obtenemos, para $x > 0$,

$$\sum_{n\le x} f(n) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} Z_f(s) x^s \frac{ds}{s} + \begin{cases} \frac{1}{2} f(x) & \text{si } x \in \mathbb{Z}, \\ 0 & \text{si } x \notin \mathbb{Z}. \end{cases} \tag{2.6}$$

Si bien todos los armónicos $x^{it}$ contribuyen a la integral, típicamente los importantes van a ser los que tienen frecuencia $t$ pequeña. Para demostrarlo, la idea es que es mejor trabajar con una función continua, antes que con la función en el lado derecho de (2.5). Podemos, por ejemplo, trabajar con la siguiente función, continua en $(0, \infty)$:

$$\psi_\delta(x) = 1_{(0,1-\delta)} + \frac{1-x}{\delta} 1_{[1-\delta,1]} \qquad 0 < \delta < \frac{1}{2}. \tag{2.7}$$

Por lo mismo que es igual a $1_{(0,1)}(x)$ cuando $0 \le x < 1-\delta$ o $x \ge 1$, está claro que, sí $|f(x)| \le 1$ para todo $x$, entonces

$$\sum_{n \le x} f(n) = O(\delta x) + \sum_{n=1}^{\infty} f(n) \psi_\delta\left(\frac{n}{x}\right) \tag{2.8}$$

Usando el teorema de inversión de Fourier (o la integral de Perron), no es difícil demostrar (ejercicio 2.5) que

$$\psi_\delta(y) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} (M\psi_\delta)(s) y^{-s} ds, \tag{2.9}$$

donde

$$(M\psi_\delta)(s) = \frac{1}{s(s+1)} \frac{1-(1-\delta)^{s+1}}{\delta}. \tag{2.10}$$

Utilizamos la notación $M\psi$ aquí pues se trata de una *transformada de Mellin* (lo cual no es sino una transformada de Fourier con un cambio complejo de variable). En general,

$$M\psi(s) := \int_0^\infty \psi(x) x^{s-1} dx,$$

y, bajo ciertas condiciones de integrabilidad sobre $\psi$ y $M\psi$,

$$\psi(x) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} M\psi(s) x^{-s} ds \qquad \text{(teorema de inversión de Mellin)},$$

lo cual se deduce del teorema de inversión de Fourier.

**Lema 2.1.** *Si $|f(n)| \le 1$ para todo $n$, entonces*

$$\sum_{n \le x} f(n) = O(\delta x \log x) + \frac{1}{2\pi i} \int_{1 + \frac{1}{\log x} - i\delta^{-2}}^{1 + \frac{1}{\log x} + i\delta^{-2}} x^s (M\psi_\delta)(s) Z_f(s) ds.$$

*Demostración.* Por (2.8) y (2.9), para $\sigma > 0$,

$$\sum_{n \le x} f(n) = O(\delta x) + \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} x^s (M\psi_\delta)(s) Z_f(s) ds.$$

Vemos que (2.10) implica que $|M(\psi_\delta)(s)| \le 2/\delta |s(s+1)|$ cuando $\Re s > 0$. Como $|f(n)| \le 1$ para todo $n$, $|Z_f(\sigma + it)| \le \zeta(\sigma) \le 1/(\sigma - 1) + O(1)$ para $\sigma > 1$. Por lo tanto,

$$\int_{\sigma + iT}^{\sigma + i\infty} x^s (M\psi_\delta)(s) Z_f(s) ds \ll \frac{|x^\sigma|}{\sigma - 1} \int_T^\infty \frac{1}{\delta |t(t+1)|} dt \ll \frac{|x^\sigma|}{\delta(\sigma - 1) T}. \tag{2.11}$$

Tomamos $\sigma = 1 + 1/\log x$ y $T = \delta^{-2}$, y el lado derecho de (2.11) se vuelve $O(\delta x \log x)$. La cota para la integral de $\sigma - i\infty$ a $\sigma - iT$ es evidentemente la misma. □

Luego efectivamente sólo las frecuencias pequeñas van a ser importantes para controlar el promedio. Veamos cómo usar este hecho para ver que hay cancelación en la suma

$$\sum_{n \leq x} \lambda(n)$$

Para ello, miramos a su función zeta correspondiente a $\lambda(n)$:

$$Z_\lambda(s) = \sum_n \lambda(n) n^{-s} = \prod_p (1 - p^{-s} + p^{-2s} - \dots) = \prod_p \frac{1}{1 + p^{-s}} = \prod_p \frac{1 - p^{-s}}{1 - p^{-2s}},$$

luego

$$Z_\lambda(s) = \frac{\zeta(2s)}{\zeta(s)}. \tag{2.12}$$

Para acotar el promedio de $f$ queremos extender $Z_f(s)$ a $s$ con parte real $< 1$ con el propósito de usar la integral en (2.6) para la $\sigma$ más pequeña que podamos, ya que $|x^s| = x^\sigma$.

Para $\Re s > 1$,

$$\zeta(s) = \frac{1}{s-1} + \sum_{n=1}^{\infty} n^{-s} - \int_1^\infty u^{-s}\, du = \frac{1}{s-1} + \sum_{n=1}^{\infty} \int_n^{n+1} (n^{-s} - u^{-s})\, du \tag{2.13}$$

y es muy sencillo ver que la última sumatoria converge en $\Re s > 0$. Por lo tanto, podemos hablar de $\zeta(s)$ como $1/(s-1)$ más una función analítica en la región $\Re s > 0$. Así, hemos extendido el numerador y el denominador en (2.12) a $\Re s > 0$; empero, también debemos controlar cuando el denominador se desvanece. Como el denominador es $\zeta(s)$, controlar cuando es cero en toda la región $\Re s > 0$ equivaldría a la hipótesis de Riemann. Vamos a citar el mejor resultado conocido en dicha dirección, debido a Vinogradov y Korobov:

**Teorema 2.2.** ([9, Thm. 8.29]) *Hay una constante $c > 0$ tal que $\zeta(s) \neq 0$ para $s = \sigma + it$ con $\sigma \geq 1 - c(\log t)^{-2/3}(\log\log t)^{-1/3}$, $|t| \geq 3$, y en dicha zona también se cumplen las cotas*

$$\frac{1}{\zeta(s)} \ll (\log t)^{2/3}(\log\log t)^{1/3},$$

$$\frac{\zeta'(s)}{\zeta(s)} \ll (\log t)^{2/3}(\log\log t)^{1/3}.$$

Con dicho resultado podemos acotar el promedio de $\lambda$:

**Teorema 2.3.**

$$\sum_{n \leq x} \lambda(n) \ll x \exp(-(\log x)^{3/5 + o(1)}).$$

*Demostración.* Aplicamos el Lema 2.1 para $f = \lambda$, y definimos $T = \delta^{-2}$. Por el teorema de Cauchy, podemos desplazar la línea de integración a los segmentos rectos $L_-$, $L_1$, $L_+$, donde $L_1$ va de $\sigma - iT$ a $\sigma + iT$, $L_-$ va de $1 + 1/\log x - iT$ a $\sigma - iT$ y $L_+$ de $\sigma + iT$ a $1 + 1/\log x + iT$; aquí $\sigma = 1 - c(\log T)^{-2/3}(\log\log T)^{-1/3}$, y utilizamos el Teorema 2.2 para asegurarnos que $\zeta(2s)/\zeta(s)$ es analítica en la región entre la

vieja y la nueva línea de integración. Así,

$$\sum_{n\le x}\lambda(n) = O(\delta x\log x) + \frac{1}{2\pi i}\int_{L_-+L_1+L_+} x^s M\psi_\delta(s)\frac{\zeta(2s)}{\zeta(s)}\,ds$$
$$= O(\delta x\log x) + O(x\delta^3) + \frac{1}{2\pi i}\int_{\sigma-iT}^{\sigma+iT} x^s M\psi_\delta(s)\frac{\zeta(2s)}{\zeta(s)}\,ds,$$

donde usamos la cota del Teorema 2.2, así como la cota $M\psi_\delta(s) = O(1/\delta|s(s+1)|)$, y el hecho que $\zeta(2s)$ está acotada por $\zeta(2\sigma)$; podemos asumir $2\sigma > 3/2$ (digamos), así que $\zeta(2\sigma) = O(1)$. Como $|x^s| = |x^\sigma| = x\exp\left(-c(\log x)(\log T)^{-2/3}(\log\log T)^{-1/3}\right)$, obtenemos, usando las mismas cotas,

$$\sum_{n\le x}\lambda(n) = O(\delta x\log x) + O\left(\delta^{-1}xe^{-c(\log x)(\log T)^{-2/3}(\log\log T)^{-1/3}}\log x\right) \tag{2.14}$$
$$\ll x\log x\left(e^{-C(\log x)^\alpha(\log\log x)^\beta} + e^{C(\log x)^\alpha(\log\log x)^\beta}\right.$$
$$\left.e^{-c(\log x)(2C(\log x)^\alpha(\log\log x)^\beta)^{-2/3}(\log 2C+\alpha\log\log x+\beta\log\log\log x)^{-1/3}}\right)$$

para $\delta = e^{-C(\log x)^\alpha(\log\log x)^\beta}$ y $T = \delta^{-2}$. Está claro que lo óptimo es escoger $\alpha$ y $\beta$ tales que $\alpha = 1 - 2\alpha/3$ y $\beta = -2\beta/3 - 1/3$, i.e., $\alpha = 3/5$ y $\beta = -1/5$. Asimismo, escogemos $C$ suficientemente pequeño para que $C < c(2C)^{-2/3}\cdot(2\alpha)^{-1/3}$, digamos. Así obtenemos

$$\sum_{n\le x}\lambda(n) \ll e^{-C(\log x)^{3/5}(\log\log x)^{-1/5}}x\log x \ll e^{-(C/2)(\log x)^{3/5}(\log\log x)^{-1/5}}x$$

para $C$ más grande que una constante. □

**Corolario 2.4.** *Sean $x \ge 1$, $t \le \exp^{(\log x)^{3/5-\epsilon}}$, $\epsilon > 0$. Entonces*

$$\sum_{x<n\le 2x}\frac{\lambda(n)}{n^{1+it}} \ll \exp(-(\log x)^{3/5+o_\epsilon(1)}).$$

*Demostración.* Por el Teorema 2.3 y sumación por partes (ejercicio 2.2), obtenemos en verdad que

$$\sum_{x<n\le 2x}\frac{\lambda(n)}{n^{1+it}} \ll (1+|t|)\exp(-(\log x)^{3/5+o(1)})\log x.$$

□

Es de suponer que la gran mayoría de lectores ya han visto por lo menos una prueba del teorema de los números primos basada sobre la integración compleja y las propiedades de $\zeta(s)$. De todas maneras, demos los primeros pasos de una prueba siguiendo las líneas generales que acabamos de sentar.

Por (2.2), la función $Z(s) = \log\frac{1}{\zeta(s)}$ es de la forma $Z_{1_P}+O(1)$ para $\Re s > 1$, donde $Z_{1_P} = \sum_p p^{-s}$ es la serie que corresponde a la función indicatriz $1_P$ del conjunto de todos los primos. Un problema con usar esta función $Z(s)$ es que $1/\zeta(1) = 0$, lo cual significa que no podemos definir $Z(s)$ como meromorfa alrededor de $s = 1$. Una manera de solucionar el problema sería integrar por partes en la integral que va de $c - i\infty$ a $c + i\infty$; luego aparecería, en vez de $Z(s)$, su derivada

$$Z'(s) = -\frac{\zeta'(s)}{\zeta(s)} = Z_\Lambda, \tag{2.15}$$

donde $\Lambda(n) = \log p$ para $n = p^\alpha$, $p$ algún primo, y $\Lambda(n) = 0$ para otros $n$. Debido a (2.13), $Z(s)$ es meromorfa en $\Re s > 0$ con un polo en $s = 1$ de resto 1.

En vez de seguir el procedimiento descrito en el párrafo anterior, es más fácil usar directamente $Z_\Lambda$ para demostrar el teorema de los números primos en la siguiente forma.

**Teorema 2.5.** *Para $x > 0$,*

$$\sum_{p \le x} \log p = x + O(x \exp(-(\log x)^{3/5+o(1)})).$$

*Demostración.* Ejercicio 2.7. □

**Corolario 2.6.** *Para $x > 0$,*

$$\pi(x) = Li(x) + O(x \exp(-(\log x)^{3/5+o(1)})), \tag{2.16}$$

$$\sum_{p \le x} \frac{\log p}{p} = \log x + \kappa_1 + O(\exp(-(\log x)^{3/5+o(1)})), \tag{2.17}$$

$$\sum_{p \le x} 1/p = \log\log x + \kappa_2 + O(\exp(-(\log x)^{3/5+o(1)})), \tag{2.18}$$

*donde $\kappa_1$ y $\kappa_2$ son constantes.*

Por cierto, las fórmulas menos precisas $\sum_{p \le x} (\log p)/p = \log x + O(1)$ y también $\sum_{p \le x} 1/p = \log\log x + \kappa_2 + O(1/\log x)$ ya eran conocidas antes del teorema de los números primos (Chebyshev-Mertens, años 1848–1874; ver [9, §2.2]).

*Esbozo de prueba del corolario.* Por sumación por partes. En el caso de (2.17), para evitar problemas debidos al hecho que $\exp(-(\log x')^{3/5+o(1)})$ puede ser bastante más grande que $\exp(-(\log x)^{3/5+o(1)})$ para $x'$ mucho más pequeño que $x$, conviene estimar las sumas

$$\sum_{n > x} \left(1_P(n) \cdot \frac{\log n}{n} - \frac{1}{n}\right), \qquad \sum_{n \le x} \frac{1}{n}, \tag{2.19}$$

la primera de ellas por sumación por partes, y la segunda por la fórmula $\sum_{n \le x} 1/n = \log x + \gamma + O(1/x)$, donde $\gamma$ es una constante (*constante de Euler*). Ésta última fórmula se establece fácilmente: las áreas que quedan entre la hipérbola $y = 1/x$ y las líneas horizontales $y = 1/n$ para $n \le x \le n+1$ pueden ponerse en pila, y entonces está claro que su suma para $n \ge 1$ es una constante, y su suma para $n > x$ es $O(1/x)$. Nótese por último que la primera suma en (2.19), tomada sobre todos los $n \ge 1$, converge a una constante, gracias a la estimación de esa misma suma por sumación por partes para $x$ general. Estas observaciones son suficientes para construir una prueba (ejercicio).

Procedemos de la misma manera para establecer 2.18: estimamos las sumas

$$\sum_{n > x} \left(\frac{1_P(n)}{n} - \frac{1}{n \log n}\right), \qquad \sum_{n \le x} \frac{1}{n \log n},$$

la primera de ellas por sumación por partes, y la segunda por una comparación con $\int_2^x 1/n \log n$. □

De manera análoga, podemos calcular la probabilidad de que un número no tenga factores primos en un rango dado.

**Lema 2.7.** *Sea* $1 \le Q^\alpha \le Q \le x^{\frac{1}{(\log\log x)^3}}$. *Entonces*

$$\sum_{n<x, p|n \Rightarrow p\notin[Q^\alpha, Q]} 1 = \alpha x + x \cdot O\left(\exp\left(-\min\left((\alpha\log Q)^{3/5+o(1)}, \frac{\log x}{3\log Q}\right)\right)\right).$$

Para $Q$ cercano a $x$, el problema comenzaría a cambiar de cariz (*función de Dickman*, *función de Buchstab*; ver, por ejemplo, [19, §7.1–7.2]).

*Demostración.* Observemos que la suma del enunciado es $\sum_{n<x} g(n)$, donde $g$ es la función totalmente multiplicativa con $Z_g(s) = \prod_{p\notin[Q^\alpha,Q]}(1-p^{-s})^{-1} = \zeta(s)\prod_{Q^\alpha\le p\le Q}(1-p^{-s})$. La función $Z_g(s)$ tiene un polo en $s=1$ con residuo $\prod_{Q^\alpha\le p\le Q}(1-p^{-1})$. Gracias a (2.18), vemos que

$$\begin{aligned}\prod_{p\le z}(1-p^{-1}) &= \exp\left(-\sum_{p\le z} p^{-1} + c + O\left(z^{-1}\right)\right)\\ &= e^{-\log\log z + c' + O(\exp(-(\log z)^{3/5+o(1)}))}\\ &= \frac{C}{\log z}(1+O(\exp(-(\log z)^{3/5+o(1)}))),\end{aligned}$$

donde $c$, $c'$ y $C$ son constantes. En consecuencia,

$$\prod_{Q^\alpha\le p\le Q}(1-p^{-1}) = \alpha\cdot(1+O(\exp(-(\alpha\log Q)^{3/5+o(1)}))). \tag{2.20}$$

Por otra parte,

$$|Z_g(\sigma+it)| \le |\zeta(\sigma+it)|\prod_{p\le Q}(1+p^{-\sigma}).$$

Para $1-1/\log Q \le \sigma \le 1$,

$$\begin{aligned}\prod_{p\le Q}(1+p^{-\sigma}) &\le \exp\left(\sum_{p\le Q} p^{-\sigma}\right) \le \exp\left(Q^{1-\sigma}\sum_{p\le Q} p^{-1}\right)\\ &\le \exp(e(\log\log Q + \kappa + o(1))) \ll (\log Q)^3,\end{aligned}$$

por el corolario 2.6. Usando (2.13), vemos que, para $s=\sigma+it$ con $\sigma \ge 1-1/\log Q$,

$$\begin{aligned}\zeta(s) &= \frac{1}{s-1} + O\left(\int_1^Q \left(\lfloor u^{-s}\rfloor - u^{-s}\right)du + \int_Q^\infty \left(\lfloor u^{-s}\rfloor - u^{-s}\right)du\right)\\ &= \frac{1}{s-1} + O\left(\int_1^Q \frac{du}{u} + |s|\int_Q^\infty u^{-\sigma-1}du\right)\\ &= \frac{1}{s-1} + O\left(\log Q + |t|/Q^\sigma\right).\end{aligned}$$

Así, para $|t| \le Q$, concluímos que $|\zeta(s)| = 1/(s-1) + O(\log Q)$. En particular, $|Z_g(\sigma+it)| \ll (\log Q)^4$ para $1 \le |t| \le Q$.

Procedamos como en la demostración del Teorema 2.3, sólo que con

$$\sigma = 1 - \min\left(c(\log T)^{-2/3}(\log\log T)^{-1/3}, 1/\log Q\right).$$

Obtenemos que

$$\sum_{n\le x} g(n) = \alpha x \cdot (1 + O(\exp(-(\alpha \log Q)^{3/5+o(1)}))) + S,$$

donde

$$S = O(\delta x \log x) + O(x\delta^3 (\log Q)^4) + O\left(\delta^{-1} x e^{-(\log x)\,\mín(c(\log T)^{-2/3}(\log\log T)^{-1/3}, 1/\log Q)} \log x\right).$$

Si $\log Q \le c(\log T)^{2/3}(\log\log T)^{1/3}$, procedemos exactamente como lo hicimos anteriormente, y obtenemos $S = xO(\exp(-(\log x)^{3/5+o(1)}))$. Si contrariamente $\log Q > c(\log T)^{2/3}(\log\log T)^{1/3}$, tomamos $\delta = x^{-1/2\log Q}$, $T = \delta^{-2} = x^{1/\log Q}$, y así

$$S = O(\delta x \log x) + O(\delta^3 x (\log x)^4) \ll \frac{x \log x}{x^{\frac{1}{2\log Q}}} \ll x^{1-\frac{1}{3\log Q}}$$

para $Q \le x^{1/(\log\log x)^3}$. □

*Ejercicios.*

**Ejercicio 2.1.** Una herramienta útil para nosotros va a ser sustituir sumas por integrales. Vamos a demostrar la *Regla del rectángulo*.

1. Demuestre que $f(n) = \int_n^{n+1} f(n)\,dt = \int_n^{n+1} f(t)\,dt + O(\int_n^{n+1} |f'(t)|\,dt)$.
2. Usando el apartado anterior, pruebe que, para enteros $a < b$ cualesquiera,

$$\sum_{a<n\le b} f(n) = \int_a^b f(t)\,dt + O\left(\int_a^b |f'(t)|\,dt\right). \qquad \text{(regla del rectángulo)} \tag{2.21}$$

   Por cierto, existen también expresiones similares con términos de error que involucran $f''$, $f'''$, etc. (*fórmula de Euler-Maclaurin*).
3. Observe que si $f$ es monótona, entonces la fórmula del apartado anterior da algo similar al *criterio integral* para series:

$$\sum_{a<n\le b} f(n) = \int_a^b f(t)\,dt + O(|f(b) - f(a)|)$$

4. Use las fórmulas anteriores para demostrar las siguientes aproximaciones:

$$\sum_{n\le x} n^3 = \frac{x^4}{4} + O(x^3), \qquad \sum_{n\le x} \operatorname{sen}\left(\frac{n^2}{x^{3/2}}\right) = cx^{3/4} + O(\sqrt{x}),$$

   con $c > 0$ la constante $c = \int_0^\infty \frac{\operatorname{sen} u}{2\sqrt{u}}\,du$.

**Ejercicio 2.2.**

1. Muestre que, para cualquier entero $N$ y $a_1, \dots, a_N \in \mathbb{C}$, $b_1, \dots, b_N \in \mathbb{C}$ arbitrarios,

$$\begin{aligned}\sum_{1\le n\le N} a_n b_n &= \sum_{1\le n\le N} (A(n) - A(n-1)) \cdot b_n \\ &= A(N)\cdot b_N - \sum_{1\le n\le N-1} A(n)\cdot(b_{n+1} - b_n),\end{aligned} \tag{2.22}$$

   donde $A(x) = \sum_{1\le n\le x} a_n$. (Se trata simplemente de recomponer los términos de una suma.) A esta técnica se la denomina *sumación por partes*. Es utilizada

cuando sabemos como estimar las sumas de tipo $A(x)$, y queremos estimar una suma $\sum_{n\le N} a_n b_n$.

2. Poner la igualdad de la forma siguiente hace más claro el paralelismo con la integración por partes (la técnica básica aprendida en un primer curso de cálculo integral). Reescriba (2.22) como sigue: para $N$, $A_0, A_1, \dots, A_N \in \mathbb{C}$ y $b_1, \dots, b_N \in \mathbb{C}$ arbitrarios,

$$\sum_{1\le n\le N} (A_n - A_{n-1})b_n = A_N b_N - A_0 b_1 - \sum_{1\le n\le N-1} A_n \cdot (b_{n+1} - b_n). \tag{2.23}$$

Alternativamente, muestre que (2.23) es un caso especial de la integración por partes, formulada para integrales de Lebesgue.

**Ejercicio 2.3.**

1. Sea $\tau(n)$ la función que cuenta el número de divisores de $n$, es decir $\tau(n) = \sum_{d|n} 1$. Demuestre, cambiando el orden de sumación, que

$$\sum_{n\le x} \tau(n) = \sum_{d\le x} \left\lfloor \frac{x}{d} \right\rfloor = x \sum_{d\le x} \frac{1}{d} + O(x) = x \log x + O(x),$$

donde $\lfloor x \rfloor$ es la parte entera de $x$. (Evidentemente, $\lfloor x \rfloor = x + O(1)$.)

2. *(Comentario)* En otras palabras, la esperanza del número de divisores de un entero aleatorio $n \le x$ es $\log x + O(1)$. Empero, la mayor parte de enteros tienen un número de divisores bastante menor. Como veremos después (ejercicio 2.4), la esperanza del número de divisores primos de un $n \le x$ aleatorio es $(1+o(1)) \log\log x$; Verifique que un número con $(1+o(1)) \log\log x$ divisores primos y sin divisores cuadrados aparte de 1 tiene $(\log x)^{(1+o(1))\log 2}$ divisores.

   Lo que sucede es que, si bien los números con $\ge C \log\log x$ divisores primos ($C > 1$) resultan estar en la minoría, tienen un número tan grande de divisores (¿cuántos?) que hacen que el promedio (o esperanza) sea bastante superior a la mediana (el valor tal que mitad de los casos son superiores y mitad son inferiores a él).

3. Use sumación por partes e integración por partes para demostrar las siguientes aproximaciones:

$$\sum_{n\le x} n\tau(n) = \frac{x^2}{2} \log x + O(x^2), \qquad \sum_{n\le x} \frac{\tau(n)}{n} = \frac{(\log x)^2}{2} + O(\log x).$$

**Ejercicio 2.4.** En este problema vamos a demostrar la asintótica $\sum_{p\le x} \frac{1}{p} = (1 + o(1)) \log\log x$ a partir del producto de Euler (2.2). Antes de ello, observa que de dicha asintótica es posible deducir que $\sum_{n\le x} w(n) = x(1 + o(1)) \log\log x$, con $w(n) = \sum_{p|n} 1$.

1. Use el criterio integral para series para demostrar que $\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \frac{1}{s-1} + O(1)$ para $s > 1$.
2. Tomando logaritmos en la identidad de Euler (2.2), usando el apartado anterior y la aproximación de Taylor $\log \frac{1}{1-x} = x + O(x^2)$ para $|x| < 1/2$, demuestre que

$$\sum_p p^{-s} = (1 + o(1)) \log \frac{1}{s-1} \qquad \text{para } s > 1.$$

En este contexto particular, $o(1)$ quiere decir "una cantidad una cantidad que tiende a 0 cuando $s \mapsto 1$".

3. Tome $s = 1 + \frac{\log\log x}{\log x}$. Usando $\sum_{p>x} p^{-s} \leq \sum_{n>x} n^{-s}$ y el criterio integral, demuestre que $\sum_{p>x} p^{-s} \ll 1$.
4. A partir de los dos apartados anteriores, demuestre que

$$\sum_{p\leq x} p^{-1} \geq \sum_{p\leq x} p^{-1-\frac{\log\log x}{\log x}} = \sum_{p} p^{-1-\frac{\log\log x}{\log x}} + O(1) = (1+o(1))\log\log x.$$

5. Para $p \leq x$, demuestre que $p^{-1-\frac{1}{\log\log x \log x}} = p^{-1}(1+o(1))$. A partir de ahí, pruebe las desigualdades

$$\sum_{p\leq x} p^{-1} \leq (1+o(1)) \sum_{p} p^{-1-\frac{1}{\log\log x \log x}} = (1+o(1))\log\log x.$$

**Ejercicio 2.5.**

1. Para $y > 0$, la ecuación de Perron (2.5) nos dice $1_{(0,1)}(y) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} y^{-s}\frac{ds}{s}$, asumiendo que definimos $1_{(0,1)}(1) = 1/2$. Usando dicha fórmula, y teniendo en cuenta que $1_{(0,b)}(y) = 1_{(0,1)}(y/b)$, demuestre que $1_{(0,b)}(y) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} b^s y^{-s}\frac{ds}{s}$.
2. Usando la fórmula del apartado anterior, demuestre la expresión (2.9).
3. Demuestre que si $\psi_\delta$ es la función definida en (2.7) entonces se cumple que $\int_0^\infty \psi_\delta(u)u^{s-1}\,du$ es igual a la función $M\psi_\delta(s)$ definida en (2.10). Así, el Teorema de inversión de Mellin también demuestra la fórmula (2.9)).
4. Recuerde que $|M\psi_\delta(s)| \leq 2/\delta|s(s+1)|$. Pruebe que $|M\psi_\delta(s)| \leq 2/s$ para $\delta \leq 1/2(s+1)$. Concluya que $|M\psi_\delta(s)| \leq 4/s$ para todo $\delta > 0$.

**Ejercicio 2.6.** Decimos que un entero $n$ es *libre de factores cuadrados* si no es divisible por ningun cuadrado $d^2$ con $d$ un entero mayor que 1. Vamos a demostrar que la proporción de números sin divisores cuadrados es $1/\zeta(2) = 0{,}6079\ldots$ Para ser precisos: sea $Q(x)$ el número de enteros $n \leq x$ libres de factores cuadrados; mostraremos que $\lim_{x\to\infty} Q(x)/x = \frac{1}{\zeta(2)}$.

Como veremos en el ejercicio 2.9 de la sección siguiente, es posible dar una prueba elemental de este mismo enunciado (con un mejor término de error). Ahora mostraremos como probarlo usando la función $\zeta(s)$, simplemente para practicar las técnicas que hemos expuesto en esta sección.

1. Demuestre que $Z_{\mu^2}(s) = \frac{\zeta(s)}{\zeta(2s)}$. Aplique el Lema 2.1 para expresar $Q(x)$ en términos de una integral sobre el segmento $V_0$ descrito por $s = 1 + \frac{1}{\log x} + it$, con $-\delta^{-c} \leq t \leq \delta^{-c}$, donde $\delta \in (0,1)$ y $c > 0$ son parámetros que luego elegiremos.
2. Usando la fórmula (2.13), demuestre que

$$\zeta(s) - \frac{1}{s-1} \ll \sum_{n=1}^{\infty} n^{-\sigma}\,\text{mín}\left(1, \frac{1+|t|}{n}\right) \ll (1+|t|)^{1-\sigma}$$

para $s = \sigma + it$, $\sigma \geq 1/2$. (Es posible probar cotas más fuertes, comenzando por la *cota de convexidad* $\zeta(s) \ll_{\sigma,\epsilon} (1+|t|)^{(1-\sigma)/2+\epsilon}$; no las necesitaremos.)
3. Deduzca del apartado anterior y del producto de Euler (2.2) para $\zeta(2s)$ que si estamos en la zona $\sigma_0 \leq \sigma < 3$ para algún $\sigma_0 > 1/2$ y a distancia $\gg 1$ de $s = 1$, entonces $Z_{\mu^2}(s) \ll_{\sigma_0} (1+|t|)^{1-\sigma}$.

4. Usando el Teorema de los residuos y el apartado $a$), demuestre que
$$Q(x) = \frac{x}{\zeta(2)} + O(\delta x \log x) + \frac{1}{2\pi i}\int_{H_+\cup V\cup H_-} x^s M\psi_\delta(s) Z_{\mu^2}(s) ds.$$
con $V$ el segmento $\sigma = \sigma_0$ ($\sigma_0 > 1/2$), $-\delta^{-c} \le t \le \delta^{-c}$ y $H_+$, $H_-$ los segmentos horizontales que unen los extremos de $V$ con los de $V_0$.
5. Use las cotas que tenemos para $Z_{\mu^2}$ y $M\psi_\delta$ para demostrar que
$$Q(x) = \frac{x}{\zeta(2)} + O(\delta x \log x) + O\left(x\delta^{(1+\sigma)c-1}\right) + O\left(x^\sigma \delta^{\sigma-1}\right).$$
6. Eliga $\delta > 0$, $\sigma > 1/2$ y $c$ adecuadamente para demostrar que $Q(x) = \frac{x}{\zeta(2)} + O_\epsilon(x^{2/3+\epsilon})$ para $\epsilon > 0$ arbitrario.

El método elemental que veremos en la siguiente seccion da un mejor exponente: $1/2$, en vez de $2/3$. Es posible obtener algunas mejoras ulteriores combinando el método elemental y el método analítico.

**Ejercicio 2.7.** Siga los pasos de la prueba del Teorema 2.3, usando la fórmula (2.15) y el Teorema de Vinogradov-Korobov (Teo. 2.2), para demostrar el teorema de los números primos en la siguiente forma:
$$\sum_{n\le x} \Lambda(n) = x + O(x\exp(-(\log x)^{3/5+o(1)})).$$
Deduzca el Teorema 2.5.

**Definición 2.8.** Mencionamos la transformada de Fourier, así que debemos precisar las constantes en nuestra definición. La *transformada de Fourier* $\widehat{f} : \mathbb{R} \to \mathbb{C}$ de una función $f : \mathbb{R} \to \mathbb{C}$ se define por
$$\widehat{f}(t) = \int_{-\infty}^{\infty} f(x) e(-xt) dx$$
donde $e(\alpha) = e^{2\pi i\alpha}$. Requerimos siempre que $f$ esté en $L^1$, es decir, $\int |f(x)|dx < \infty$ ("$f$ es integrable").

**2.2. Las cribas y sus limitaciones: el problema de paridad.** Hagamos una breve pausa, no para ver más resultados que necesitamos, sino para comprender mejor por qué los problemas con los que lidiaremos son difíciles, y no pueden ser resueltos sólo mediante cribas.

Digamos que tenemos un conjunto finito $A$ y un conjunto finito de propiedades $P$. Para cada subconjunto de $P$, se nos da una fórmula aproximada – con algún término de error – para el número de elementos de $A$ que satisfacen todas las propiedades en el subconjunto. Se nos pide dar una cota, superior o inferior, para el número de elementos de $A$ que no satisfacen ninguna de las propiedades.

La estrategia evidente sería aplicar el principio de inclusión-exclusión (ejercicio 2.8). Empero, el número de términos en una inclusión-exclusión completa es exponencial en el número de propiedades, por lo cual el error total podría ser enorme. Nos podemos preguntar si existe alguna manera de usar un número mucho más pequeño de términos para obtener cotas superiores o inferiores, o aún expresiones asintóticas. Una *criba* da una respuesta afirmativa a esta pregunta en contextos comunes en la teoría de números. (En particular, $A$ será un conjunto de enteros, y las propiedades a considerar estarán dadas con una biyección natural con los primos en un conjunto finito.)

Una amplia clase de cribas (*cribas pequeñas*) se pueden poner en el formalismo siguiente. Para permitir multiplicidades, en vez de trabajar con un conjunto $A$, trabajaremos con reales no negativos $a_n$ para $1 \leq n \leq x$. Las propiedades que consideraremos serán la divisibilidad por un conjunto $\mathscr{P}$ de primos $p \leq z$, donde $z$ es un parámetro. Para cada $d \leq D$ (donde $D \geq z$) libre de factores cuadrados y con factores primos sólo en $\mathscr{P}$, se nos da que

$$\sum_{\substack{1 \leq n \leq x \\ d|n}} a_n = g(d)X + r_d, \tag{2.24}$$

donde $g : \mathbb{Z}^+ \to \mathbb{R}$ es una función multiplicativa, $r_d$ es el término de error y $X$ depende sólo de $x$. Aquí $0 \leq g(p) < 1$ para $p \in \mathscr{P}$. Los términos de error $r_d$ son tal que $\sum_{d \leq D} |r_d|$ sea pequeño comparado con $X$ (digamos, $O(X/(\log X)^A)$).

Hay diversas cribas para diversos problemas. Una criba muy simple nos permite calcular $\sum_{n \leq x : \forall d > 1, d^2 \nmid n} a_n$. Asimismo, es posible usar una criba para dar cotas superiores e inferiores para $\sum_{n \leq x : n \text{ tiene } \leq 3 \text{ factores primos}} a_n$, digamos, o para dar una cota superior para $\sum_{p \leq x} a_p$. ¿Podemos, empero, dar una cota inferior no trivial para $\sum_{p \leq x} a_p$, o una estimación no trivial de $\sum_{n \leq x} \lambda(n) a_n$, puramente a través de una criba?

La respuesta es no, como lo muestra el siguiente argumento de Selberg. Por una parte, la secuencia $a_n = 1/2$ satisface (2.24) con $g(d) = 1/d$, $|r_d| \leq 1$. Por otra parte, consideremos la secuencia $a'_n = (\lambda(n) + 1)/2$. Vemos que, para $d \leq D = x^{1-\epsilon}$,

$$\sum_{\substack{n \leq x \\ d|n}} a'_n = \frac{1}{2} \sum_{n \leq x/d} 1 + \frac{\lambda(d)}{2} \sum_{n \leq x/d} \lambda(n) = \frac{x}{2d} + O\left(\frac{x/d}{(\log x)^{A+1}}\right)$$

gracias a (1.2). Por lo tanto, (2.24) se cumple para $d \leq D$, nuevamente con $g(d) = 1/d$ y $\sum_{d \leq D} |r_d| = O(x/(\log x)^A)$. En otras palabras, una criba con el dato (2.24) para $d \leq D = x^{1-\epsilon}$ no puede distinguir entre $a_n$ y $a'_n$. Ahora bien, $\sum_{p \leq x} a'_p = 0$ y $\sum_{n \leq x} \lambda(n) a'_n = \sum a'_n \sim x/2$, sumas que difieren drásticamente de $\sum_p a_p = \sum_{p \leq x} 1/2 \sim x/2 \log x$ y $\sum_{n \leq x} \lambda(n) a_n = \sum_{n \leq x} \lambda(n) = o(x)$.

Por lo tanto, una criba, aún con datos válidos para todo $d \leq x^{1-\epsilon}$, no puede, *por si sóla*, darnos una cota inferior para $\sum_p a_p$, o una estimación no trivial para $\sum_n \lambda(n) a_n$. A este hecho se le llama "problema de la paridad"; fue formulado en la forma que acabamos de ver por Selberg.

La razón del nombre se vuelve más clara si consideramos también la secuencia $a''_n = (1 - \lambda(n))/2$. Exactamente por el mismo argumento que acabamos de ver, esta secuencia satisface (2.24) con los mismos valores de $g(d)$ para $d \leq x^{1-\epsilon}$ que las secuencias $\{a_n\}$ y $\{a'_n\}$. Por lo tanto, una criba, por si sola, no puede distinguir entre $a'_n$ y $a''_n$. Ahora bien, $a'_n = 1$ cuando $n$ tiene un número par de factores primos, y $a'_n = 0$ de lo contrario, mientras que $a''_n = 1$ cuando $n$ tiene un número impar de factores primos, y $a''_n = 0$ de lo contrario.

No utilizaremos métodos de cribas en el trabajo presente. Empero, se trata de herramientas útiles y completamente usuales en la teoría de números, por lo cual un cierto conocimiento de ellas es de suma importancia.

*Ejercicios.*

**Ejercicio 2.8.**

1. Sea $A$ un conjunto finito. Entonces el número de elementos de $A$ que no son ni rojos ni grandes es igual a

$$|A| - |\{a \in A : a \text{ es rojo}\}| - |\{a \in A : a \text{ es grande}\}| + |\{a \in A : a \text{ es rojo y grande}\}|.$$

   Éste es un caso especial del *principio de inclusión-exclusión*.
2. He aquí un enunciado general del principio de inclusión-exclusión. Sea $A$ un conjunto finito y $\mathscr{P}$ un conjunto finito de propiedades que los elementos de $A$ pueden o no tener. Para $P \subset \mathscr{P}$, denotemos por $A_P$ el conjunto de elementos de $A$ que satisfacen todas las propiedades en $P$, y por $A_{\setminus \mathscr{P}}$ el conjunto de elementos de $A$ que no satisfacen ninguna de las propiedades en $\mathscr{P}$. Muestre que

$$\left|A_{\setminus \mathscr{P}}\right| = \sum_{i=0}^{|\mathscr{P}|} (-1)^i \sum_{P \subset \mathscr{P} : |P| = i} |A_P| \,. \tag{2.25}$$

**Ejercicio 2.9.** Veamos como aún el principio de inclusión-exclusión, sin más elaboración, puede dar una solución a un problema de criba bastante particular. Recordemos que $Q(x)$ denota el número de enteros $n \leq x$ libres de factores cuadrados.

1. Usando (2.25), muestre que, para $\mathbf{P}$ el conjunto de primos $p \leq \sqrt{x}$,

$$\begin{aligned} Q(x) &= \sum_{i=0}^{|\mathbf{P}|} (-1)^i \sum_{P \subset \mathbf{P} : |P| = i} \left|\{n \leq x : p^2 | n \ \ \forall p \in P\}\right| \\ &= \sum_{d \mid \prod_{p \in \mathbf{P}} p} \mu(d) |\{n \leq x : d^2 | n\}| = \sum_{d \leq \sqrt{x}} \mu(d) \left|\{n \leq x : d^2 | n\}\right|. \end{aligned}$$

   Sugerencia: utilice dos veces el hecho que $d^2 | n$ implica $d \leq \sqrt{x}$.
2. Tomando en cuenta que, para todo $\ell \in \mathbb{Z}^+$,

$$|\{n \leq x : \ell | n\}| = \frac{x}{\ell} + O(1),$$

   muestre que

$$Q(x) = x \cdot \sum_{d \leq \sqrt{x}} \frac{\mu(d)}{d^2} + O(\sqrt{x}).$$

   Concluya que

$$Q(x) = x \cdot \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} + O(\sqrt{x}) = \frac{x}{\zeta(2)} + O(\sqrt{x}).$$

**Ejercicio 2.10.** Demuestre las dos siguientes desigualdades, las cuales podríamos llamar versiones "incompletas" de (2.25): para $j \leq |\mathscr{P}|$ par,

$$\left|A_{\setminus \mathscr{P}}\right| \leq \sum_{i=0}^{j} (-1)^i \sum_{P \subset \mathscr{P} : |P| = i} |A_P| \,, \tag{2.26}$$

mientras que, para $j \leq |\mathscr{P}|$ impar, la desigualdad va en la otra dirección:

$$\left|A_{\setminus \mathscr{P}}\right| \geq \sum_{i=0}^{j} (-1)^i \sum_{P \subset \mathscr{P} : |P| = i} |A_P| \,. \tag{2.27}$$

El primer resultado no trivial de cribas (*criba pura de Brun*) se basó en estas desigualdades. En la teoría de probabilidades, se llaman *desigualdades de Bonferroni*, aunque el trabajo de Bonferroni [1] es posterior a aquel de Brun [3].

Para ver porque estas desigualdades pueden ser útiles para las cribas, acote el número de términos del las sumas en el lado derecho de (2.26) o (2.27), y compárelo, para $j$ pequeño, con el número $2^{|\mathscr{P}|}$ de términos de la suma para $j = |\mathscr{P}|$. Típicamente, cada cantidad $|A_P|$ se estima de manera aproximada, con un error, y por lo tanto el término de error total crece con el número de términos. Así, las desigualdades (2.26) y (2.27), tomadas juntas, pueden terminar dando un resultado más preciso que la igualdad (2.25).

**Ejercicio 2.11.** Discutamos algunos resultados generales, concretos y no triviales de cribas, utilizando la notación que acabamos de explicar.

Suponemos siempre que hay una constante $\kappa > 0$ (llamada la *dimensión* de la criba) tal que

$$\frac{V(w)}{V(z)} \leq K \left(\frac{\log z}{\log w}\right)^{\kappa} \tag{2.28}$$

para todo $1 < w \leq z$ y alguna constante $K$, donde $V(y) = \prod_{p \in \mathscr{P}: p \leq y} (1 - g(p))$, $g$ cumpliendo (2.24). Entonces tenemos el siguiente resultado.

**Lema 2.9** (Lema fundamental de las cribas)**.** *Sean dados* $\{a_n\}_{n \leq x}$, $\mathscr{P}$, $z$, $D$, $g$, $X$, $r_d$, $\kappa$ *con las propiedades mencionadas. Entonces, para* $s = (\log D)/\log z$,

$$\sum_{\substack{n \leq x \\ p|n \Rightarrow p \notin \mathscr{P}}} a_n = (1 + O_\kappa(K^{O(1)} e^{-s}))XV(z) + R, \tag{2.29}$$

*donde*

$$|R| \leq R(D) = \sum_{\substack{d \leq D \\ p|d \Rightarrow p \in \mathscr{P}}} |r_d|.$$

*Demostración.* Por la criba (no tan pura) de Brun (o varias otras). Ver, por ejemplo, [6, Cor. 6.10] o [7, §2.8]. □

Para el ejercicio siguiente, es suficiente saber que

$$\sum_{\substack{n \leq x \\ p|n \Rightarrow p \notin \mathscr{P}}} a_n \ll_\kappa K^{O(1)} XV(z) + R(D), \tag{2.30}$$

lo cual se deduce inmediatamente del Lema 2.9. A tal resultado se le puede llamar *criba de cota superior*. Se puede obtener (con definiciones del término de resto $R(D)$ ligeramente distintas) de muchos procedimientos de criba distintos, e.g., la criba de Selberg ([9, §6.4] o cualquier otra fuente), la cual da (2.30) sin la dependencia en $K$, bajo condiciones muy generales.

**Ejercicio 2.12.**

1. Sea $k \geq 2$. Veamos como utilizar una criba de cota superior para acotar el número de primos $p$ tal que $p + k$ también es primo.

2. Como práctica, utilizando (2.30) con $a_n = 1$ para $n \le x$, muestre que

$$\pi(x) \ll \frac{x}{\log x}.$$

Use la *fórmula de Mertens*

$$\prod_{p \le y} \left(1 - \frac{1}{p}\right) = \frac{c + o(1)}{\log y},$$

la cual se deduce del Corolario 2.6, pero es históricamente anterior al teorema de los números primos. (Sólo requiere el método de Chebyshev en su prueba; ver [9, §2.2]. Por cierto, la constante $c$ es igual a $e^{-\gamma}$, donde $\gamma$ es la constante de Euler.)

3. Sean $k, X \in \mathbb{Z}^+$, $x = X(X + k)$. Para $n \le x$, sea $a_n$ la sucesión indicatriz de los números $n = m(m + k)$ para algún $1 \le m \le X$. Muestre que la condición (2.24) se cumple para todo $d \ge 1$ libre de factores cuadrados con

$$g(d) = \prod_{\substack{p \mid d \\ p \mid k}} \frac{1}{p} \cdot \prod_{\substack{p \mid d \\ p \nmid k}} \frac{2}{p}$$

y $|r_d| \le \tau(d)$. Verifique también que (2.28) se cumple con $\kappa = 2$, $K = O(1)$ y $\mathscr{P} = \{p \le z : p \text{ primo}\}$, para $z$ arbitrario. Aquí

$$V(y) = \prod_{p \le y} (1 - g(p)).$$

4. Usando el Lema 2.9 (o (2.30)), deduzca que, para $Y > X > 0$ y $k \in \mathbb{Z}^+$ arbitrarios, el número de enteros $Y < m \le Y + X$ tales que $m(m + k)$ no tiene factores primos de tamaño $\le \sqrt{X}$ es

$$\sum_{\substack{n \le x \\ p \mid n \Rightarrow p > \sqrt{X}}} a_n \ll_\kappa XV(z) + \sqrt{X} \log(X) \ll C(k) \frac{X}{(\log X)^2},$$

con $C(k) = \prod_{p \mid k, p > 2} (1 - 1/p)/(1 - 2/p) \ll \prod_{p \mid k} (1 + 1/p)$. (Claro está, $\sum_{d \le D} \tau(d) = \sum_{d \le D} \sum_{l \mid d} 1 = \sum_{l \le D} \lfloor D/l \rfloor \ll D \log D$.) Concluya que el número de primos $Y < p \le Y + X$ tal que $p + k$ también es primo es

$$\ll C(k) \frac{X}{(\log X)^2}. \tag{2.31}$$

Se cree que la asintótica correcta es proporcional a $C(k)\, X/(\log X)^2$ para $k$ par; más precisamente, la asintótica sería

$$\prod_{p \nmid k} \frac{1 - 2/p}{(1 - 1/p)^2} \prod_{p \mid k} \frac{1}{1 - 1/p} \cdot \frac{X}{(\log X)^2}.$$

Este es un caso especial de la *primera conjetura de Hardy y Littlewood*. (Para $k$ impar, la asintótica es claramente 0, pues existe a lo más un primo $p$ tal que $p + k$ también es primo: $p = 2$.) No sabemos, empero, siquiera si el número de primos $p$ con $p + 2$ primo es infinito o no (*problema de los primos gemelos*). Lo mismo vale para $p$ y $p + k$, $k > 2$.

### 2.3. Estimaciones de valor medio.

**Lema 2.10.** ([16, Thm. 6.1])*; see also* [9, Thm. 9.1]) *Sean* $a_1, a_2, \dots, a_N \in \mathbb{C}$. *Entonces, para todo* $T \geq 0$,

$$\int_0^T \left| \sum_{n \leq N} a_n n^{it} \right|^2 dt = (T + O(N)) \sum_{n \leq N} |a_n|^2. \tag{2.32}$$

Esta es una cota que no deja de recordar a la *gran criba* (de la cual no precisaremos). Como en ciertas pruebas de la gran criba, es más sencillo probar una cota con un factor de log de más. Expliquemos la prueba así, comenzando con una demostración de esa cota más débil.

*Demostración.* Hagamos un intento directo e ingenuo. Expandiendo el cuadrado, vemos que

$$\int_0^T \left| \sum_{n \leq N} a_n n^{it} \right|^2 dt = \sum_{n_1, n_2 \leq N} a_{n_1} \overline{a_{n_2}} \int_0^T (n_1/n_2)^{it} dt. \tag{2.33}$$

Ahora bien, para todo $r > 0$,

$$\int_0^T r^{it} dt = \frac{r^{iT} - 1}{i \log r} = O\left(\frac{1}{\log r}\right). \tag{2.34}$$

Puesto que $\log r \geq (r-1)/(r+1)$ para $r \geq 1$ (como podemos verificar comparando derivadas), $\log(n_1/n_2) \geq (n_1 - n_2)/(n_1 + n_2) \geq (n_1 - n_2)/2N$ para $1 \leq n_2 \leq n_1 \leq N$, y así $\log(n_1/n_2) \geq |n_1 - n_2|/2N$ para $1 \leq n_1, n_2 \leq N$. Por lo tanto, la expresión en el lado derecho de (2.33) es

$$\begin{aligned} &T \sum_{\substack{n_1, n_2 \leq N \\ n_1 = n_2}} a_{n_1} \overline{a_{n_2}} + O\left( \sum_{\substack{n_1, n_2 \leq N \\ n_1 \neq n_2}} |a_{n_1} \overline{a_{n_2}}| \frac{N}{|n_1 - n_2|} \right) \\ &\quad = T \sum_{n \leq N} |a_n|^2 + O\left( N \sum_{\substack{n_1, n_2 \leq N \\ n_1 \neq n_2}} \frac{|a_{n_1}|^2 + |a_{n_2}|^2}{2} \frac{1}{|n_1 - n_2|} \right) \\ &\quad = (T + O(N \log N)) \sum_{n \leq N} |a_n|^2. \end{aligned} \tag{2.35}$$

Hemos obtenido casi lo que queríamos. Veamos ahora porqué hay un factor de $\log N$ espurio. La integral en (2.34) no es sino $\widehat{1_{[0,T]}}(-(\log r)/2\pi)$, donde $1_I$ es la función característica del intervalo $I$. Estamos en verdad en una situación muy similar a la de la sección §2.1, cuando examinamos la fórmula de Perron (2.5).

Como (2.34) nos muestra, la transformada de Fourier de una función $1_I(t)$ decae como $1/|t|$ cuando $t \to \pm\infty$, mientras que, como es bien conocido y fácil de probar, la transformada de Fourier de una función continua $f$ decae por lo menos tan rápidamente como $1/t^2$ (esto es, es $O(1/t^2)$ cuando $t \to \pm\infty$). El factor de $\log N$ viene de una suma $\sum_{n \leq N} 1/n$; nos lo ahorraremos si tenemos una suma de tipo $\sum_{n \leq N} 1/n^2$, que converge.

Por lo tanto, lo que hacemos es elegir una función $f$ continua tal que $f(t) \geq 1_{[0,T]}(t)$ para todo $t$, y, al mismo tiempo, $\int f(t)dt$ no es mucho mayor que $\int 1_{[0,T]}(t)dt = T$, y la derivada $f'$, aparte de estar definida en casi todas partes, no tiene grandes fluctuaciones (pues esto afectaría la constante en la cota $\widehat{f}(t) = O(1/t^2)$). Algo estándar es la función "trapecio" $f(t)$ igual a 0 cuando $t \leq -N$ o $t \geq T + N$, igual a 1 para $0 \leq t \leq T$, y lineal afín en $[-N, 0]$ y $[T, T + N]$ (igual a $1 + t/N$ en el primer intervalo, y a $1 - (t - T)/N$ en el segundo). (Es en verdad el mismo tipo de elección que en (2.7). Otras funciones continuas también servirían; ésta es simplemente conveniente.)

Como $f(t) \geq 1_{[0,T]}(t)$,

$$\begin{aligned} \int_0^T \left| \sum_{n \leq N} a_n n^{it} \right|^2 dt &\leq \int_0^T f(t) \left| \sum_{n \leq N} a_n n^{it} \right|^2 dt \\ &\leq \sum_{n_1, n_2 \leq N} a_{n_1} \overline{a_{n_2}} \widehat{f}\left( -\frac{\log n_1/n_2}{2\pi} \right). \end{aligned} \tag{2.36}$$

Ahora bien, para $f$ la función "trapecio", $\widehat{f}(0) = \int f(t)dt = T + 2N$, y (ejercicio 2.14)

$$\widehat{f}(x) = O\left( \frac{1}{Nx^2} \right).$$

Por lo tanto, el valor absoluto del lado derecho de (2.36) es

$$\begin{aligned} \sum_{n \leq N} |a_n|^2 (T + 2N) + O\left( \sum_{\substack{n_1, n_2 \leq N \\ n_1 \neq n_2}} \frac{|a_{n_1}|^2 + |a_{n_2}|^2}{2} \frac{N}{|n_1 - n_2|^2} \right) \\ = (T + O(N)) \sum_{n \leq N} |a_n|^2. \end{aligned} \tag{2.37}$$

□

Si en la parte izquierda de la igualdad (2.32) integrásemos en un subconjunto $\mathscr{T}$ de $[0, T]$ en vez de en el total $\mathscr{T} = [0, T]$, entonces parece razonable que se siguiera cumpliendo la desigualdad ($\ll$) sustituyendo en la parte derecha $T$ por algo menor. Esto es lo que demostraremos en la Proposición 2.12. El problema es que para conseguirlo no sirve la técnica de la prueba del Lema 2.10, ya que ésta dependía de la cancelación que se obtiene al integrar $(n_1/n_2)^{it}$ en $[0, T]$, y no está claro que se pueda obtener al integrar sobre subconjuntos $\mathscr{T}$ generales de $[0, T]$. Sin embargo, veremos que es posible sustituir dicha cancelación por la de las sumas $\sum_n n^{it}$ para $t$ fijo, que es lo que establecemos a continuación.

**Lema 2.11.** *Sea $f : \mathbb{R}^+ \to \mathbb{R}$ definido por $f(x) = 1$ para $0 < x \leq 1$, $f(x) = 2 - x$ para $1 < x \leq 2$, y $f(x) = 0$ de lo contrario. Entonces, para $x \geq 1$ y $t \in \mathbb{R}$ arbitrario,*

$$\sum_n f(n/x) n^{it} \ll \frac{x}{(1 + |t|)^2} + \sqrt{|t|} \log(1 + |t|). \tag{2.38}$$

Como de costumbre, la elección de $f$ no tiene nada de particular; cualquier $f$ doblemente diferenciable con $f''$ en $L^1$ nos daría una cota de la misma forma que (2.38).

*Demostración.* (O más bien dicho un comentario; la demostración está en los ejercicios.) En los ejercicios 2.15–2.19, para ser precisos. Se trata de una aplicación de la fórmula de sumación de Poisson, seguidas de cotas generales para integrales del tipo $\int_a^b \eta(x)e(\theta(x))dx$, $\eta$ contínua y $\theta$ en $C^1$. (Se trata de cotas superiores, no de asintóticas, para los dos casos principales – fase no estacionaria y fase estacionaria con $\theta'' \neq 0$.) Algunos notarán que las mismas cotas aparecen en el método de van der Corput. De hecho, los ejercicios están cercanos al tratamiento clásico de tal método ([17, §3.3] o [9, §8.3]), con las mejoras que son de esperarse dado que usamos un peso liso $f$. Si no se usa tal $f$, se obtiene un resultado ligeramente peor:

$$\sum_{n\leq x} n^{it} \ll \frac{x}{1+|t|} + \sqrt{|t|}\log(1+|t|). \tag{2.39}$$

□

*Esbozo de otra demostración.* Por el teorema de inversión de Mellin,

$$\sum_n f(n/x)n^{it} = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} x^s Mf(s)\zeta(s-it)ds$$

para $\sigma > 1$. Desplazamos la integral hacia la izquierda, pasando por polos en $s = 1$ y en $s = 0$. Las contribuciones de los polos dan los términos del lado derecho de (2.38).

(Aquí estamos utilizando el hecho que la función zeta tiene continuación analítica a todo el plano complejo, así como la cota estándar $|\zeta(it)| \ll \sqrt{t}\log t$, y cotas similares para $|\zeta(\sigma+it)|$, $\Re\sigma < 0$. Para obtener tales cotas, es necesario utilizar la ecuación funcional de $\zeta(s)$ (la cual también da la continuación analítica). Como la prueba usual de la ecuación funcional se basa sobre la formulación de Poisson, esta demostración y la anterior no son tan distintas como parecen a primera vista.) □

**Proposición 2.12** (Halász–Montgomery)**.** *Sean $T \geq 2$ y $\mathscr{T} \subset [-T, T]$ medible. Sean $a_1, a_2, \dots, a_N \in \mathbb{C}$. Entonces*

$$\int_{\mathscr{T}} \left|\sum_{n\leq N} a_n n^{it}\right|^2 dt \ll (N + |\mathscr{T}|\sqrt{T}\log T)\sum_{n\leq N}|a_n|^2. \tag{2.40}$$

El lector avisado notará que la prueba tiene gran similitud con una de las pruebas de la *gran criba* (distinta de las *pequeñas cribas* que consideramos en §2.2). La diferencia consistirá en que aplicamos el Lema 2.11 en vez de cotas sobre sumas exponenciales. Parte del procedimiento (utilización de un $f$ continuo) es como en la prueba del Lema 2.10 que acabamos de ver.

*Demostración.* Por el principio de dualidad (ejercicio 4), basta mostrar que, para $a \in L^2(\mathscr{T})$,

$$\sum_{n\leq N}\left|\int_{\mathscr{T}} a(t)n^{it}\right|^2 \ll (N + |\mathscr{T}|\sqrt{T}\log T)\int_{\mathscr{T}}|a(t)|^2 dt.$$

Claro está, para cualquier $f : [0,\infty) \to [0,\infty)$ con $f(x) \geq 1$ para $0 \leq x \leq 1$,

$$\sum_{n\leq N}\left|\int_{\mathscr{T}} a(t)n^{it}\right|^2 \leq \sum_n f\left(\frac{n}{N}\right)\left|\int_{\mathscr{T}} a(t)n^{it}\right|^2.$$

Expandiendo, vemos que

$$\sum_n f\left(\frac{n}{N}\right)\left|\int_{\mathscr{T}} a(t)n^{it}\right|^2 = \int_{\mathscr{T}}\int_{\mathscr{T}} \overline{a(t_1)}a(t_2)\sum_n f\left(\frac{n}{N}\right)n^{i(t_2-t_1)}dt_1dt_2. \tag{2.41}$$

Usando la desigualdad $|a_1a_2| \le (|a_1|^2+|a_2|^2)/2$ y el Lema 2.11, concluímos que el lado derecho de (2.41) es a lo más

$$\begin{aligned}\int_{\mathscr{T}} |a(t_1)|^2 \int_{\mathscr{T}} \sum_n f\left(\frac{n}{N}\right)n^{i(t_2-t_1)}dt_1dt_2 \\ \ll \int_{\mathscr{T}} |a(t_2)|^2 \int_{\mathscr{T}} \left(\frac{N}{(1+|t_1-t_2|)^2}+\sqrt{T}\log T\right)dt_1dt_2 \\ \ll \left(N+|\mathscr{T}|\sqrt{T}\log T\right)\int_{\mathscr{T}} |\overline{a(t_2)}|^2dt_2,\end{aligned} \tag{2.42}$$

que era lo que queríamos demostrar. □

*Ejercicios.*

**Ejercicio 2.13.** En este problema vamos a ver que es sencillo demostrar el Lema 2.11 en cierto rango para $t$. Demostrarlo para todo $t$ será más complicado, como veremos en los siguientes ejercicios.

1. Demuestre que podemos extender la regla del rectángulo (2.21) a funciones complejas $f : \mathbb{R} \to \mathbb{C}$ usando la desigualdad $|a| + |b| \le \sqrt{2}\sqrt{|a|^2+|b|^2}$.
2. Aplicando el apartado anterior, pruebe la desigualdad (2.39) en el rango $|t| \le \sqrt{x/\log x}$.
3. Demuestre el Lema 2.11 en el rango $|t| \le (x/\log x)^{1/3}$.

**Ejercicio 2.14.** Sean $T, D > 0$. Sea

$$f(t) = \begin{cases} 0 & \text{si } t \le -D \text{ o } t \ge T+D, \\ 1+t/D & \text{si } -D \le t \le 0, \\ 1 & \text{si } 0 \le t \le T, \\ 1-\frac{t-T}{D} & \text{si } T \le t \le T+D. \end{cases} \tag{2.43}$$

1. Muestre, por integración por partes, que, para $t \neq 0$,
$$\begin{aligned}\widehat{f}(t) &= \frac{1}{2\pi it}\left(\int_{-D}^0 + \int_0^T + \int_T^{T+D} f'(x)e(-xt)dx\right) \\ &= \frac{1}{2\pi iDt}\left(\int_{-D}^0 e(-xt) - \int_T^{T+D} e(-xt)dx\right).\end{aligned}$$
2. Concluya que, para $t \neq 0$,
$$|\widehat{f}(t)| \le \frac{1/D}{(\pi t)^2}.$$
3. En general, sea $f : \mathbb{R} \to \mathbb{C}$ tal que (a) $f(x), f'(x) \to 0$ cuando $x \to \pm\infty$, (b) $f''$ es continua e integrable (i.e., $|f''|_1 = \int_{-\infty}^{\infty} |f''(x)|dx$ es finita). Usando integración por partes dos veces, muestre que, para $t \neq 0$,
$$|\widehat{f}(t)| \le \frac{|f''|_1}{(2\pi t)^2}.$$

**Nota.** La condición que $f''$ sea continua puede ser reemplazada por condiciones bastante más débiles; basta con usar versiones más generales de la integración por partes (mediante integrales de Riemann-Stieltjes, o distribuciones, medidas, etc.). Basta con que $f$ sea continua, $f'$ sea definida fuera de un conjunto finito de puntos, $f'$ sea de *variación total* finita y $f(x), f'(x) \to 0$ cuando $x \to \pm\infty$. Un enunciado de esta generalidad cubre la función $f$ en (2.43).

**Ejercicio 2.15.** Sea $f : \mathbb{R} \to \mathbb{C}$ continua e integrable, así como diferenciable fuera de un conjunto finito de puntos. Asumamos también que $f'$ es integrable (en todo intervalo donde está definida) y que $\sum_{n\in\mathbb{Z}} |\widehat{f}(n)| < \infty$. Entonces

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \widehat{f}(n). \qquad \text{(fórmula de sumación de Poisson)} \tag{2.44}$$

Veamos cómo probar esta fórmula.

1. Muestre que la función $F : \mathbb{R} \to \mathbb{C}$ dada por $F(t) = \sum_{n\in\mathbb{Z}} f(t+n)$ está bien definida (es decir, la suma converge). Verifique también que $F$ es de período 1 y que $\int_0^1 |F(t)|dt < \infty$.
2. El teorema de inversión de Fourier (para funciones de período 1) nos dice que
$$F(t) = \sum_{n\in\mathbb{Z}} a_n e(nt),$$
donde $a_n = \int_0^1 F(t)e(-nt)dt$, bajo la condición que $\sum_n |a_n| < \infty$. Muestre que $a_n = \widehat{f}(n)$.
3. Concluya que
$$\sum_{n\in\mathbb{Z}} f(n) = F(0) = \sum_{n\in\mathbb{Z}} \widehat{f}(n).$$

**Ejercicio 2.16.** *Fase no estacionaria.* Sean $\theta, \eta : [a,b] \to \mathbb{R}$ tales que $\eta$ y $\theta'$ son continuas. Asumamos que $\theta'(x) \neq 0$ para todo $x \in [a,b]$. (Ésta es la suposición crucial, llamada "fase no estacionaria", pues $\theta(x)$ es la "fase".) Asumamos también que $g(x) = \eta(x)/\theta'(x)$ es monótona en $[a,b]$.

Muestre, por intregración por partes, que

$$\int_a^b \eta(x)e(\theta(x))dx = \frac{1}{2\pi i}\left(g(x)e(\theta(x))|_a^b + \int_a^b e(\theta(x))dg(x)\right).$$

Concluya que

$$\left|\int_a^b \eta(x)e(\theta(x))dx\right| \leq \frac{\text{máx}(|g(a)|, |g(b)|)}{\pi}. \tag{2.45}$$

En particular, para $\theta'$ monótona con $\theta'(x) \neq 0$ para todo $x \in [a,b]$,

$$\left|\int_a^b e(\theta(x))dx\right| \leq \frac{1/\pi}{\text{mín}(|\theta'(a)|, |\theta'(b)|)}. \tag{2.46}$$

Deduzca de (2.45) la siguiente variante: sea $\theta : [a,b] \to \mathbb{R}$ diferenciable, con $\theta'$ continua y $\theta'(x) \neq 0$ para todo $x \in [a,b]$, y $\eta_1, \eta_2 : [a,b] \to \mathbb{R}$ tales que $\eta_1(x)$ y

$\eta_2(x)/\theta'(x)$ son monótonas. Entonces, para $\eta = \eta_1 \cdot \eta_2$,

$$\left|\int_a^b \eta(x) e(\theta(x)) dx\right| \leq \frac{c}{\pi} \cdot \max\left(\frac{|\eta_2(a)|}{|\theta'(a)|}, \frac{|\eta_2(b)|}{|\theta'(b)|}\right) \tag{2.47}$$

con $c = \max(|\eta_1(a)|, |\eta_1(b)|, |\eta_1(a) - \eta_1(b)|)$. Es fácil obtener otras variantes: por ejemplo, si, en vez de $\eta_1$ monótona en $[a, b]$, tenemos $\eta_1$ monótona en $[a, x_0]$ y en $[x_0, b]$ para algún $x_0 \in [a, b]$, y $\eta_1(a) = \eta_1(b) = 0$, entonces (2.47) vale con $c = |\eta_1(x_0)|$.

**Ejercicio 2.17.** Sea $\theta : [a, b] \to \mathbb{R}$ doblemente diferenciable. Asumamos que $\theta''(x) \geq \rho > 0$ para todo $x \in [a, b]$.

Para $\delta > 0$ arbitrario, podemos tener $|\theta'(x)| < \delta$ sólo dentro de un intervalo $I$ de longitud $\leq 2\delta/\rho$. (¿Por qué?) Por lo tanto,

$$\int_I e(\theta(x)) dx \leq 2\delta/\rho.$$

(Ésta es la contribución de la región de *fase estacionaria*, es decir, de una vecindad del punto en el cual $\theta'(x) = 0$, si tal punto existe.) Aplique (2.46) para mostrar que

$$\int_{I \setminus [a,b]} e(\theta(x)) dx \leq 2/\pi\delta.$$

Escoja $\delta$ de manera óptima para así concluir que

$$\int_a^b e(\theta(x)) dx \leq \frac{2}{\sqrt{\pi\rho}}. \tag{2.48}$$

Aquí también podemos introducir un peso $\eta$. Por ejemplo, sea $\eta : [a, b] \to [0, \infty)$ continua tal que $\eta$ es monótona en $[a, x_0]$ y en $[x_0, b]$ para algún $x_0 \in [a, b]$, y además $\eta(a) = \eta(b) = 0$. Deduzca de (2.48) que

$$\int_a^b \eta(x) e(\theta(x)) dx \leq \frac{2 \max_{x \in [a,b]} |\eta(x)|}{\sqrt{\pi\rho}}. \tag{2.49}$$

**Ejercicio 2.18.** Sean $\theta : [a, b] \to \mathbb{R}$ diferenciable tal que $\theta'$ es monótona y $\theta'(x) \neq 0$ para todo $x \in [a, b]$. Sea $\eta : [a, b] \to [0, \infty)$ continua, así como diferenciable fuera de a lo más un conjunto finito de puntos; asumamos también que $\eta'$ es decreciente y acotada, y que $\eta(a) = \eta(b) = 0$.

Como el signo de $\theta'(x)$ es constante, y como podemos hacer un cambio de variables $x \mapsto b + a - x$ sin cambiar lo que suponemos sobre $\eta$, podemos asumir sin pérdida de generalidad que $\theta'$ es creciente y positiva en $[a, b]$. Sea $\alpha = \theta'(a)$.

Muestre que, por integración por partes,

$$\int_a^b \eta(x) e(\theta(x)) dx = -\frac{1}{\alpha} \int_a^b \left(\frac{\eta'(x)}{2\pi i} + \eta(x)(\theta'(x) - \alpha)\right) e(\theta(x)) dx. \tag{2.50}$$

Aplique (2.47) y obtenga que

$$\left|\int_a^b \left(\frac{\eta'(x)}{2\pi i} + \eta(x)(\theta'(x) - \alpha)\right) e(\theta(x)) dx\right| \leq \frac{\eta'(a) - \eta'(b)}{2\pi^2\alpha} + \frac{\max_{x \in [a,b]} |\eta(x)|}{\pi\alpha}.$$

Por lo tanto

$$\left|\int_a^b \eta(x) e(\theta(x)) dx\right| \leq \frac{c}{\min(|\theta'(a)|, |\theta'(b)|)^2} \tag{2.51}$$

donde $c = (\eta'(a) - \eta'(b))/2\pi^2 + (\max_{x\in[a,b]} |\eta(x)|)/\pi$.

**Ejercicio 2.19.** Sea $\eta : \mathbb{R} \to \mathbb{R}$ dada por $\eta(x) = 2x - 1$ para $1/2 \le x \le 1$, $\eta(x) = 2 - x$ para $1 \le x \le 2$ y $\eta(x) = 0$ cuando $x < 1/2$ o $x > 2$. Quisiéramos estimar

$$\sum_n \eta(n/X)x^{it}$$

para $t \neq 0$ arbitrario y $X > 1/2$. (Para $0 < X \le 1/2$, la suma es trivialmente igual a 0.)

Por la formula de sumación de Poisson (2.44),

$$\sum_n \eta(n/X)x^{it} = \sum_{n\in\mathbb{Z}} \int_0^\infty \eta(x/X)x^{it}e(-nx)dx,$$

asumiendo que la suma en el lado derecho converge absolutamente.

1. Estimemos primero la contribución de $n = 0$. Por integración por partes,
$$\int_0^\infty \eta(x/X)x^{it}dx = -\int_0^\infty \frac{\eta'(x/X)}{X}\frac{x^{it+1}}{it+1}dx.$$
Utilice integración por partes una vez más para concluir que
$$\int_0^\infty \eta(x/X)x^{it}dx \ll \frac{X}{(|t|+1)^2}.$$
2. Sea $n \in \mathbb{Z}$ tal que ya sea $|n| \ge 2t/\pi X$ o $n$ es de signo contrario a $t$. Use (2.51) para mostrar que
$$\int_0^\infty \eta(x/X)x^{it}e(-nx)dx \ll \frac{1}{n^2}. \tag{2.52}$$
3. Concluya que, para $X \ge 2|t|/\pi$,
$$\sum_n \eta(n/X)n^{it} \ll \frac{X}{(|t|+1)^2} + 1. \tag{2.53}$$
4. Sea ahora $X < 2|t|/\pi$. Muestre que (2.49) nos da que
$$\sum_n \eta(n/X)x^{it}e(-nx) \ll \frac{X}{\sqrt{|t|}}$$
para $n$ arbitrario, y en particular para $|n| \ge 2t/\pi X$. Concluya, usando también (2.52), que
$$\sum_n \eta(n/X)n^{it} \ll \sqrt{|t|}. \tag{2.54}$$
5. Sea $x \ge 1$. Sea $f_x$ la función $t \mapsto f(t/x)$, donde $f$ es como en el enunciado del Lema 2.11. Exprese $f_x$ como una suma de funciones del tipo $\eta(n/X)$ para diversos valores de $X$, y deduzca de (2.53) y (2.54) que
$$\sum_n f(n/x)n^{it} \ll \frac{x}{(|t|+1)^2} + \sqrt{|t|}(\log|t| + 1) + \log x$$
para $t \ge 1$. En otras palabras, el Lema 2.11 es cierto.

**Ejercicio 2.20.** *Dualidad.*

1. Sean $V$, $W$ espacios de Hilbert.[1] Definimos la norma $|A|$ de un operador (es decir, una función lineal) $A : V \to W$ por

$$|A| = \sup_{\substack{v \in V \\ v \neq 0}} \frac{|Av|_2}{|v|_2}.$$

Se dice que $A$ es acotado si su norma es finita. Todo $A$ acotado tiene un (único) *operador dual* $A^* : W \to V$, definido como la función lineal tal que

$$\langle A^* v, w \rangle = \langle v, Aw \rangle.$$

2. Sean $X$, $Y$ son dos espacios de Lebesgue de medida finita. Sea $V = L^2(X)$ y $W = L^2(Y)$. Definamos $A : V \to W$ por

$$(Av)(y) = \int_X K(x,y) v(x) dx,$$

donde $K : X \times Y \to \mathbb{C}$ es una función de imagen acotada. (Se comprende que $v \in V$ es una función $v : X \to \mathbb{C}$, y $Av \in W$ es una función $w : Y \to \mathbb{C}$.) Muestre que la función $A^* : W \to V$ definida por

$$(A^* w)(x) = \int_Y \overline{K(x,y)} w(y) dy$$

es el operador dual de $A$.

3. Sean $V$, $W$ espacios de Hilbert. Sea $A : V \to W$ un operador acotado. La desigualdad de Cauchy-Schwarz[2] nos dice que $\langle w_1, w_2 \rangle \leq |w_1|_2 |w_2|_2$ para $w_1, w_2 \in W$, con igualdad si $w_1 = \lambda w_2$ para algún $\lambda \in \mathbb{C}$. Deduzca que, para $v \in V$ arbitrario, $|Av|_2 = \sup_{w \in W : w \neq 0} \langle w, Av \rangle / |w|_2$. Usando este hecho, muestre que

$$\sup_{\substack{v \in V \\ v \neq 0}} \frac{|Av|_2}{|v|_2} = \sup_{\substack{v \in V \\ v \neq 0}} \sup_{\substack{w \in W \\ w \neq 0}} \frac{\langle w, Av \rangle}{|v|_2 |w|_2} = \sup_{\substack{w \in W \\ w \neq 0}} \sup_{\substack{v \in V \\ v \neq 0}} \frac{\langle A^* w, v \rangle}{|v|_2 |w|_2} = \sup_{\substack{w \in W \\ w \neq 0}} \frac{|A^* w|_2}{|w|_2}.$$

Concluya que

$$|A| = |A^*| \qquad \text{(principio de dualidad)}. \tag{2.55}$$

4. Sean $X$ e $Y$ dos espacios de Lebesgue de medida finita. Sea $K : X \times Y \to \mathbb{C}$ una función acotada y $C \geq 0$ una constante tal que, para todo $v \in L^2(X)$,

$$\int_Y \left| \int_X K(x,y) v(x) dx \right|^2 dy \leq C \int_X |v(x)|^2 dx.$$

Concluya que, para todo $w \in L^2(Y)$,

$$\int_X \left| \int_Y K(x,y) w(y) dy \right|^2 dx \leq C \int_Y |w(y)|^2 dy.$$

---

[1] Recordamos que un *espacio de Hilbert* es un espacio lineal $V$ sobre $\mathbb{R}$ o $\mathbb{C}$ con un producto escalar $\langle \cdot, \cdot \rangle$ tal que $V$ es completo con respecto a la distancia $d(x,y) := |x-y|_2 := \sqrt{\langle x-y, x-y \rangle}$. El único ejemplo que necesitaremos es el siguiente: sea $X$ un espacio de Lebesgue de medida finita; definamos el producto escalar $\langle v_1, v_2 \rangle = \int_X \overline{v_1(x)} v_2(x) dx$ para $v_1, v_2 : X \to \mathbb{C}$ medibles; entonces el espacio lineal $L^2(X)$ de funciones $v : X \to \mathbb{C}$ con $|v|_2 < \infty$ es un espacio de Hilbert.

[2] Ya sabemos que el nombre históricamente correcto es *Cauchy* o *Cauchy-Bunyakovsky-Schwarz*, pero todo el mundo sabe lo que *Cauchy-Schwarz* quiere decir.

## 3. Cancelación de $\lambda$ en intervalos cortos, en promedio

**3.1. Un primer tratamiento.** Queremos demostrar el Teorema 1.1. La demostración pasará por la función $Z_\lambda(s) = \zeta(2s)/\zeta(s)$, asociada con $\lambda$. En este contexto, es más fácil usar sumas donde los intervalos estén expresados en términos de $n/x$.

**Teorema 3.1.** *Sea $h = h(X)$ tal que $h(X) \to \infty$ cuando $X \to \infty$. Entonces, cuando $X \to \infty$,*

$$\mathbb{E}_{(1-\frac{h}{X})N<n\leq N}\lambda(n) = o_h(1)$$

*para todo entero $N \in [X, 2X]$ fuera de un conjunto de $o(X)$ excepciones.*

Aquí hemos usado la notación $\mathbb{E}_{n\in A} f(n) = \frac{1}{|A|}\sum_{n\in A} f(n)$. En verdad podemos leer

$$\mathbb{E}_{X<n\leq 2X}, \quad \mathbb{E}_{N<n\leq(1+\frac{h}{X})N}, \quad \text{etc.},$$

como

$$(1/X)\sum_{X<n\leq 2X}, \quad (1/(X/h))\sum_{N<n\leq(1+\frac{h}{X})N}, \quad \text{etc.},$$

ya que $|\{X < n \leq 2X\}| = X + O(1)$, y estamos trabajando asintóticamente.

Es sencillo ver que el Teorema 1.1 es una consecuencia del Teorema 3.1 (ejercicio 3.1), así que vamos a concentrarnos en demostrar este último. Es una consecuencia directa de la siguiente cota de varianza.

**Teorema 3.2.** *Sea $h = h(X)$ tal que $h(X) \to \infty$ cuando $X \to \infty$. Entonces, cuando $X \to \infty$,*

$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}\lambda(n)|^2 = o_h(1).$$

Nuestra tarea en esta sección será probar este teorema, con, por cierto, una cota precisa en vez de $o_h(1)$ (Teorema 3.17). Para empezar, veamos la relación entre las sumas $\mathbb{E}_{(1-\frac{h}{X})x<n\leq x} f(n)$ y $Z_f(s)$.

**Proposición 3.3.** *Sea $f : \mathbb{N} \to \mathbb{C}$ una función con soporte en $(X, 2X]$ tal que $|f(x)| \leq 1$ para todo $x$. Entonces si $x \in (X, 2X]$*

$$\mathbb{E}_{(1-\frac{h}{X})x<n\leq x} f(n) = \frac{1}{2\pi i}\int_{1-i\frac{X}{h\delta^2}}^{1+i\frac{X}{h\delta^2}} x^{s-1} M\psi_{\delta,h}(s) Z_f(s)\, ds + O(\delta)$$

*para cierta función $\psi_{\delta,h} : (0, +\infty) \to \mathbb{R}$ con $M\psi_{\delta,h}(s) \ll \min(1, \frac{X/h}{|s|}, \frac{X/h}{\delta|s|^2})$ para $\Re s > 0$.*

*Demostración.* La suma inicial puede escribirse como $\frac{1}{x}\sum_n f(n)\frac{1}{h/X}1_{(1-\frac{h}{X},1]}(\frac{n}{x})$. Lo que vamos a hacer es, como en el Lema 2.1, aproximar la función $\frac{1}{h/X}1_{(1-\frac{h}{X},1)}$ por una función $\psi_{\delta,h}$ que (a) tenga soporte en $[1-\frac{h}{X}, 1]$, (b) valga $\frac{1}{h/X}$ en $[1-(1-\delta)\frac{h}{X}, 1-\delta\frac{h}{X}]$, y entre 0 y $\frac{1}{h/X}$ en el resto del soporte. Podemos tomar una función "trapecio" similar a (2.7). Entonces

$$\psi_{\delta,h}(y) = \frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} M\psi_{\delta,h}(s) y^{-s} ds$$

con $M\psi_{\delta,h}$ cumpliendo las condiciones del enunciado. Por lo tanto,

$$\mathbb{E}_{(1-\frac{h}{X})x<n\leq x} f(n) = \frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} x^{s-1} M\psi_{\delta,h}(s) Z_f(s)\, ds + O(\delta).$$

Como $f$ es acotada y tiene soporte en $[X, 2X]$, tenemos que $|Z_f(1+it)| \ll 1$. Por el decaimiento de $M\psi_{h,\delta}$ podemos cortar la integral a altura $\delta^{-2}X/h$ perdiendo sólo $O(\delta)$. □

Ahora vamos a aplicar esta expresión para evaluar el promedio de sumas cortas de $f$.

**Proposición 3.4.** *Sea $f : \mathbb{N} \to \mathbb{C}$ una función con soporte en $(X, 2X]$ tal que $|f(x)| \leq 1$ para todo $x$. Entonces*

$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}f(n)|^2 \ll \int_{1-i\frac{X}{h\delta^2}}^{1+i\frac{X}{h\delta^2}} |Z_f(s)|^2 ds + O(\delta).$$

*Demostración.* Queremos usar la Proposición 3.3, expandir el cuadrado e intentar aprovechar la sumación en $x$. Para que esto funcione mejor, de nuevo es conveniente introducir una función suave que reemplaze a $1_{[X,2X]}$. En este caso

$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n<x}f(n)|^2 \ll \int \eta\left(\frac{x}{X}\right)|\mathbb{E}_{(1-\frac{h}{X})x<n<x}f(n)|^2\frac{dx}{x} \tag{3.1}$$

para cierta función $\eta$ con $0 \leq \eta(t) \leq 1$, soporte en $[1/2, 3]$, $C^2$ y tal que $\eta''$ es integrable. Ahora, usando la Proposición 3.3 y expandiendo el cuadrado tenemos que la parte derecha de (3.1) es

$$O(\delta) + \frac{1}{(2\pi)^2}\int_{1-i\frac{X}{h\delta^2}}^{1+i\frac{X}{h\delta^2}}\int_{1-i\frac{X}{h\delta^2}}^{1+i\frac{X}{h\delta^2}} (M\psi_{\delta,h})(s_1)\overline{M\psi_{\delta,h}(s_2)}Z_f(s_1)\overline{Z_f(s_2)}I(t_1,t_2)ds_1ds_2,$$

donde $t_i = \Re s_i$, y con

$$I(t_1,t_2) = \int \eta\left(\frac{x}{X}\right)x^{i(t_1-t_2)}\frac{dx}{x} = X^{i(t_1-t_2)}\int \eta(e^y)e^{i(t_1-t_2)y}\,dy.$$

Por integración por partes, $I(t_1,t_2) \ll (1+|t_1-t_2|)^{-2}$. (En general, para $f$ doblemente diferenciable tal que $f$ y $f''$ son integrables, tenemos $\widehat{f}(t) \ll (1+|t|)^{-2}$, por integración por partes.) Así, usando la cota $M\psi_{\delta,h} \ll 1$ y $|Z_f(s_1)Z_f(s_2)| \leq |Z_f(s_1)|^2 + |Z_f(s_2)|^2$ obtenemos el resultado. □

Una estimación de valor medio (Lema 2.10) permite mostrar que tenemos una desigualdad en la otra dirección. Esta desigualdad nos dice que no hemos perdido nada con respecto a la cota trivial.

**Proposición 3.5.** *Si $f$ tiene soporte en $(Y, 2Y]$, entonces*

$$\int_{1-iY}^{1+iY} |Z_f(s)|^2\,|ds| \ll \mathbb{E}_{Y<n\leq 2Y}|f(n)|^2.$$

*Demostración.* Por el Lema 2.10. □

La ganancia sobre esta cota trivial va a venir de factorizar $Z_f(s) = Z_{f_1}(s)Z_{f_2}(s)$.

**Proposición 3.6.** *Si $f_1$ es una función con soporte en $(Q, 2Q]$ y $f_2$ es una función con soporte en $\left(\frac{X}{Q}, \frac{2X}{Q}\right]$, entonces, para todo intervalo $\mathcal{T} \subset [1-i\frac{X}{Q}, 1+i\frac{X}{Q}]$,*

$$\int_{\mathcal{T}} |Z_{f_1}(s)Z_{f_2}(s)|^2|ds| \ll M_1^2\,\mathbb{E}_{\frac{X}{Q}<n\leq 2\frac{X}{Q}}|f_2(n)|^2$$

*con $M_1 = \max_{s\in\mathcal{T}}|Z_{f_1}(s)| \leq \max|f_1|$.*

*Demostración.* Sacamos el máximo fuera de la integral y aplicamos la proposición 3.5. □

Para usar el resultado anterior, queremos factorizar $Z_f(s)$. La idea es usar la factorización de un número $n$ en $n = pm$ con $p$ su primo más pequeño en cierto intervalo $P_0 < p < Q_0$. Para poder hacer esto primero necesitamos ver que la mayoría de números $n$ tienen algún primo en dicho intervalo, lo cual está asegurado por el Lema 2.7 si tomamos $P_0 = Q_0^\alpha$ grande con $\alpha$ pequeño. Ahora, si consideramos los números con primos en un intervalo $(P_0, Q_0]$, $P_0 = Q_0^\alpha$, y si sacamos el primo más pequeño en dicho intervalo tenemos la factorización

$$\sum_{\substack{X<n\le 2X\\ \exists p\in(P_0,Q_0],p|n}} f(n)n^{-s} = \sum_{P_0<p\le Q_0} f(p)p^{-s} \sum_{\substack{m\\ X<mp\le 2X\\ p'|m\Rightarrow p'\notin[P_0,p)}} f(m)m^{-s} \tag{3.2}$$

para $f$ totalmente multiplicativa (con $p'$ primo). El problema es que esta no es una factorización en dos funciones zeta, ya que en el sumatorio de dentro aparece la variable $p$ en dos condiciones.

Eliminar la dependencia en $p$ de la condición $X/p < m \le 2X/p$ es algo completamente de rutina. Eliminar la dependencia en $p$ de la condición $p' \notin (P_0, p)$ también es factible. Lo haremos de manera ligeramente distinta a [11] (o al artículo expositorio [24], el cual también da su propia versión).

Primero, un poco de notación. Para $\eta_1, \eta_2 : \mathbb{R}^+ \to \mathbb{C}$, denotaremos por $\eta_1 *_M \eta_2$ la convolución multiplicativa en $\mathbb{R}^+$ ("convolución de Mellin"):

$$(\eta_1 *_M \eta_2)(x) = \int_0^\infty \eta_1(x/t)\eta_2(t)\frac{dt}{t}.$$

Para $\upsilon_1, \upsilon_2 : \mathbb{Z}^+ \to \mathbb{C}$, denotaremos simplemente por $\upsilon_1 * \upsilon_2$ la convolución multiplicativa en $\mathbb{Z}^+$ ("convolución de Dirichlet"):

$$(\upsilon_1 * \upsilon_2)(n) = \sum_{d|n} \upsilon_1(n/d)\upsilon_2(d) = \sum_{d|n} \upsilon_1(d)\upsilon_2(n/d).$$

**Lema 3.7.** *Sean $X \ge 1$, $0 < \alpha, \beta, \delta < 1$, $Q_0 \le \exp((\log X)^{1-\beta})$ y $P_0 = Q_0^\alpha$. Definamos $u_X : \mathbb{Z}^+ \to \mathbb{R}$ por*

$$u_X(n) = \frac{1}{\log(1+\delta)} \int_{P_0}^{Q_0} \left(\upsilon_{1,Q,\delta} * \upsilon_{2,Q(1+\delta),X/Q}\right)(n) \frac{dQ}{Q}, \tag{3.3}$$

*donde*

$$\upsilon_{1,Q,\delta}(n) = \begin{cases} 1 & \text{si } n \text{ es primo y } Q < n \le (1+\delta)Q, \\ 0 & \text{de otra manera,} \end{cases} \tag{3.4}$$

$$\upsilon_{2,r,Y}(m) = \begin{cases} 1 & \text{si } m \in (Y, 2Y] \text{ y } p'|m \Rightarrow p' \notin [P_0, r), \\ 0 & \text{de otra manera.} \end{cases} \tag{3.5}$$

*Entonces existe un conjunto $Err \subset (X, 2(1+\delta)X]$ con*

$$\frac{|Err|}{X} \ll \alpha + \delta + \exp(-(\log P_0)^{\min(3/5,\beta)+o(1)})$$

*tal que $u_X(n) = 1_{(X,2X]}(n)$ para $n \notin Err$. Más aún, $0 \le u_X(n) \le 1$ para todo $n \in \mathbb{Z}^+$.*

*Demostración.* Por el Lema 2.7, existen a lo más

$$\alpha X + X \cdot O\left(\max\left(\exp(-(\log P_0)^{3/5+o(1)}), \exp\left(-\frac{\log X}{3\log Q_0}\right)\right)\right) + \frac{X}{P_0} \ll (\alpha + O(\exp(-(\log P_0)^{\min(3/5,\beta)+o(1)})))X$$

elementos de $(X, 2X]$ tales que no hay ningún primo $p \in (P_0, Q_0]$ tal que $p|n$. Incluímos todos estos elementos en el conjunto $Err$.

Es fácil ver que la diferencia

$$\eta_\Delta(x) = \left(1_{(X,2X]} - \frac{1_{(1,1+\delta]} *_M 1_{(X,2X]}}{\log(1+\delta)}\right)(x)$$

se desvanece cuando $x \le X$, $(1+\delta)X < x \le 2X$ o $x > 2(1+\delta)X$. Incluímos en $Err$, entonces, todos los elementos de $(X, (1+\delta)X]$ y $(2X, 2(1+\delta)X]$.

Por definición y un cambio de variables $t = mQ$, tenemos

$$\begin{aligned}\left(1_{(1,1+\delta]} *_M 1_{(X,2X]}\right)(pm) &= \int_0^\infty 1_{(1,1+\delta]}\left(\frac{p}{Q}\right) 1_{(X,2X]}(mQ)\frac{dQ}{Q} \\ &= \int_{(1-\delta)P_0}^{Q_0} 1_{(Q,(1+\delta)Q]}(p) 1_{(\frac{X}{Q}, \frac{2X}{Q}]}(m)\frac{dQ}{Q}\end{aligned}$$

para $P_0 < p \le Q_0$ y $m$ arbitrario. Así, para todo $n$ que no hayamos ya incluído en $Err$,

$$1_{(X,2X]}(n) = \frac{1}{\log(1+\delta)} \int_{(1-\delta)P_0}^{Q_0} \sum_{p|n} 1_{(Q,(1+\delta)Q]}(p)\upsilon_{2,p,X/Q}(n/p)\frac{dQ}{Q}, \tag{3.6}$$

donde $\upsilon_{2,r,Y}$ es como en (3.5). (Está claro que la suma $\sum_{p|n}$ puede tener a lo más un término no nulo, ya que $\upsilon_{2,p,X/Q}(n/p)$ se anulará a menos que $p$ sea el factor primo $\ge P_0$ más pequeño de $n$.)

Como el conjunto $S_2$ de enteros $X < n \le 2(1+\delta)X$ con por lo menos un factor primo en el rango $((1-\delta)P_0, (1+\delta)P_0]$ tiene a lo más

$$\begin{aligned}&\sum_{p\in((1-\delta)P_0,(1+\delta)P_0]} \left(\frac{(1+2\delta)X}{p} + O(1)\right) \\ &\qquad = (1+2\delta)X\log\frac{\log(1+\delta)P_0}{\log(1-\delta)P_0} + O\left((1+\delta)P_0\right) \ll \frac{\delta X}{\log P_0} + P_0\end{aligned}$$

elementos, podemos permitirnos cambiar el rango de la integral en (3.6) de $[(1-\delta)P_0, Q_0]$ a $[P_0, Q_0]$, añadiendo $O(\delta X/\log P_0 + P_0)$ elementos a $Err$.

Finalmente, debemos examinar lo que sucede cuando cambiamos $\upsilon_{2,p,X/Q}$ en (3.6) (con la integral de $P_0$ a $Q_0$) por $\upsilon_{2,Q(1+\delta),X/Q}$. Los únicos enteros afectados están en el conjunto $S_3$ de enteros $X < n \le 2(1+\delta)X$ divisibles por algún producto $pp'$ de dos primos $p$, $p'$ con $P_0 \le p \le Q_0$ y $p \le p' < (1+\delta)p$, y tales que ningún primo $P_0 \le p'' \le p$ divide $n$.

Ahora bien,

$$\begin{aligned}|S_3| &\leq \sum_{P_0\leq p\leq Q_0}\sum_{p_1<p_1'<(1+\delta)p_1} |\{X<n\leq 2(1+\delta)X : p_1, p_1'|n\}|\\ &\ll \sum_{P_0\leq p\leq Q_0}\sum_{p<p'<(1+\delta)p}\left(\frac{X}{pp'}+O(1)\right)\\ &\ll X\sum_{P_0\leq p\leq Q_0}\frac{1}{p}\left(\log\frac{\log(1+\delta)p}{\log p}+O(\exp(-(\log p)^{3/5+o(1)}))\right)+O(Q_0^2)\\ &\ll X\sum_{P_0\leq p\leq Q_0}\frac{\delta}{p\log p}+X\sum_{P_0\leq p\leq Q_0}\frac{1}{pe^{(\log p)^{3/5+o(1)}}}+Q_0^2\\ &\ll \frac{\delta X}{\log P_0}+X\exp(-(\log P_0)^{3/5+o(1)}),\end{aligned}$$

donde acotamos las dos sumas en la penúltima línea por el teorema de los números primos y sumación por partes. Incluímos $S_3$ en $Err$. Tenemos ahora

$$1_{(X,2X]}(n) = u_X(n)$$

para todo $n\notin Err$, donde

$$u_X(n) = \frac{1}{\log(1+\delta)}\int_{(1-\delta)P_0}^{Q_0}\sum_{p|n}1_{(Q,(1+\delta)Q]}(p)\upsilon_{2,Q(1+\delta),X/Q}(n/p)\frac{dQ}{Q}.$$

Por último, como $\upsilon_{2,Q(1+\delta),X/Q}(n/p)$ puede ser no nulo sólo cuando $p$ es el divisor primo $\geq P_0$ más pequeño de $n$, y $1_{(Q,(1+\delta)Q]}(p)\neq 0$ sólo cuando $Q$ está entre $p/(1+\delta)$ y $p$, vemos que, para cualquier $n$, $0\leq u_X(n)\leq 1$. □

**Corolario 3.8.** *Sea $f:\mathbb{N}\to\mathbb{C}$ una función completamente multiplicativa tal que $|f(n)|\leq 1$ para todo $n\in\mathbb{N}$. Entonces, para $0<\alpha,\beta,\delta<1$, $Q_0\leq\exp((\log X)^{1-\beta})$ y $P_0 = Q_0^\alpha$,*

$$Z_{f\cdot 1_{(X,2X]}} = Z_{err}+\frac{1}{\log(1+\delta)}\int_{P_0}^{Q_0} Z_{f_{1,Q,\delta}}Z_{f_{2,Q(1+\delta),X/Q}}\frac{dQ}{Q}$$

*donde*

$$f_{1,Q,\delta}(n) = \begin{cases} f(n) & \text{si } n \text{ es primo y } Q<n\leq(1+\delta)Q,\\ 0 & \text{de otra manera,}\end{cases}$$

$$f_{2,Q,Y}(m) = \begin{cases} f(m) & \text{si } m\in(Y,2Y] \text{ y } p'|m\Rightarrow p'\notin[P_0,Q),\\ 0 & \text{de otra manera,}\end{cases}$$

*y $err:\mathbb{N}\to\mathbb{C}$ es una función con soporte en $(X,2(1+\delta)X]$ tal que*

$$\frac{1}{X}\sum_{X<n\leq 2(1+\delta)X}|err(n)|^2\ll\alpha+\delta+\exp(-(\log P_0)^{\min(3/5,\beta)+o(1)}). \tag{3.7}$$

*Demostración.* Se sigue inmediatamente del Lema 3.7. □

Ahora que tenemos esencialmente una factorización de $Z_{\lambda 1_{[X,2X]}}$, la idea es mostrar que el máximo del factor

$$Z_{\lambda_{1,Q,\delta}}(1+it) = \sum_{Q<p\leq Q+\delta Q}\lambda(p)p^{-1-it} = -\sum_{Q<p<Q+\delta Q}p^{-1-it}.$$

es pequeño. En realidad esto no es siempre cierto, ya que para $t$ pequeña el signo de $\Re p^{-it}$ va a ser positivo, por lo cual no va a haber cancelación. Afortunadamente, para $t$ pequeña, ya sabemos que $Z_{\lambda 1_{[X,2X]}}(1+it)$ es pequeño por el Corolario 2.4. Cuando $t$ no es pequeña, la cancelación en la suma $\sum_{Q<p<Q+\delta Q} p^{-1-it}$ nos permite acotarla de la manera siguiente.

**Proposición 3.9.** *Para* $\exp((\log x)^a) \le t \le \exp((\log x)^{(3/2)(1-a)})$, $a > 0$, $0 < \delta \le 1$,

$$\sum_{x<p\le(1+\delta)x} p^{-1-it} \ll \exp(-(\log x)^{a+o(1)}).$$

Es, por cierto, necesario poner no sólo una cota inferior sobre $t$ como condición, sino también alguna cota superior, tal y como lo hemos hecho aquí. [3]

*Prueba de la Proposición 3.9.* Por sumación por partes es suficiente demostrar la desigualdad

$$\sum_{n\le x} \Lambda(n) n^{-it} \ll x \exp(-(1/2+o(1))(\log x)^a).$$

Como la función zeta correspondiente es $-\zeta'(s)/\zeta(s)$, esta desigualdad puede demostrarse como en la prueba del Teorema 2.5, pero evitando el polo mediante la elección $T = |t|/2$ (digamos), $\delta = 1/\sqrt{T}$. Usando las cotas en el Teorema 2.2, así como la cota $|M\psi_\delta(s)| \ll 1/s$, obtenemos

$$\begin{aligned}\sum_{n\le x} \Lambda(n) n^{-it} &\ll \frac{x(\log T)^{2/3+o(1)}}{\sqrt{T}} + x \exp\left(-(\log x)(\log T)^{-2/3+o(1)}\right)(\log T)^{2/3+o(1)} \\ &\ll \frac{x}{e^{(1+o(1))(\log x)^a/2}} + \frac{x}{e^{(\log x)^{a+o(1)}}} = \frac{x}{e^{(\log x)^{a+o(1)}}}.\end{aligned}$$

□

**Proposición 3.10.** *Sean* $Q_0 \le \exp((\log X)^{1-\beta})$, $P_0 = Q_0^\alpha$, $0 < \alpha, \beta < 1$. *Sean* $h \ge \delta^{-2} Q_0$ *y* $0 < \delta < 1$. *Entonces*

$$\mathbb{E}_{X<x\le 2X} |\mathbb{E}_{(1-\frac{h}{X})x<n\le x} \lambda(n)|^2 \ll \alpha + \frac{\eta^2}{\delta^2\alpha^2} + \delta + \exp(-(\log P_0)^{\min(3/5,\beta)+o(1)}),$$

*donde*

$$\eta = \max_{\substack{t\in[\exp(\sqrt{\log X}), X/h\delta^2] \\ P_0\le Q\le Q_0}} |Z_{\lambda_{1,Q}}(1+it)|. \tag{3.8}$$

[3]Para ver (de manera informal) la necesidad de una cota superior para obtener un resultado no trivial, notamos que tendremos cancelación en la suma $\sum_{x<p\le(1+\delta)x} p^{-1-it}$ sólo si el número complejo

$$p^{-it} = \exp(it \log p)$$

cambia de argumento lo suficiente. No es plausible que esto pase para todo $t$ grande, por el siguiente argumento heurístico. Deberíamos tener que $\{t \log p/2\pi\} < 1/8$ con probabilidad $1/8$, para $t$ tomado al azar en un intervalo grande. Si tales eventos probabilísticos (uno por primo) son aproximadamente independientes – lo cual es generalmente visto como plausible – entonces, con probabilidad $\gtrsim (1/8)^{\delta Q/\log Q}$, tendremos que $\{t \log p/2\pi\} < 1/8$ para todo $Q < p \le (1+\delta)Q$. Por tanto, debería haber algún $t \le 8^Q$ (digamos) para el cual este es el caso, lo cual implicaría que $\Re p^{-it} \ge 1/\sqrt{2}$ para todo $x < p \le (1+\delta)x$, y por lo tanto no habría suficiente cancelación: $\Re \sum_{x<p\le(1+\delta)x} p^{-1-it}$ sería $\ge (1/\sqrt{2})|\{x < p \le (1+\delta)x\}|$.

*Demostración.* Aplicamos la Proposición 3.4 con $f = \lambda 1_{[X,2X]}$. Luego aplicamos el Corolario 2.4 para concluir que $Z_{\lambda 1_{[X,2X]}}(1+it) \ll \exp(-(\log X)^{3/5})$ para $t \leq \exp(\sqrt{\log X})$ (digamos). Así, podemos quitar esa parte de la integral, y en el resto usamos el Corolario 3.8 para factorizar $Z_{\lambda 1_{[X,2X]}}$. Aplicando la Proposición 3.5 para controlar $Z_{err}$, llegamos a

$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n<x}\lambda(n)|^2 \ll \int_{1+i\exp(\sqrt{\log X})}^{1+i\frac{X}{h\delta^2}} \left|\frac{1}{\delta}\int_{P_0}^{Q_0} Z_{\lambda_{1,Q}} Z_{\lambda_{2,Q,X/Q}} \frac{dQ}{Q}\right|^2 ds$$

$$+ \alpha + \delta + \exp(-(\log P_0)^{\min(3/5,\beta)+o(1)}).$$

Por Cauchy-Schwarz y la Proposición 3.6,

$$\begin{aligned}
&\int_{1+i\exp(\sqrt{\log X})}^{1+i\frac{X}{h\delta^2}} \left|\frac{1}{\delta}\int_{P_0}^{Q_0} Z_{\lambda_{1,Q}} Z_{\lambda_{2,Q,X/Q}} \frac{dQ}{Q}\right|^2 ds \\
&\quad\leq \frac{1}{\delta^2}\left(\log\frac{Q_0}{P_0}\right)\int_{P_0}^{Q_0}\int_{1+i\exp(\sqrt{\log X})}^{1+i\frac{X}{h\delta^2}} \left|Z_{\lambda_{1,Q}} Z_{\lambda_{2,Q,X/Q}}\right|^2 ds\frac{dQ}{Q} \\
&\quad\leq \frac{\eta^2}{\delta^2}\left(\log\frac{Q_0}{P_0}\right)\int_{P_0}^{Q_0}\int_{1}^{1+i\frac{X}{h\delta^2}} \left|Z_{\lambda_{2,Q,X/Q}}\right|^2 ds\frac{dQ}{Q} \\
&\quad\leq \frac{\eta^2}{\delta^2}\left(\log\frac{Q_0}{P_0}\right)\int_{P_0}^{Q_0} \mathbb{E}_{\frac{X}{Q}<n\leq 2\frac{X}{Q}}|\lambda_{2,Q,X/Q}|^2\frac{dQ}{Q},
\end{aligned} \tag{3.9}$$

donde $\eta$ es como en (3.8). (En la última línea, estamos usando la suposición $h\delta^2 \geq Q$.) Usando la cota trivial $\mathbb{E}_{\frac{X}{Q}<n\leq 2\frac{X}{Q}}|\lambda_{2,Q,X/Q}|^2 \leq 1$, obtenemos que la expresión en la última línea de (3.9) es

$$\ll \frac{\eta^2}{\delta^2}\left(\log\frac{Q_0}{P_0}\right)^2 = \frac{\eta^2}{\delta^2\alpha^2}.$$

□

Usando este resultado y la Proposición 3.10 con $Q_0 = \delta^3 h$ y $\delta = (\log h)^{-1/2}$, podemos completar nuestro objetivo en cierto rango.

**Teorema 3.11.** *Para* $\exp((\log X)^{2/3+\epsilon}) \leq h \leq \exp((\log X)^{1-\epsilon})$, $0 < \epsilon \leq 1/6$, *tenemos*

$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n<x}\lambda(n)|^2 \ll_\epsilon \frac{(\log X)^{\frac{2}{3}+\frac{\epsilon}{2}}}{\log h} \leq \frac{1}{(\log X)^{\epsilon/2}}. \tag{3.10}$$

*Demostración.* Aplicaremos la Proposición 3.9 para acotar la cantidad $\eta$ definida en (3.8). Escogemos $a = 1-1/(1+3\epsilon/4)$, $\delta \geq h^{-1/3}$, $Q_0 = \delta^2 h$, $P_0 = \exp((\log X)^{2/3+\epsilon/2})$ y $\alpha = (\log P_0)/\log Q_0 = (\log X)^{2/3+\epsilon/2}/\log Q_0$. De esta manera,

$$(\log P_0)^{\frac{3}{2}(1-a)} = (\log X)^{\left(\frac{2}{3}+\frac{\epsilon}{2}\right)\cdot\frac{3}{2}(1-a)} = \log X.$$

Así, $t \leq X$ implica $t \leq \exp((\log P_0)^{(3/2)(1-a)})$. Como $\epsilon \leq 1/6$, vemos que $a \leq 1/9 < 1/2$, y por lo tanto $t \geq \exp(\sqrt{\log X})$ implica $t \geq \exp((\log Q_0)^a)$. Concluímos que sí podemos aplicar la Proposición 3.9, la cual nos da que

$$\eta \ll \exp(-(\log P_0)^{a+o(1)}) = \exp(-(\log X)^{(2/3+\epsilon/2)(a+o(1))}).$$

Resulta sensato escoger $\delta = \exp(-(\log X)^{2a/3})$. (La condición $\delta \geq h^{-1/3}$ se cumple para $X$ más grande que una constante, ya que $h \geq \exp((\log X)^{2/3+\epsilon})$ y $a \leq 1/9$.)

Ahora aplicamos la Proposición 3.10 con $\beta = \epsilon$, y obtenemos que

$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n<x}\lambda(n)|^2 \ll \frac{(\log X)^{\frac{2}{3}+\frac{\epsilon}{2}}}{\log h} + \frac{\exp(-(\log X)^{(2/3+\epsilon/2)(a+o(1))})}{\exp(-(\log X)^{2a/3})(\log X)^{-2/3}}$$
$$+ \exp(-(\log X)^{2a/3}) + \exp\left(-(\log X)^{2\epsilon/3+o(1)}\right) \ll_\epsilon \frac{(\log X)^{\frac{2}{3}+\frac{\epsilon}{2}}}{\log h}.$$

□

Es fácil deducir el resultado para $h > \exp((\log X)^{1-\epsilon})$ del Teorema 3.11 (ejercicio 3.3).

*Ejercicios.*

**Ejercicio 3.1.** Queremos demostrar que el Teorema 3.1 implica el Teorema 1.1.

1. Sea $|f| \leq 1$. Demuestre que para $h, Y > 1$, $0 < \delta < 1$ cualesquiera se cumple
$$\int_Y^{(1+\delta)Y} \left|\sum_{x<n\leq x+h} f(n)\right| dx = O(\delta)\delta Y h + \int_{Y+h}^{Y+h+\delta Y} \left|\sum_{(1-\frac{h}{Y+h})x<n\leq x} f(n)\right| dx.$$
2. Aplique el apartado anterior para demostrar que
$$\int_X^{2X} \left|\sum_{x<n\leq x+h} f(n)\right| dx \ll \delta X h + \frac{1}{\delta} \max_{X\leq X'\leq 3X} \int_{X'}^{2X'} \left|\sum_{(1-\frac{h}{X'})x<n\leq x} f(n)\right| dx.$$
3. Tomando el $\delta$ adecuado, concluya que el Teorema 3.1 implica el Teorema 1.1.

**Ejercicio 3.2.** Para cierto $B > 1$ y todo $h \geq X^{1/6}(\log X)^B$, vamos a ver una demostración del resultado

$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}\Lambda(n) - 1|^2 = o(1),$$

usando diferentes resultados conocidos sobre la función $\zeta(s)$. Se trata de un resultado clásico, conocido ya mucho antes de Matomäki y Radziwiłł.

Nótese que este resultado implica que hay primos en casi todos los intervalos de longitud $h$. Usando técnicas similares (pero un poco más complicadas) es posible demostrar lo mismo para $\lambda(n)$ en vez de $\Lambda(n) - 1$. Por otra parte: ¿por qué es que la demostración que hemos hecho en esta sección para $\lambda(n)$ no vale para $\Lambda(n) - 1$? (Hay, por otra parte, resultados sobre $\Lambda$ que pueden ser probados de manera indirecta usando los trabajos de Matomäki y Radziwiłł sobre $\lambda(n)$: ver [15, Thm. 1.3].)

1. Como en la prueba de la Proposición 3.3, demuestre que, si $x \in (X, 2X]$,
$$\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}\Lambda(n) = \frac{1}{2\pi i}\int_{1+\frac{1}{\log X}-i\frac{X}{h\delta^2}}^{1+\frac{1}{\log X}+i\frac{X}{h\delta^2}} x^{s-1} M\psi_{\delta,h}(s)\frac{-\zeta'(s)}{\zeta(s)}\, ds + O(\delta \log X)$$
para cierta función analítica $M\psi_{\delta,h}(s) \ll \min(1, \frac{X/h}{|s|}, \frac{X/h}{\delta|s|^2})$ para $\Re s > -1$, con $M\psi_{\delta,h}(1) = 1 + O(\delta)$.

2. En la zona $-1 < \Re s < 2$, los únicos ceros de la función $\zeta(s)$ están en $0 \leq \Re s \leq 1$, y si $N(T)$ es el número de ceros con $|\Im s| \leq T$, entonces $N(T) \sim \frac{T}{\pi} \log T$ y $N(T+1) - N(T) \ll \log T$. Además, si $|s-1| \gg 1$ y $s$ está a distancia $\epsilon$ del cero de $\zeta(s)$ más cercano, entonces $\frac{\zeta'(s)}{\zeta(s)} \ll \frac{\log(2+|s|)}{\min(1,\epsilon)}$. Usando esa información, el apartado anterior y el teorema de los residuos demuestre que, para algún $|w| \leq 1$,
$$\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}\Lambda(n) = 1 - \sum_\rho x^{\rho-1} M\psi_{\delta,h}(\rho) + O(\delta(\log X)^2),$$
donde $\rho$ son los ceros de $\zeta(s)$ que satisfacen $|\Im\rho| < \frac{X}{h\delta^2} + w$.
3. Usando la fórmula del apartado anterior y procediendo como en la demostración de la Proposición 3.4, demuestre que
$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}\Lambda(n) - 1|^2 \ll \log X \sum_\rho X^{-2(1-\Re\rho)} + O(\delta(\log X)^2).$$
4. Se sabe que el número $N(\sigma, T)$ de ceros de $\zeta(s)$ con $\sigma \leq \Re\rho \leq 1$ y $|\Im\rho| \leq T$ satisface $N(\sigma,T) \ll T^{\frac{12}{5}(1-\sigma)}(\log T)^A$, para cierto $A > 1$ (*teorema de densidad*). Use ese hecho para acotar $\sum_{\sigma\leq\Re\rho\leq\sigma+\frac{1}{\log X}} X^{-2(1-\Re\rho)}$ por cierta función $S(\sigma)$ creciente si $h\delta^2 > X^{1/6}$.
5. Usando los dos apartados anteriores y la región libre de ceros del Teorema 2.2, demuestre el resultado que se menciona al principio del problema. Observe también que si se cumpliera la hipótesis de Riemann, podríamos demostrarlo para $h > (\log X)^6$.

**Ejercicio 3.3.** Usando el Teorema 3.11, vamos a mostrar que, cuando $0 < \epsilon \leq 1/6$ y $\exp((\log X)^{1-\epsilon}) \leq h \leq X/2$,
$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}\lambda(n)|^2 \ll_\epsilon \frac{1}{(\log X)^{\frac{1}{3}-\epsilon}}.$$

1. Sea $j \in \mathbb{N}$. Si definimos $h_j$ por $1 - \frac{h_j}{X} = (1-\frac{h}{X})^{1/j}$, dividiendo la suma interior en $j$ trozos, demuestre que para $|f| \leq 1$ se cumple
$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n\leq x} f(n)|^2 \leq j^2 \max_{X'\in[X-h,X]} \mathbb{E}_{X'<x\leq 2X'}|\mathbb{E}_{(1-\frac{h_j}{X})x<n\leq x} f(n)|^2.$$
2. Usando el apartado anterior y el Teorema 3.11, demuestre la desigualdad del enunciado del problema.

**3.2. El caso general: valores excepcionales.** Cuando $h$ es más pequeño que un cierto nivel – digamos, $\exp((\log X)^{2/3})$ – nos encontramos con un obstáculo. No podemos hacer que $Q$ sea más pequeño que $\exp((\log x)^{2/3+\epsilon})$, pues estaríamos fuera del rango en el cual podemos aplicar la Proposición 3.9. Por otra parte, tampoco parecería que podemos tomar $h$ más pequeño que $Q$: el rango de integración en
$$\int_0^{X/h} |Z_{\lambda_{2,Q},X/Q}(1+it)|^2 dt$$
es demasiado grande como para que podamos aplicar la Proposición 3.6 (es decir, la estimación de valor medio que viene del Lema 2.10).

Existe la alternativa siguiente. Nuestra ganancia viene del factor $|Z_{\lambda_{1,Q}}|^2 \leq M_{1,Q}^2$ en (3.9). Podemos estudiar por separado los valores de $t$ para los cuales $|Z_{\lambda_{1,Q}}|$ es excepcionalmente grande – digamos $|Z_{\lambda_{1,Q}}| > Q^{-1/9}$. Si demostramos que tal cosa

sucede sólo para muy pocos valores de $s = 1 + it$, podremos usar el teorema del valor medio de Halász-Montgomery (Proposición 2.12) en vez del Lema 2.10 para acotar

$$M_{1,Q}^2 \int\limits_{\substack{0<t<X/h \\ |Z_{\lambda_{1,Q}}|>Q^{-1/9}}} |Z_{\lambda_{2,Q,X/Q}}(1+it)|^2 dt.$$

Mostremos, entonces, que el conjunto de valores de $t$ para los cuales $|Z_{\lambda_{1,Q}}(1+it)| > Q^{-1/9}$ es suficientemente pequeño.

**Lema 3.12.** *Sea $|f| \leq 1$ con soporte en los primos de $(Q, 2Q]$ y $(\log T)^9 < Q < T^{1/3}$. En ese caso, el conjunto de $t$ en $[0, T]$ tales que $|Z_f(1 + it)| > Q^{-1/9}$ tiene medida $O(T^{4/9})$.*

El exponente 4/9 no es óptimo, pero nos será suficiente.

*Demostración.* El teorema del valor medio (Lema 2.10) aplicado a $|Z_f(1 + it)|^2$ no controla bien los valores grandes de $|Z_f(1 + it)|$. Para conseguir tal control, es mejor usarlo sobre potencias superiores de $Z_f$. Lo aplicaremos a $Z_f(1 + it)^\ell$ con $\ell \in \mathbb{Z}^+$ tal que $Q^{\ell-1} < T \leq Q^\ell$. Así, como

$$Z_f(1 + it)^\ell = \sum_n a_n n^{-1-it}$$

donde $a_n$ tiene soporte en $(Q^\ell, (2Q)^\ell]$ y $|a_n| \leq \ell!$, por el Lema 2.10 tenemos

$$\int_0^T |Z_f(1 + it)^\ell|^2 = \int_0^T |\sum_{Q^\ell < n \leq (2Q)^\ell} a_n n^{-1-it}|^2 dt \ll (2Q)^\ell \sum_n \frac{|a_n|^2}{n^2}$$
$$\ll \ell! 2^\ell Q^{-\ell} (\sum_{Q<p<2Q} 1)^\ell \ll \ell! 2^\ell \ll \ell^\ell.$$

Así, si $A$ es el conjunto del enunciado, tenemos que $(Q^{-1/9})^{2\ell}|A| \ll \ell^\ell$. Como $Q > (\log T)^9$, sabemos que $\ell < Q^{1/9}$, y por lo tanto

$$|A| \ll (Q^\ell)^{3/9} \ll Q^{3/9} T^{3/9} \ll T^{4/9}.$$

□

En el siguiente resultado mostramos que podemos controlar la integral sobre los valores excepcionales de $Z_{\lambda 1_{(X,2X]}}$.

**Teorema 3.13.** *Sea $h \geq 1$ y $Q_0 = \exp((\log X)^\beta)$, donde $2/3 < \beta < 1$. Entonces, para $0 < \rho < 1 - \frac{2}{3\beta}$, $\alpha = 1/(\log Q_0)^\rho$ y $1/(\log X) \leq \delta \leq 1/(\log Q_0)^\rho$,*

$$\int_{[1,1+i\frac{X}{h}] \setminus \mathscr{T}_{\lambda, Q_0, \frac{1}{9}, \alpha, \delta}} |Z_{\lambda \cdot 1_{(X,2X]}}(s)|^2 ds \ll_{\beta,\rho} \frac{1}{(\log Q_0)^\rho}$$

*con $\mathscr{T}_{f,A,\gamma,\alpha,\delta}$ el conjunto de todos los valores de $s = 1 + it$, $0 \leq t \leq X/h$, tales que $|Z_{f_{1,Q,\delta}}(s)| \leq Q^{-\gamma}$ para todo $Q \in [A^\alpha, A] \cap \mathbb{Z}$, donde $f_{1,Q,\delta}(n) = f(n) \cdot 1_{(Q,(1+\delta)Q]}(n)$ si $n$ es primo y 0 en otro caso.*

*Demostración.* Comenzamos como en la demostración de la Proposición 3.10; es decir, aplicamos el Corolario 2.4 para concluir $Z_{\lambda 1_{(X,2X]}}(1+it) \ll \exp(-(\log X)^{3/5+o(1)})$ para $t \leq \exp(\sqrt{\log X})$ (digamos). Así, podemos quitar esa parte de la integral, y

en el resto usamos el Corolario 3.8 para factorizar $Z_{\lambda 1_{(X,2X]}}$. Aplicando (3.7) y la Proposición 3.5 para controlar $Z_{err}$, obtenemos

$$\int_{[1,1+iX/h]\setminus \mathscr{T}_{\lambda,Q_0,1/9,(\log Q_0)^{-\rho},(\log Q_0)^{-\rho'}}} |Z_{\lambda 1_{(X,2X]}}(s)|^2 ds$$
$$\ll O_\rho\left(\frac{1}{(\log Q_0)^\rho}\right) + \frac{1}{\delta^2}\int_B \left|\int_{Q_0^\alpha}^{Q_0} Z_{f_{1,Q,\delta}}(s) Z_{f_{2,Q(1+\delta),X/Q}}(s)\frac{dQ}{Q}\right|^2 ds$$

con $B = [1 + i\exp(\sqrt{\log X}), 1 + iX/h] \setminus \mathscr{T}_{\lambda,Q_0,1/9,\alpha,\delta}$. Como $\rho < 1$, el término de error $\exp\left(-(\log P_0)^{3/5+o(1)}\right)$ en (3.7) es absorbido por $O_\rho\left(1/(\log Q_0)^\rho\right)$. Por Cauchy-Schwarz, de manera similar a (3.9),

$$\int_B \left|\int_{Q_0^\alpha}^{Q_0} Z_{f_{1,Q,\delta}}(s) Z_{f_{2,Q(1+\delta),X/Q}}(s)\frac{dQ}{Q}\right|^2 ds \tag{3.11}$$
$$\leq (\log Q_0^{1-\alpha})^2 \max_Q \int_B \left|Z_{f_{1,Q,\delta}}(s) Z_{f_{2,Q(1+\delta),X/Q}}(s)\right|^2 ds,$$

donde el máximo de $Q$ se toma en el intervalo $[Q_0^\alpha, Q_0]$. (Después de aplicar Cauchy-Schwarz, invertimos el orden de las integrales, y luego reemplazamos la integral ahora exterior por un máximo.)

Ahora sacamos el máximo de $Z_{\lambda_{1,Q}}$ en $B$, aplicando la Proposición 3.9 con $a = 1 - 1/((3/2)\beta(1-\rho))$ (Nótese que $\beta(1-\rho)\cdot(3/2)\cdot(1-a) = 1$, por lo cual $\exp((\log Q)^{(3/2)(1-a)}) \geq X$). Obtenemos que

$$\int_B |Z_{\lambda_{1,Q,\delta}} Z_{\lambda_{2,(1+\delta)Q,X/Q}}|^2 ds \ll \exp(-2(\log X)^a) \int_B |Z_{\lambda_{2,Q(1+\delta),X/Q}}|^2 ds.$$

Finalmente, tenemos que $B = \cup_{Q'} B_{Q'}$ con $Q' \in [Q_0^\alpha, Q_0] \cap \mathbb{Z}$ y $B_{Q'}$ contenido en el conjunto de $t \in [0, X]$ tales que $|Z_{\lambda_{1,Q',\delta}}(1+it)| > Q'^{-1/9}$. Por el Lema 3.12 tenemos que $|B_{Q'}| \ll X^{4/9}$, lo cual implica que $|B| \ll Q_0 X^{4/9} = X^{4/9+o(1)}$. Luego, por Halász-Montgomery (Proposición 2.12), concluímos que

$$\int_B |Z_{\lambda_{2,Q(1+\delta),X/Q}}|^2 ds \ll \left(\frac{X}{Q} + X^{4/9+o(1)}\sqrt{X}\log X\right)\cdot \sum_{n\leq X/Q}(Q/X)^2 \ll 1,$$

lo que demuestra el teorema. □

*Ejercicios.*

**Ejercicio 3.4.** Sea $|f| \leq 1$, $X \geq 1$ y $\epsilon > 0$.

1. Use el Teorema del valor medio para demostrar que la medida del conjunto de $t \in [0, X]$ para los que $|\sum_{X\leq p\leq 2X} f(p)p^{-it}| > X^{1-\epsilon}$ es $O(X^{2\epsilon})$.
2. Use el Teorema del valor medio, junto con una elevación a una potencia (como en la prueba del Lema 3.12) para demostrar que, para todo $k \geq 1$, la medida del conjunto de $t \in [0, X^k]$ para los que $|\sum_{X\leq p\leq 2X} f(p)p^{-it}| > X^{1-\epsilon}$ es $O_k(X^{2k\epsilon})$.

**3.3. El caso general: valores típicos.** Es imposible controlar una integral

$$\int_{\mathscr{T}_{\lambda,Q_0,\gamma,\alpha,\delta}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds$$

factorizándola con un factor del tamaño $Q_0$, ya que $Q_0$ es mucho mayor que $h$. (Recuérdese que $\mathscr{T}_{\lambda,Q_0,\gamma,\alpha,\delta}$ es el conjunto de los valores "típicos" dentro del intervalo $[0, X/h]$.) La idea crucial ahora es usar que es posible sustituir dicha integral por otra del tipo

$$\int_{\mathscr{T}_{\lambda,q_0,\gamma-\epsilon,\alpha,\delta}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds,$$

con $q_0$ sustancialmente más pequeño que $Q_0$, si pagamos aumentado el tamaño que permitimos para el factor zeta correspondiente. Dicho resultado está basado en el siguiente lema, que de nuevo usa el teorema del valor medio.

**Lema 3.14.** *Sea* $0 < \varepsilon < \gamma < \frac{1}{2}$, $(\log Q)^{4/\varepsilon} < q < Q^{\varepsilon/4} < X^{\frac{1}{(\log\log X)^3}}$. *Sean* $f_q, f_Q, g_{X/Q} : \mathbb{Z} \to \mathbb{C}$ *funciones con* $|f_q(n)|, |f_Q(n)|, |g_{X/Q}(n)| \le 1$ *para todo* $n$. *Asumamos que* $g_{X/Q}$, $f_Q$ *y* $f_q$ *tienen soporte en* $[X/Q, 2X/Q]$, *en* $[Q, 2Q]$ *y en los primos de* $[q, 2q]$, *respectivamente. Entonces*

$$\int\limits_{\substack{s\in[1,1+iX]\\ |Z_{f_Q}(s)|\ll Q^{-\gamma}\\ |Z_{f_q}(s)|\ge q^{-\gamma+\varepsilon}}} |Z_{f_Q}(s) Z_{g_{X/Q}}(s)|^2 \, ds \ll Q^{-\varepsilon}.$$

*Demostración.* Podemos acotar la integral por

$$\ll Q^{-2\gamma} \int\limits_{\substack{s\in[1,1+iX]\\ |Z_{f_q}(s)|\ge q^{-\gamma+\varepsilon}}} |Z_{g_{X/Q}}(s)|^2 \, ds \ll Q^{-2\gamma} \int_1^{1+iX} \left| \left( \frac{Z_{f_q}(s)}{q^{-\gamma+\varepsilon}} \right)^{\ell} \right|^2 |Z_{g_{X/Q}}(s)|^2 \, ds.$$

Tomando $\ell$ natural tal que $q^\ell < Q \le q^{\ell+1}$, como $(q^{\gamma-\varepsilon})^{2\ell} Q^{-2\gamma} \le Q^{2\gamma-2\varepsilon} Q^{-2\gamma} = Q^{-2\varepsilon}$, vemos que la integral del enunciado es

$$\ll Q^{-2\varepsilon} \int_1^{1+iX} |Z_{f_q}^{\ell}(s) Z_{g_{X/Q}}(s)|^2 \, ds.$$

Ahora se trata de acotar la integral por el teorema del valor medio (Lema 2.10). Tenemos

$$\int_1^{1+iX} |Z_{f_q}^{\ell}(s) Z_{g_{X/Q}}(s)|^2 \, ds \ll 2^\ell X \sum_{X/q \le n \le 2^{\ell+1} X} \frac{a_n^2}{n^2}$$

con

$$a_n = \sum_{\substack{p_1 p_2 \dots p_\ell m = n\\ q \le p_j \le 2q\\ X/Q \le m \le 2X/Q}} 1 \le \ell! \sum_{\substack{rm=n\\ p|r \Rightarrow q \le p \le 2q}} 1 = \ell! \, w(n)$$

con $w(n)$ la función totalmente multiplicativa con $w(p^k) = 1$ si $p \notin [q, 2q]$ y $w(p^k) = k+1$ si $p \in [q, 2q]$. Como $Z_{w^2}(s) = \zeta(s) \prod_{q\le p\le 2q} \frac{1+4p^{-s}+9p^{-2s}+\dots}{1+p^{-s}+p^{-2s}+\dots}$, luego con residuo en $s = 1$ de la forma $\prod_{q\le p\le 2q}(1 + O(1/p)) \ll 1$, podemos proceder

como en la demostración del Lema 2.7 para demostrar que $\sum_{n<x} w^2(n) \ll x$ para $x \geq X/q$. Por lo tanto

$$\int_1^{1+iX} |Z_{f_q}^\ell(s) Z_{g_{X/Q}}(s)|^2 \, ds \ll q^2 2^{2\ell} (\ell!)^2 \ll q^2 \ell^{2\ell}.$$

Como $q < Q^{\varepsilon/4}$ y $\ell^{2\ell} \ll (\log Q)^{2\ell} \ll (q^{\varepsilon/4})^{2\ell} \ll Q^{\varepsilon/2}$, hemos acabado. □

Ahora veamos que efectivamente podemos sustituir la integral de $Z_f$ sobre $\mathscr{T}_{f,Q_0,c,\alpha,\delta}$ ("valores típicos con respecto a $Q_0$") por la integral de $Z_f$ sobre $\mathscr{T}_{f,q_0,c,\alpha,\delta}$ ("valores típicos con respecto a $q_0$"), donde $q_0$ es mucho más pequeño que $Q_0$.

**Proposición 3.15.** *Sea $f$ totalmente multiplicativa con $|f(n)| \leq 1$ para todo $n$. Sea $\mathscr{T}_{f,A,\gamma,\alpha,\delta}$ definido como en el Teorema 3.13. Sea $Q_0 \leq \exp((\log X)^{1-\epsilon})$, $\epsilon \in (0,1)$. Entonces, para $\gamma \in (1/\log\log Q_0, 1/2)$, $0 < \epsilon < 1$, $0 < \rho < 1$, $\rho < \kappa < 1 - \rho$, así como $\alpha \leq 1/(\log Q_0)^\rho$, $\alpha' = 1/(\log q_0)^{1-\epsilon}$, $1/(\log Q_0) \leq \delta \leq 1/(\log Q_0)^\rho$, $0 < \delta' \leq 1$ y $\log q_0 = (\log Q_0)^\kappa$, tenemos que*

$$\int_{\mathscr{T}_{f,Q_0,\gamma,\alpha,\delta}} |Z_{f_X}(s)|^2 ds \leq \int_{\mathscr{T}_{f,q_0,\gamma',\alpha',\delta'}} |Z_{f_X}(s)|^2 ds + O_{\rho,\kappa,\epsilon}\left(\frac{1}{(\log Q_0)^\rho}\right).$$

*con $f_X = f \cdot 1_{(X,2X]}$ y $\gamma' = \gamma - 1/\log\log Q_0$.*

Cualquier función del tipo $(\log Q_0)^{o(1)}$ podría usarse en vez de $\log\log Q_0$.

*Demostración.* Podemos suponer que $\alpha = 1/(\log Q_0)^\rho$, ya que $\alpha \leq \alpha_0 = 1/(\log Q_0)^\rho$ implica $\mathscr{T}_{f,Q_0,\gamma,\alpha,\delta} \subset \mathscr{T}_{f,Q_0,\gamma,\alpha_0,\delta}$.

Para comenzar,

$$\int_{\mathscr{T}_{f,Q_0,\gamma,\alpha,\delta}} |Z_{f_X}(s)|^2 ds \leq \int_{\mathscr{T}_{f,q_0,\gamma',\alpha',\delta'}} |Z_{f_X}(s)|^2 ds + \int_{\mathscr{T}_{f,Q_0,\gamma,\alpha,\delta} \setminus \mathscr{T}_{f,q_0,\gamma',\alpha',\delta'}} |Z_{f_X}(s)|^2 ds.$$

Apliquemos el Corolario 3.8 para factorizar $Z_{f_X}$, usando (3.7) y la Proposición 3.5 para controlar $Z_{err}$. Obtenemos, procediendo como en (3.9) o (3.11),

$$\int_{\mathscr{T}_{f,Q_0,\gamma,\alpha,\delta} \setminus \mathscr{T}_{f,q_0,\gamma',\alpha',\delta'}} |Z_{f_X}(s)|^2 ds \ll O_\rho\left(\frac{1}{(\log Q_0)^\rho}\right) + \frac{\left(\log Q_0^{1-\alpha}\right)^2}{\delta^2} \cdot \max_{Q \in [Q_0^\alpha, Q_0]} \int_{\mathscr{T}_{f,Q_0,\gamma,\alpha,\delta} \setminus \mathscr{T}_{f,q_0,\gamma',\alpha',\delta'}} \left|Z_{f_{1,Q,\delta}}(s) Z_{f_{2,Q(1+\delta),X/Q}}(s)\right|^2 ds.$$

Ahora bien, para todo $s \in \mathscr{T}_{f,Q_0,\gamma,\alpha,\delta}$, tenemos, por definición, que $|Z_{f_{1,Q,\delta}}(s)| < Q^{-\gamma}$ para todo $Q \in [Q_0^\alpha, Q_0] \cap \mathbb{Z}$, luego $|Z_{f_1,Q,\delta}| \leq \frac{2}{Q} + Q^{-\gamma} \ll Q^{-\gamma}$ para todo $Q \in [Q_0^\alpha, Q_0]$; más aún, si $s \notin \mathscr{T}_{f,q_0,\gamma',\alpha,\delta}$, entonces $|Z_{f_{1,q,\delta}}(s)| \geq q^{-\gamma'}$ para algún

$q \in [q_0^\alpha, q_0] \cap \mathbb{Z}$. Así, vemos que

$$\int\limits_{\mathscr{T}_{f,Q_0,\gamma,\alpha,\delta} \setminus \mathscr{T}_{f,q_0,\gamma',\alpha',\delta'}} \left|Z_{f_{1,Q,\delta}}(s) Z_{f_{2,Q(1+\delta),X/Q}}(s)\right|^2 ds$$

$$\leq 2q_0 \max_{q \in [q_0^\alpha, q_0]} \int\limits_{\substack{s \in [1,1+iX] \\ |Z_{f_{1,Q,\delta}}| \ll Q^{-\gamma} \\ |Z_{f_{1,q,\delta'}}| \geq q^{-\gamma'}}} |Z_{f_{1,Q,\delta}}(s) Z_{f_{2,Q(1+\delta),X/Q}}(s)|^2 ds,$$

donde recordamos que $\mathscr{T}_{f,Q_0,\gamma,\alpha,\delta} \subset [1, 1+iX/h] \subset [1, 1+iX]$. Aplicando el Lema 3.14, tenemos la cota

$$q_0 \int\limits_{\substack{s \in [1,1+iX] \\ |Z_{f_{1,Q,\delta}}| \ll Q^{-\gamma} \\ |Z_{f_{1,q,\delta'}}| \geq q^{-\gamma'}}} |Z_{f_{1,Q}}(s) Z_{f_{2,Q,X/Q}}(s)|^2 ds \ll q_0 Q^{-\varepsilon} \leq q_0 (Q_0^\alpha)^{-\varepsilon} < Q_0^{-3\alpha\varepsilon/4}$$

para $\gamma' = \gamma - \varepsilon$ y cualquier $\varepsilon \in (0, \gamma)$ tal que $(\log Q_0)^{4/\varepsilon} < q_0^\alpha$ y $q_0 < Q_0^{\varepsilon\alpha/4}$. Para $\varepsilon = 1/\log\log Q_0$, estas desigualdades se cumplen (debido a $\kappa + \rho < 1$), si asumimos que $Q_0$ es más grande que una constante $c$ que depende sólo de $\rho$ y $\kappa$. Bajo las mismas condiciones, $Q_0^{-3\alpha\varepsilon/4} < 1/(\log Q_0)^\rho$ (por un amplio margen). □

**Teorema 3.16.** *Supongamos que* $0 < \epsilon < 1/3$, $Q_0 = \exp((\log X)^{1-\epsilon})$ *y* $1 \leq h \leq \exp((\log X)^{2/3-\epsilon})$. *Entonces, para* $\alpha = \delta = 1/(\log Q_0)^\rho$, $1/6 < \rho < 1/3$,

$$\int_{\mathscr{T}_{\lambda,Q_0,\frac{1}{9},\alpha,\delta}} |Z_{\lambda \cdot 1_{(X,2X]}}(s)|^2 ds \ll_\epsilon \max\left(\frac{1}{(\log Q_0)^\rho}, \frac{1}{(\log h)^{1-\epsilon}}\right).$$

*Demostración.* Supongamos primero que $h \geq \exp(\sqrt{\log Q_0})$. Aplicamos la Proposición 3.15 una vez con $\rho$, $\gamma = 1/9$, $\delta' = 1/\log h$, $\alpha' = 1/(\log h)^{1-\epsilon}$ y $\kappa = (\log\log h)/(\log\log Q_0)$, de tal manera que $q_0 = h$. Es fácil ver que $1/2 \leq \kappa \leq (2/3 - \epsilon)/(1 - \epsilon) < 2/3$, así que $\rho < \kappa < 1 - \rho$. Obtenemos

$$\int_{\mathscr{T}_{\lambda,Q_0,\frac{1}{9},\alpha,\delta}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds \leq O_\epsilon\left(\frac{1}{(\log Q_0)^\rho}\right) + \int_{\mathscr{T}_{\lambda,h,\frac{1}{18},\alpha',\delta'}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds,$$

puesto que podemos suponer que $\gamma - 1/\log\log Q_0 \geq 1/18$.

Ahora utilizamos el Corolario 3.8 para factorizar $Z_{\lambda 1_{[X,2X]}}$ con $h$ en vez de $Q_0$, $\alpha = (\log h)^{-(1-\epsilon)}$. Por Cauchy-Schwarz, la Proposición 3.6 y la definición de $\mathscr{T}_{\lambda,h,\frac{1}{18},\alpha',\delta'}$, concluímos que

$$\int_{\mathscr{T}_{\lambda,h,\frac{1}{18},\alpha',\delta'}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds$$

$$\ll O_\epsilon\left(\frac{1}{(\log h)^{1-\epsilon}}\right) + (\log h)^4 (h^\alpha)^{-\frac{2}{18}} = O_\epsilon\left(\frac{1}{(\log h)^{1-\epsilon}}\right).$$

Supongamos ahora que $h < \exp(\sqrt{\log Q_0})$. Procederemos exactamente como antes, excepto que iteraremos el uso de la Proposición 3.15. Definimos $Q_{i+1} = \exp(\sqrt{\log Q_i})$ para $i = 0, 1, \ldots$, hasta que llegamos a un $i$ tal que $(\log h)^2 < \log Q_i \leq (\log h)^4$. Entonces definimos $m = i + 1$, $Q_m = h$.

Comenzamos aplicando la Proposición 3.15 con $\rho$, $\gamma = 1/9$, $\delta' = \delta_1 = 1/\log Q_1$, $\alpha' = \alpha_1 = 1/(\log Q_1)^{1-\epsilon}$, $\kappa = (\log\log Q_1)/(\log\log Q_0) = 1/2$ y los valores iniciales de $\alpha$ y $\delta$. Obtenemos

$$\int_{\mathscr{T}_{\lambda,Q_0,\frac{1}{9},\alpha,\delta}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds \leq O_\epsilon\left(\frac{1}{(\log Q_0)^\rho}\right) + \int_{\mathscr{T}_{\lambda,Q_1,\gamma_1,\alpha_1,\delta_1}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds,$$

donde $\gamma_1 = \gamma - 1/\log\log Q_0$. Luego aplicamos la Proposición 3.15 para $i = 1, 2, \dots, m-1$, con $\rho = 1/2 - \epsilon$, $\gamma = \gamma_i$, $\delta = \delta_i$, $\alpha = \alpha_i$, $\delta' = \delta_{i+1} = 1/\log Q_{i+1}$, $\alpha' = \alpha_{i+1} = 1/(\log Q_{i+1})^{1-\epsilon}$ y $\kappa = (\log\log Q_{i+1})/\log\log Q_i = 1/2$. Así,

$$\begin{aligned}\int_{\mathscr{T}_{\lambda,Q_i,\gamma_i,\alpha_i,\delta_i}} &|Z_{\lambda 1_{[X,2X]}}(s)|^2 ds \\ &\leq O_\epsilon\left(\frac{1}{(\log Q_i)^{1/2}}\right) + \int_{\mathscr{T}_{\lambda,Q_{i+1},\gamma_{i+1},\alpha_{i+1},\delta_{i+1}}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds,\end{aligned}$$

donde $\gamma_{i+1} = \gamma_i - 1/\log\log Q_i$. Está claro que $1/(\log Q_i)^{1/2} \leq 1/(\log Q_{i+1})$. Verificamos que

$$\sum_{i<m} \frac{1}{\log\log Q_i} \leq \frac{1}{\log\log Q_m} \sum_{k\geq 0} \frac{1}{2^k} = \frac{2}{\log\log h} < \frac{1}{18},$$

ya que podemos asumir que $h$ es más grande que una constante. Por lo tanto, $\gamma_m \geq 1/18$.

Terminamos usando el Corolario 3.8 exactamente como antes ($h$ en vez de $Q_0$, $\alpha = (\log h)^{-(1-\epsilon)}$). Por Cauchy-Schwarz y la Proposición 3.6, concluímos que

$$\begin{aligned}\int_{\mathscr{T}_{\lambda,Q_m,\gamma_m,\alpha_m,\delta_m}} &|Z_{\lambda 1_{[X,2X]}}(s)|^2 ds \\ &\ll O_\epsilon\left(\frac{1}{(\log h)^{1-\epsilon}}\right) + (\log h)^4 \left(\exp\left((\log h)^\epsilon\right)\right)^{-\frac{2}{18}} = O_\epsilon\left(\frac{1}{(\log h)^{1-\epsilon}}\right).\end{aligned}$$

□

Ahora ya podemos concluir nuestro resultado principal.

**Teorema 3.17.** *Sea $1 \leq h \leq X$. Entonces, para $\epsilon > 0$ arbitrariamente pequeño,*

$$\mathbb{E}_{X<x\leq 2X} |\mathbb{E}_{(1-\frac{h}{X})x<n\leq x} \lambda(n)|^2 \ll_\epsilon \text{máx}\left(\frac{1}{(\log X)^{\frac{1-\epsilon}{3}}}, \frac{1}{(\log h)^{1-\epsilon}}\right). \tag{3.12}$$

*Demostración.* Podemos asumir que $h \leq \exp((\log X)^{2/3-\epsilon})$ (como en el Ejercicio 3.3). Aplicando la Proposición 3.4 tenemos la cota

$$\mathbb{E}_{X<x\leq 2X} |\mathbb{E}_{(1-\frac{h}{X})x<n\leq x} \lambda(n)|^2 \ll \int_1^{1+i\frac{X}{h\delta^2}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds + O(\delta).$$

Tomemos $\delta = h^{-1/4}$ (digamos). El resultado (con un múltiplo constante de $\epsilon$ en vez de $\epsilon$, lo cual es claramente inofensivo) se sigue del Teorema 3.13 y el Teorema 3.16, aplicados con $h\delta^2$ en vez de $h$, $\alpha = \delta = 1/(\log Q_0)^\rho$, $Q_0 = \exp((\log X)^{1-\epsilon})$ y $\rho = 1 - \frac{2}{3(1-2\epsilon)}$. □

Es posible obtener una cota con $(\log\log h)/\log h$ en vez de $1/(\log h)^{1-\epsilon}$ [11]. Hemos optado por nuestra cota por simplicidad.

El Teorema 3.17 implica inmediatamente el Teorema 3.1, por un breve y conocido argumento (Chebyshev / cota de varianza) que ahora mostraremos. Escribamos la cota en (3.12) de la manera $\leq c_\epsilon/\operatorname{máx}(\log h, (\log X)^{1/3})^{1-\epsilon}$. Entonces, para todo $C$, la proporción de valores de $X < x \leq 2X$ tales que

$$\left|\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}\lambda(n)\right| \geq \frac{C\sqrt{c_\epsilon}}{\operatorname{máx}(\log h, (\log X)^{1/3})^{(1-\epsilon)/2}}$$

es a lo más $1/C^2$.

Si, alternativamente, siguiéramos [11], podríamos mostrar que, para $h$ en cierto rango, una cota más fuerte vale: la proporción de $x$ tales que

$$|\mathbb{E}_{(1-h/X)x<n\leq x}\lambda(n)| \geq C/(\log h)^{1-\epsilon} \tag{3.13}$$

es pequeña. La diferencia principal consiste en que [11] separa desde un principio el conjunto $\mathcal{S}$ de enteros $X < n \leq 2X$ con factores primos del tamaño deseado, y prueba una cota de varianza para ellos; luego reintroduce los enteros que quedaron fuera de $\mathcal{S}$ al último momento. De esta manera consigue mostrar que pocos $x$ satisfacen (3.13), aún si la forma de la cota de varianza es en esencia la misma.

*Ejercicios.*

**Ejercicio 3.5.** Sea $2 \leq h = o(X)$, $h \in \mathbb{N}$, $h \to \infty$.

1. Sea $f(n) = \cos(\frac{2\pi n}{100h})$.
   *a*) Demuestre que, si bien oscila en intervalos largos (i.e., $\sum_{X<n\leq 2X} f(n) = o(X)$), no oscila en intervalos cortos:
   $$|\sum_{x<n\leq x+2h} f(n)| \gg h$$
   para todo $x \in [X, 2X]$. Usando la Proposición 3.4, deduzca que
   $$\int_0^{cX/h} |\sum_{X<n\leq 2X} f(n)n^{-1-it}|^2\,dt \gg 1$$
   para alguna constante $c > 1$.
   *b*) Muestre que
   $$\sum_{X<n\leq u} f(n) \leq 100h$$
   para todo $u$. (Consejo: la función $f$ es periódica, con período 100.) Deduzca, por sumación por partes, que
   $$\sum_{X<n\leq 2X} f(n)n^{-1-it} \ll th/X$$
   para todo $t$.
   *c*) Deduzca que, para $h$ tal que $(\log h)^{150} \leq X/h$,
   $$\int_{(\log h)^{100}}^{cX/h} |\sum_{X<n\leq 2X} f(n)n^{-1-it}|^2\,dt \gg 1$$
   para alguna constante $c > 1$.

2. Sea $S$ el conjunto de números que son producto de dos primos, uno en $[P, 2P]$ y otro en $[X/P, 2X/P]$. Sea $g(n)$ la función multiplicativa tal que $g(p) = \cos(\frac{2\pi p}{100h})$. Asumiendo que $h = X^{1/10}$ y $P = \sqrt{h}$, demuestre que

$$\int\limits_{(\log h)^{100}}^{cX/h} |\sum_{n\in S} g(n)n^{-1-it}|^2\, dt \ll_c \frac{1}{(\log X)^{20}}.$$

para cualquier constante $c > 1$. Sugerencia: use el Teorema del valor medio (Lema 2.10) y la Proposición 3.9 (o más bien su prueba, pues consideramos valores de $t$ por debajo de lo que la Proposición 3.9 cubre).

**Ejercicio 3.6.** Sea $4 \le h \le (\frac{X}{2})^{1/5}$. Sea $S$ el conjunto de números que son producto de tres primos, uno en $[h, 2h]$, otro en $[h^2, 2h^2]$ y otro en $[X/h^3, 2X/h^3]$. Sea $f(n)$ una función multiplicativa tal que $Z(t) = \sum_{q\in[h^2,2h^2]} f(q)q^{-1-it} \ll (h^2)^{-1/9}$ para todo $t \in [(\log h)^{100}, X/h]$.

1. Muestre que

$$\sum_{n\in S} f(n)n^{-1-it} = \sum_{p\in[h,2h]} f(p)p^{-1-it} \sum_{q\in[h^2,2h^2]} f(q)q^{-1-it} \sum_{r\in[X/h^3,2X/h^3]} f(r)r^{-1-it}$$

con $p, q, r$ primos.

2. Use la cota para $Z(t)$ y el Teorema del valor medio (Lema 2.10) para probar que

$$\int\limits_{\substack{(\log h)^{100} \\ |\sum_{p\in[h,2h]} f(p)p^{-1-it}|>h^{-1/18}}}^{X/h} |\sum_{n\in S} f(n)n^{-1-it}|^2\, dt \ll \frac{1}{(h^{2/9})^2} \cdot \frac{1}{(h^{-1/18})^4} = \frac{1}{h^{2/9}}.$$

3. Use el Teorema del valor medio nuevamente para concluír que

$$\int\limits_{(\log h)^{100}}^{X/h} |\sum_{n\in S} f(n)n^{-1-it}|^2\, dt \ll \frac{1}{h^{2/9}}.$$

## 4. Coeficientes de Fourier de $\lambda$ en intervalos cortos, en promedio

Nuestra tarea en esta sección es acotar promedios de coeficientes de Fourier de $\lambda \cdot 1_{(x,x+h]}$:

$$\int_0^X \left| \sum_{x<n\le x+h} \lambda(n)e(\alpha n) \right| dx.$$

Al final de la sección, veremos como tales cotas implica que el promedio de expresiones como $\lambda(n)\lambda(n+c)$ tiende a 0 para una proporción que tiende a 1 de todos los $c$ en un pequeño intervalo ("Chowla en promedio"). Las líneas generales del argumento siguen las que se seguían para el mismo problema cuando $c$ varía en un intervalo grande, así como para problemas análogos. (El artículo [12] menciona las fuentes [5], [10], [18] y [2].)

El tratamiento será distinto dependiendo de si $\alpha \in \mathbb{R}$ está en un *arco mayor* o en un *arco menor*. (En verdad la definición depende sólo de $\alpha \bmod \mathbb{Z}$.) Tomemos $Q = \sqrt{h}$. Por el teorema de aproximación de Dirichlet (Ejercicio 4.1), existen

$a, q \in \mathbb{Z}^+$, $1 \le q \le Q$, $(a,q) = 1$, tales que $|\alpha - a/q| \le 1/qQ$. Si tales $a$, $q$ existen con $q \le R$ para cierto $R$ pequeño (será una potencia de $\log h$), decimos que $\alpha$ está en un *arco mayor*; de lo contrario, existen tales $a$, $q$ con $q > R$, y decimos que $\alpha$ está en un *arco menor*.

*Ejercicios.*

**Ejercicio 4.1.** (Teorema de aproximación de Dirichlet)

1. Sean dados $\alpha \in \mathbb{R}$ y $M \in \mathbb{Z}^+$. Muestre que existen $0 \le m_1 < m_2 \le M$ tales que $\alpha m_1$ y $\alpha m_2$ están a distancia $\le 1/(M+1)$ en $\mathbb{R}/\mathbb{Z}$ el uno del otro. Deduzca que $|(m_2 - m_1)\alpha - a| \le 1/(M+1)$ para algún $a \in \mathbb{Z}^+$.
2. Sean dados $\alpha \in \mathbb{R}$ y un real $Q \ge 1$. Concluya que existen $a, q \in \mathbb{Z}^+$, $1 \le q \le Q$, $(a,q) = 1$, tales que
$$\left|\alpha - \frac{a}{q}\right| \le \frac{1}{qQ}.$$
(Aplique la primera parte del ejercicio con $M = \lfloor Q \rfloor$.)

**Ejercicio 4.2.** En este problema vamos a discutir, a vuelo de pájaro, una variante relativamente sencilla del problema ternario de Goldbach, para ilustrar el uso de sumas exponenciales $\sum_n \lambda(n) e(\alpha n)$ similares a las que estudiaremos en esta sección (si bien en verdad más antiguas).

Para $N \in \mathbb{Z}^+$, consideraremos la suma
$$S_N = \sum_{\substack{a,b,c \ge 1 \\ a+b+c=N}} \lambda(a)\lambda(b)\lambda(c).$$
Si cambiáramos $\lambda$ por la función indicatriz de los primos estaríamos calculando el número de maneras de escribir $N$ como suma de tres primos. Nosotros queremos demostrar que $S_N = o(N^2)$, lo que nos diría que entre las tripletas de números que suman $N$ hay la misma probabilidad de tener un número par que uno impar de primos.

1. Demuestre la identidad $\int_0^1 e(j\alpha)\, d\alpha = 1_{j=0}$ para $j \in \mathbb{Z}$, con $e(x) = e^{2\pi i x}$. (Es la identidad sobre la cual se basa el análisis de Fourier sobre $\mathbb{R}/\mathbb{Z}$.)
2. Usando el apartado anterior, demuestre que
$$S_N = \int_0^1 (A_\lambda(\alpha))^3 e(-N\alpha)\, d\alpha,$$
con $A_\lambda(\alpha) = \sum_{1 \le n \le N} \lambda(n) e(n\alpha)$.
3. Es posible demostrar que $\max_{\alpha \in [0,1]} |A_\lambda(\alpha)| = o(N)$. (Ésta es, claro está, la parte difícil, la cual omitimos.) Usando esta cota, demuestre que
$$S_N \le o(N) \int_0^1 |A_\lambda(\alpha)|^2 d\alpha = o(N)N = o(N^2).$$
4. Sea $T_N = \sum_{a,b \ge 1 : a+b=N} \lambda(a)\lambda(b)$ la suma correspondiente a Goldbach con dos primos. Observe que si intentamos demostrar que $T_N = o(N)$ con las técnicas de los apartados anteriores, no funciona. ¿Por qué? Esto es parecido a lo que ocurre en la integral de la Proposición 3.4, y por eso allí fue necesario factorizarla.

### 4.1. Arcos menores.

**Proposición 4.1.** *Sean $X \geq 1$, $1 \leq h \leq X$, $1 \leq R \leq \log h$. Sea $\alpha \in \mathbb{R}$ tal que existen $a, q \in \mathbb{Z}^+$, $R \leq q \leq \sqrt{h}$, $(a,q) = 1$, tales que $|\alpha - \frac{a}{q}| \leq \frac{1}{q\sqrt{h}}$. Entonces*

$$\frac{1}{hX} \int_X^{2X} \left| \sum_{x < n \leq x+h} \lambda(n) e(\alpha n) \right| dx \ll \frac{(\log R)^3}{R^{1/4}}. \tag{4.1}$$

*Demostración.* Probar (4.1) es equivalente a probar que, para todo $\theta : \mathbb{R} \to \mathbb{C}$ medible, con soporte en $[X, 2X]$, y tal que $|\theta(x)| \leq 1$ para todo $x$,

$$\int_{\mathbb{R}} \theta(x) \sum_{x < n \leq x+h} \lambda(n) e(\alpha n) dx \ll hX \frac{(\log R)^3}{R^{1/4}}. \tag{4.2}$$

Sean $\beta = 1/2$, $\delta = R^{-1/4}$, $Q_0 = \exp(R^{1/4})$ y $P_0 = \exp((\log R)^3)$. Por el Lema 3.7, el lado izquierdo de (4.2) es igual a

$$O\left(\frac{(\log R)^3}{R^{1/4}}\right) \cdot hX + \int_{\mathbb{R}} \theta(x) \sum_{x < n \leq x+h} \lambda(n) u_X(n) e(\alpha n) dx, \tag{4.3}$$

donde $u_X$ es como en (3.3) (el error proviene de que cada valor de $n$ satisface $x < n \leq x + h$ sólo para $x$ dentro de un intervalo de largo $\leq h$). Entonces, por la definición (3.3) de $u_X$, es suficiente demostrar que

$$\frac{1}{\log(1+\delta)} \int_{P_0}^{Q_0} I(X, Q, \delta) \frac{dQ}{Q} \ll \frac{(\log R)^3}{R^{1/4}} \cdot hX, \tag{4.4}$$

donde

$$\begin{aligned} I(X,Q,\delta) &= \int_{\mathbb{R}} \theta(x) \sum_{x < n \leq x+h} \lambda(n) \left(\upsilon_{1,Q,\delta} * \upsilon_{2,Q(1+\delta),X/Q}\right)(n) e(\alpha n) dx \\ &= -\sum_m \upsilon_{2,Q(1+\delta),X/Q}(m) \lambda(m) \sum_{Q < p \leq (1+\delta)Q} e(\alpha m p) \int_{\mathbb{R}} \theta(x) 1_{(x,x+h]}(mp) dx, \end{aligned}$$

y $\upsilon_{1,Q,\delta}$, $\upsilon_{2,r,Y}$ están definidos como en (3.4) y (3.5). Por Cauchy-Schwarz,

$$I(X,Q,\delta)^2 \ll \frac{X}{Q} \sum_{X/Q < m \leq 2X/Q} \left| \sum_{Q < p \leq (1+\delta)Q} e(\alpha m p) \int_{\mathbb{R}} \theta(x) 1_{(x,x+h]}(mp) dx \right|^2, \tag{4.5}$$

ya que $\upsilon_{2,Q,(1+\delta),X/Q}$ tiene soporte en $[X/Q, 2X/Q]$ y $|\upsilon_{2,Q,(1+\delta),X/Q}(m)| \leq 1$ para todo $m$.

Ahora expandimos el cuadrado, y cambiamos el orden de la suma:[4]

$$I(X,Q,\delta)^2 \ll \frac{X}{Q} \sum_{Q < p_1, p_2 \leq (1+\delta)Q} \int_{\mathbb{R}} \int_{\mathbb{R}} \theta(x_1) \overline{\theta(x_2)} \sum_{\substack{X/Q < m \leq 2X/Q \\ \frac{x_i}{p_i} < m \leq \frac{x_i+h}{p_i}}} e(\alpha(p_1 - p_2)m) dx_1 dx_2.$$

[4]A partir de este momento, algunos lectores reconocerán ciertos paralelos con la versión de Linnik de la prueba del trabajo de Vinogradov sobre los tres primos.

La suma sobre $m$ es una suma sobre un intervalo $I$ contenido en $(x_1/p_1, x_1/p_1 + h/p_1] \subset (x_1/p_1, x_1/p_1+h/Q)$. Puede ser acotada trivialmente por $\leq |I|+1 \leq h/Q+1$; también vemos que se anula a menos que $(x_1/p_1, (x_1+h)/p_1] \cap (x_2/p_2, (x_2+h)/p_2] \neq \emptyset$, lo cual implica

$$\frac{p_2}{p_1}x_1 - h < x_2 < \frac{p_2}{p_1}x_1 + (1+\delta)h. \tag{4.6}$$

También sabemos que, para cualquier $\beta \in \mathbb{R}$ no entero y cualquier intervalo $I$ escrito en la forma $[m_0, m_1]$, $m_0, m_1 \in \mathbb{Z}$,

$$\sum_{m\in I} e(\beta m) = \frac{e(\beta(m_1+1)) - e(\beta m_0)}{e(\beta) - 1} \qquad \text{(suma geométrica)}$$

y, por lo tanto,

$$\left|\sum_{m\in I} e(\beta m)\right| \leq \frac{2}{|e(\beta)-1|} = \frac{1}{\sin \pi\beta} \leq \frac{2/\pi}{d(\beta, \mathbb{Z})}, \tag{4.7}$$

donde $d(\beta, \mathbb{Z})$ es la distancia entre $\beta$ y el entero más cercano. Por lo tanto,

$$\sum_{\substack{X/Q<m\leq 2X/Q \\ \frac{x_i}{p_i}<m\leq \frac{x_i+h}{p_i}}} e(\alpha(p_1-p_2)m) \ll \text{mín}\left(\frac{h}{Q}, \frac{1}{d(\alpha(p_1-p_2), \mathbb{Z})}\right),$$

y así, por la condición (4.6),

$$I(X,Q,\delta)^2 \ll \frac{X}{Q}Xh \sum_{Q<p_1,p_2\leq(1+\delta)Q} \text{mín}\left(\frac{h}{Q}, \frac{1}{d(\alpha(p_1-p_2), \mathbb{Z})}\right).$$

Ahora bien, para todo entero $l$, el número de parejas $Q \leq p_1, p_2 \leq Q + \delta Q$ tales que $p_1 - p_2 = l$ es $O(\delta Q)$. Vemos entonces que

$$I(Q,X,\delta)^2 \ll \delta X^2 h \sum_{\substack{l\in\mathbb{Z} \\ |l|\leq \delta Q}} \text{mín}\left(\frac{h}{Q}, \frac{1}{d(\alpha l, \mathbb{Z})}\right).$$

Ahora bien, por un lema de Vinogradov (ejercicio 4.3),

$$I(Q,X,\delta)^2 \ll \delta X^2 h \left(\frac{\delta Q}{q} + 1\right)\left(\frac{h}{Q} + q\log q\right).$$

Como $R \leq q \leq \sqrt{h}$, $\delta = R^{-1/4}$ y $\frac{R}{\delta} \ll P_0 \leq Q \leq Q_0 = \exp(R^{1/4}) \leq h^{1/4}$, tenemos que $q \log q \ll \frac{h}{Q}$ y

$$I(Q,X,\delta)^2 \ll \delta X^2 h \left(\frac{\delta h}{q} + \frac{h}{Q}\right) \ll \frac{\delta^2 X^2 h^2}{R}.$$

Por lo tanto

$$\frac{1}{\log(1+\delta)} \int_{P_0}^{Q_0} I(X,Q,\delta)\frac{dQ}{Q} \ll (\log Q_0)\,\frac{Xh}{R^{1/2}} = \frac{Xh}{R^{1/4}}$$

lo que demuestra (4.4). □

*Ejercicios.*

**Ejercicio 4.3.** (Lema de Vinogradov)

1. Para $a, q \in \mathbb{Z}^+$ coprimos y $n = n_0, n_0+1, \dots, n_0+q-1$, $n_0$ arbitrario, muestre que las fracciones $n/q \bmod \mathbb{Z}$ no son sino
$$0, 1/q, 2/q, \dots, (q-1)/q$$
en algún orden.
2. Sean $\alpha \in \mathbb{R}$, $a, q \in \mathbb{Z}^+$, $1 \le q \le Q$, $(a,q) = 1$, tales que $|\alpha - a/q| \le 1/qQ$. Muestre que, para $n = n_0, n_0+1, \dots, n_0+q-1$, $n_0$ arbitrario, las distancias $d(n\alpha, \mathbb{Z})$ son a lo más
$$0, 0, \frac{1/2}{q}, \frac{1}{q}, \frac{3/2}{q}, \dots, \frac{(q-2)/2}{q},$$
en algún orden.
3. Sean $\alpha \in \mathbb{R}$, $a, q \in \mathbb{Z}^+$, $1 \le q \le Q$, $(a,q) = 1$, tales que $|\alpha - a/q| \le 1/qQ$. Deduzca que, para $n_0$ y $X$ arbitrarios,
$$\sum_{n=n_0}^{n_0+q-1} \text{mín}\left(X, \frac{1}{d(n\alpha, \mathbb{Z})}\right) \le 2X + 2q(\log(q-2)+1) \ll X + q\log q.$$
Concluya que, para $N$ arbitrario,
$$\sum_{|n|\le N} \text{mín}\left(X, \frac{1}{d(n\alpha, \mathbb{Z})}\right) \ll \left(\frac{N}{q}+1\right)(X + q\log q).$$

**Ejercicio 4.4.** En este problema vamos a reproducir el esquema descrito en el ejercicio 4.2 de §4 para estimar la suma
$$S_N = \sum_{p+q+b=N} \log p \log q$$
que cuenta el número de maneras (ponderadas) de escribir $N$ como suma de dos primos y un número natural cualquiera. Esta es una versión muy sencilla del problema ternario de Goldbach, y podríamos estimarla sin introducir exponenciales $e(n\alpha)$, pero por una cuestión didáctica veamos que es posible hacerlo con ellas.

1. Procediendo como en el ejercicio 4.2 de §4, muestre que
$$S_N = \int_{-1/2}^{1/2} F(\alpha) G(\alpha)^2 e(-N\alpha)\, d\alpha$$
con $F(\alpha) = \sum_{b=1}^{N} e(b\alpha)$ y $G(\alpha) = \sum_{p\le N} \log p\, e(p\alpha)$.
2. En este caso los arcos menores van a ser la zona donde $F$ es pequeña. Tomemos como arcos menores $\mathfrak{m} = [-1/2, 1/2] \setminus (-\frac{1}{N\delta}, \frac{1}{N\delta})$, con $0 < \delta < 1/2$. Demuestre usando (4.7) que si $\alpha \in \mathfrak{m}$ entonces
$$|F(\alpha)| \ll \delta N.$$
3. Use el apartado anterior, el razonamiento expuesto en el ejercicio 3 de §4 y el TNP en la forma del Teorema 2.5 para demostrar que
$$\int_{\mathfrak{m}} F(\alpha) G(\alpha)^2 e(-N\alpha)\, d\alpha \ll \delta N^2 \log N.$$

En los ejercicios de la siguiente sección veremos que la integral sobre los arcos mayores $\mathfrak{M} = (-\frac{1}{N\delta}, \frac{1}{N\delta})$ da la contribución principal para $S_N$ (para cierto $\delta$).

**4.2. Arcos mayores y conclusión.** Examinemos ahora el caso de $\alpha = a/q$ con $q$ un entero pequeño. La idea es entonces que $e(\alpha n)$ es una función $q$-periódica. Así, lo que nos gustaría es sustituir $\lambda(n)e(\alpha n)$ por $\lambda(n)f(n)$ con $f(n)$ alguna función multiplicativa, para intentar usar las técnicas de la sección anterior. Esto es posible, usando análisis de Fourier para funciones $F : G \to \mathbb{C}$, con $G = (\mathbb{Z}/q\mathbb{Z})^\times$ el grupo (abeliano) de residuos módulo $q$ coprimos con $q$, con la operación de multiplicación. Dicho análisis permite expresar $F$ como suma de varios *carácteres módulo* $q$. Un carácter módulo $q$ no es sino un homomorfismo $\chi_* : G \to \mathbb{C}^\times$. Podemos extender tal homomorfismo a una función multiplicativa y $q$-periódica $\chi : \mathbb{Z} \to \mathbb{C}$, simplemente poniendo $\chi(n) = 0$ para $n$ no coprimo con $q$.

Nosotros sólo vamos a usar que $|\chi_*(g)| = 1$ y que los carácteres $\chi_*$ módulo $q$ forman una base ortonormal $\widehat{G}$ del espacio de funciones de $G = (\mathbb{Z}/q\mathbb{Z})^\times$ a $\mathbb{C}$. En verdad, $\widehat{G}$ es también es un grupo. Todos estos hechos son ciertos en general para grupos abelianos finitos $G$. Pueden, por cierto, ser probados fácilmente (ejercicio 4.5).

Si $\alpha$ está cerca de un racional con denominador pequeño, vamos a ver cómo pasar de sumas cortas de $\lambda(n)e(\alpha n)$ a sumas cortas de $\lambda(n)\chi(n)$ para algún carácter $\chi$ de módulo pequeño.

**Proposición 4.2.** *Sea* $1 < R^4 < h \le X$, $|\alpha - \frac{a}{q}| \le \frac{1}{q\sqrt{h}}$ *con* $q \le R$. *Entonces*

$$\frac{1}{hX}\int_X^{2X} \left|\sum_{x<n\le x+h} \lambda(n)e(\alpha n)\right| dx \ll \frac{1}{R} + R \max_{\chi, h', X'} \frac{1}{h'X'}\int_{X'}^{3X'} \left|\sum_{x<n\le x+h'} \lambda(n)\chi(n)\right| dx$$

*donde el máximo se toma sobre los carácteres* $\chi$ *de módulo menor o igual que* $R$ *y sobre* $X' \in [X/R, X]$, $h' \in [\sqrt{h}/R^2, \sqrt{h}/R]$.

*Demostración.* Comenzamos con la desigualdad

$$\frac{1}{hX}\int_X^{2X} \left|\sum_{x<n\le x+h} f(n)\right| dx \le O\left(\frac{h_1}{h}\right) + \frac{1}{h_1 X}\int_X^{3X} \left|\sum_{x<n\le x+h_1} f(n)\right| dx$$

para $1 \le h_1 \le h$, $|f| \le 1$, que simplemente proviene de dividir la suma interior en sumas de longitud $h_1$. La usamos con $f(n) = \lambda(n)e(\alpha n)$ y $h_1 = \sqrt{h}/R$. Ahora bien,

$$\left|\sum_{x<n\le x+h_1} \lambda(n)e(\alpha n)\right| = \left|\sum_{x<n\le x+h_1} \lambda(n)e\left(\frac{an}{q}\right) e\left(\left(\alpha - \frac{a}{q}\right)(n-x)\right)\right|$$

y como $|(\alpha - a/q)(n-x)| \le \frac{1}{q\sqrt{h}} h_1 \le R^{-1}$ tenemos que

$$|\sum_{x<n\le x+h_1} \lambda(n)e(\alpha n)| = \left|\sum_{x<n\le x+h_1} \lambda(n)e\left(\frac{a}{q}n\right)\right| + O(R^{-1})h_1.$$

Así

$$\frac{1}{hX}\int_X^{2X}\left|\sum_{x<n\leq x+h}\lambda(n)e(\alpha n)\right|dx \ll R^{-1}+\frac{1}{h_1X}\int_X^{3X}\left|\sum_{x<n\leq x+h_1}\lambda(n)e\left(\frac{a}{q}n\right)\right|dx.$$

Como $n\mapsto e(an/q)$ es $q$-periódica, podemos escribir (ejercicio 4.6)

$$e\left(\frac{a}{q}n\right)=\sum_{d|q}1_{d|n}\sum_{\chi\in(\mathbb{Z}/((q/d)\mathbb{Z}))^\times}a_\chi\chi(n/d),$$

donde $|a_\chi|\leq 1$. Podríamos tener una mejor cota para $a_\chi$ (*sumas de Gauss*), pero no nos importa. Obtenemos que

$$\begin{aligned}\frac{1}{X}\int_X^{3X}&\left|\sum_{x<n\leq x+h_1}\lambda(n)e\left(\frac{a}{q}n\right)\right|dx\\ &\leq\frac{1}{X}\sum_{d|q}\sum_{\chi\in(\mathbb{Z}/((q/d)\mathbb{Z}))^\times}\int_X^{3X}\left|\sum_{\frac{x}{d}<m\leq\frac{x}{d}+\frac{h_1}{d}}\lambda(m)\chi(m)\right|dx\\ &\leq q\max_{\substack{d|q\\(\mathbb{Z}/((q/d)\mathbb{Z}))^\times}}\frac{1}{X/d}\int_{X/d}^{3X/d}\left|\sum_{x'<m\leq x'+\frac{h_1}{d}}\lambda(m)\chi(m)\right|dx'.\end{aligned}$$

□

Con la proposición anterior vemos que sólo nos queda controlar el promedio de sumas cortas de $\lambda(n)\chi(n)$. Para ello, vamos a ver que funcionarían las técnicas de la sección anterior, igual que para $\lambda(n)$. Lo único que necesitaríamos usar en la prueba es cancelación para las sumas

$$\sum_{n<x}\lambda(n)\chi(n)\qquad\sum_{p<Q}\chi(p)p^{it_0},$$

y el control de dicha sumas depende de las funciones $Z_{\lambda\chi}(s)=Z_{\chi^2}(2s)/Z_\chi(s)$ y $Z'_\chi(s)/Z_\chi(s)$ con $Z_\chi(s)=L(s,\chi)$, donde $s\mapsto L(s,\chi)$ son las así llamadas funciones $L$ de Dirichlet:

$$L(s,\chi)=\sum_n\chi(n)n^{-s}.$$

Por lo tanto, necesitaremos control sobre los ceros y el tamaño de dichas funciones zeta. Para el caso de $\lambda$ usamos [9, Thm. 8.29] (que es nuestro Teorema 2.2) y en el lema anterior a dicho teorema puede verse que dicho control proviene de cotas superiores para $|L(s,\chi)|$ en dicha zona. Como

$$L(s,\chi)=q^{-s}\sum_{b(q)}\chi(b)\sum_{m=0}^{\infty}\left(m+\frac{b}{q}\right)^{-s}$$

y $\sum_{m=0}^{\infty}(m+\alpha)^{-s}$ es muy similar a $\zeta(s)$, esencialmente podemos conseguir para $|L(s,\chi)|$ las mismas cotas que para $|\zeta(s)|$, y así tenemos un equivalente al Teorema 2.2 para este caso.

**Teorema 4.3.** *Hay una constante $c > 0$ tal que $L(s,\chi) \neq 0$ para $s = \sigma + it$ con $\sigma \geq 1 - \frac{c}{\log q + (\log t)^{2/3}(\log\log t)^{1/3}}$, $|t| \geq 3$, y en dicha zona también se cumplen las cotas*

$$\frac{1}{L(s,\chi)} \ll \log q + (\log t)^{2/3}(\log\log t)^{1/3},$$
$$\frac{L'(s,\chi)}{L(s,\chi)} \ll \log q + (\log t)^{2/3}(\log\log t)^{1/3}.$$

*Además, en $|t| \leq 3$ se cumple que $L(s,\chi) \neq 0$ en $\sigma > 1 - \frac{c}{\sqrt{q}(\log q)^2}$ y en esa zona $\frac{1}{L(s,\chi)} \ll \sqrt{q}(\log q)^2$, $\frac{L'(s,\chi)}{L(s,\chi)} \ll \sqrt{q}(\log q)^2$.*

Las cotas para $|t| \leq 3$ son clásicas y provienen de otras técnicas [19, Capítulo 11]. Éstos resultados son todos efectivos; no esconden constantes no especificables – en particular, no lidiamos con los *ceros de Siegel*.

Para demostrar el Teorema 2.3 usamos las cotas del Teorema 2.2 con $t$ escogido de manera tal que $(\log t)^{2/3}(\log\log t)^{1/3}$ sea igual a $(\log x)^{2/5}(\log\log x)^{1/5}$, por lo que, si $\sqrt{q}(\log q)^2 \ll (\log x)^{2/5}(\log\log x)^{1/5}$, tendremos las mismas cotas en el Teorema 4.3, y luego serán ciertos los resultados análogos al Teorema 2.3 y al Corolario 2.4 para $\lambda(n)\chi(n)$. En particular

**Corolario 4.4.** *Sean $x \geq 1$, $t \leq e^{(\log x)^{3/5-\epsilon}}$, $\epsilon > 0$, y $\chi$ un carácter de módulo $q \leq (\log x)^{4/5-\epsilon}$. Entonces*

$$\sum_{x<n\leq 2x} \frac{\lambda(n)\chi(n)}{n^{1+it}} \ll \exp(-(\log x)^{3/5+o_\epsilon(1)}).$$

Además, en la zona $\log t > (\log Q)^a$ con $a > 0$, si $q \leq (\log Q)^2$ tenemos que las cotas del Teorema 4.3 serían las mismas que sin $\log q$, y luego el resultado análogo a la Proposición 3.9 también será cierto.

**Proposición 4.5.** *Sea $\chi$ un carácter de módulo $q \leq (\log Q)^2$. Para $\exp((\log Q)^a) \leq t_0 \leq \exp((\log Q)^{(3/2)(1-a)})$, $a > 0$, $0 < \delta \leq 1$,*

$$\sum_{Q<p\leq(1+\delta)Q} \chi(p)p^{-1-it_0} \ll \exp(-(\log Q)^{a+o(1)}).$$

Como en la demostración del Teorema 3.17 se usa el equivalente al Corolario 4.4 con $x \approx X$ y el equivalente a la Proposición 4.5 con $\log Q = (\log X)^{1-\epsilon}$, tenemos que

**Teorema 4.6.** *Sea $1 < h \leq X$, $\epsilon > 0$. Para $q \leq (\log X)^{4/5-\epsilon}$,*

$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}\lambda(n)\chi(n)|^2 \ll_\epsilon \frac{1}{(\log h)^{1-\epsilon}} + \frac{1}{(\log X)^{1/3-\epsilon}}.$$

**Corolario 4.7.** *Sea $1 < h \leq X$, $\epsilon > 0$. Para $q \leq (\log X)^{4/5-\epsilon}$,*

$$\mathbb{E}_{X<x\leq 2X}\,|\mathbb{E}_{x<n\leq x+h}\lambda(n)\chi(n)| \ll_\epsilon \frac{1}{(\log h)^{1/3-\epsilon}} + \frac{1}{(\log X)^{1/9-\epsilon}}.$$

*Demostración.* Notemos que es suficiente demostrar la desigualdad

$$\mathbb{E}_{X<x\leq 2X}\,|\mathbb{E}_{x-h<n\leq x}\lambda(n)\chi(n)| \ll_\epsilon \frac{1}{(\log h)^{1/3-\epsilon}} + \frac{1}{(\log X)^{1/9-\epsilon}}. \tag{4.8}$$

Ahora, para $0 < \delta \leq 1$ tenemos que

$$\mathbb{E}_{X<x\leq 2X} |\mathbb{E}_{x-h<n\leq x}\lambda(n)\chi(n)| \ll \mathbb{E}_{j:1<(1+\delta)^j\leq 2} \mathbb{E}_{\substack{Y<x\leq Y+\delta Y \\ Y=(1+\delta)^j X}} |\mathbb{E}_{x-h<n\leq x}\lambda(n)\chi(n)|$$

y como para cualquier $x \in (Y, Y+\delta Y]$ se cumple que

$$\mathbb{E}_{x-h<n\leq x}\lambda(n)\chi(n) = O(\delta) + \mathbb{E}_{x-h\frac{x}{Y}<n\leq x}\lambda(n)\chi(n)$$

deducimos que

$$\mathbb{E}_{X<x\leq 2X} |\mathbb{E}_{x-h<n\leq x}\lambda(n)\chi(n)| \ll \delta + \max_{Y\in[X,2X]} \mathbb{E}_{Y<x\leq Y+\delta Y}|\mathbb{E}_{x-h\frac{x}{Y}<n\leq x}\lambda(n)\chi(n)|.$$

Ahora aplicamos Cauchy-Schwarz:

$$\mathbb{E}_{Y<x\leq Y+\delta Y}|\mathbb{E}_{x-h\frac{x}{Y}<n\leq x}\lambda(n)\chi(n)| \ll \sqrt{\mathbb{E}_{Y<x\leq Y+\delta Y}|\mathbb{E}_{x-h\frac{x}{Y}<n\leq x}\lambda(n)\chi(n)|^2},$$

y luego, por el Teorema 4.6 (acotando brutalmente la integral de una cantidad no negativa sobre $(Y, Y+\delta Y]$ por la integral de la misma cantidad sobre $(Y, 2Y]$) vemos que

$$\mathbb{E}_{X<x\leq 2X} |\mathbb{E}_{x-h<n\leq x}\lambda(n)\chi(n)| \ll \delta + \max_{Y\in[X,2X]} \sqrt{\delta^{-1}\left(\frac{1}{(\log h)^{1-\epsilon}} + \frac{1}{(\log Y)^{1/3-\epsilon}}\right)}.$$

Finalmente, tomando

$$\delta = \min\left(1, \max\left(\frac{1}{(\log h)^{1/3-\epsilon}}, \frac{1}{(\log X)^{1/9-\epsilon}}\right)\right)$$

demostramos (4.8).

□

Usando los resultados anteriores, obtenemos lo siguiente.

**Proposición 4.8.** *Sea* $1 < R^5 \leq h \leq X$, $|\alpha - \frac{a}{q}| \leq \frac{1}{q\sqrt{h}}$ *con* $q \leq R \leq (\log X)^{4/5-\epsilon}$. *Entonces, para* $\epsilon > 0$,

$$\frac{1}{hX}\int_X^{2X} \left|\sum_{x<n\leq x+h} \lambda(n)e(\alpha n)\right| dx \ll_\epsilon \frac{1}{R} + \frac{R}{(\log h)^{\frac{1}{3}-\epsilon}} + \frac{R}{(\log X)^{\frac{1}{9}-\epsilon}}.$$

*Demostración.* Por la Proposición 4.2 y el Corolario 4.7. □

Finalmente, usando el Teorema de aproximación de Dirichlet y tomando $R$ igual a $\min((\log h)^{1/3}, (\log X)^{1/9})^{4/5}$ en la Proposiciones 4.1 y 4.8, concluímos que

**Teorema 4.9.** *Sean* $X > 1$, $1 < h \leq X$. *Entonces, para todo* $\alpha \in \mathbb{R}$ *y* $\epsilon > 0$,

$$\frac{1}{hX}\int_X^{2X} \left|\sum_{x<n\leq x+h} \lambda(n)e(\alpha n)\right| dx \ll_\epsilon \frac{1}{(\log h)^{\frac{1}{15}-\epsilon}} + \frac{1}{(\log X)^{\frac{1}{45}-\epsilon}}.$$

Los exponentes 1/15, 1/45 no son de ninguna manera óptimos; el lector interesado puede divertirse mejorándolos.

Vamos a ver como el teorema anterior implica la conjetura de Chowla en promedio (Teorema 1.2). Para ello, veamos la relación entre las sumas cortas de $f(n)$ y

sumas largas de $\overline{f(n)}f(n+h)$. Supongamos que $f(n)$ tiene soporte finito. Entonces, expandiendo el cuadrado y cambiando el orden de sumación obtenemos

$$\sum_x |\sum_{x<n\leq x+h} f(n)|^2 = \sum_{|n-m|<h} f(m)\overline{f(n)}(h-|n-m|).$$

Por tanto, escribiendo $m = n+j$, tenemos que

$$\sum_x |\sum_{x<n\leq x+h} f(n)|^2 = \sum_{|j|<h} (h-|j|)\sum_n f(n+j)\overline{f(n)}.$$

Ahí vemos directamente que las sumas cortas de $f(n)$ controlan un promedio de sumas largas de $\overline{f(n)}f(n+j)$. El problema es que podría ser que hubiera cancelación en la suma en $j$ en vez de en la suma en $n$. Para evitarlo, usamos la misma identidad con $f(n) = f_0(n)e(n\alpha)$, que da

$$\sum_x |\sum_{x<n\leq x+h} f_0(n)e(n\alpha)|^2 = \sum_{|j|<h} \left[(h-|j|)\sum_n f_0(n+j)\overline{f_0(n)}\right] e(j\alpha).$$

Ahora, teniendo en cuenta la identidad de Parseval

$$\int_0^1 |\sum_j a_j e(j\alpha)|^2 d\alpha = \sum_j |a_j|^2, \tag{4.9}$$

la cual simplemente proviene de que $\int_0^1 e(m\alpha)\,d\alpha = 1_{m=0}$, tenemos que

$$\sum_{|j|<h} |(h-|j|)|^2 \left|\sum_n f_0(n+j)\overline{f_0(n)}\right|^2 = \int_0^1 \left|\sum_x |\sum_{x<n\leq x+h} f_0(n)e(n\alpha)|^2\right|^2 d\alpha.$$

Así, si

$$M = \max_\alpha \sum_x |\sum_{x<n\leq x+h} f_0(n)e(n\alpha)|^2$$

entonces sacándolo fuera de la integral tenemos

$$\sum_{|j|<h} |(h-|j|)|^2 \left|\sum_n f_0(n+j)\overline{f_0(n)}\right|^2 \leq M\sum_x \int_0^1 |\sum_{x<n\leq x+h} f_0(n)e(n\alpha)|^2\,d\alpha,$$

y usando de nuevo (4.9) sobre la parte derecha concluímos que

$$\sum_{|j|<h} |(h-|j|)|^2 \left|\sum_n f_0(n+j)\overline{f_0(n)}\right|^2 \leq M\sum_x \sum_{x<n\leq x+h} |f_0(n)|^2.$$

Finalmente, aplicando esta desigualdad con $f_0(n) = \lambda(n)1_{(X,2X]}(n)$ y usando el Teorema 4.9, obtenemos

**Corolario 4.10.** *Para $1 \leq h \leq X$ y para todo $\epsilon > 0$*

$$\frac{1}{hX^2}\sum_{j\leq h/2} |\sum_{X<n,n+j\leq 2X} \lambda(n+j)\lambda(n)|^2 \ll_\epsilon \frac{1}{(\log h)^{\frac{1}{15}-\epsilon}} + \frac{1}{(\log X)^{\frac{1}{45}-\epsilon}}.$$

A partir de ahí es muy sencillo eliminar la condición $X < n+j \leq 2X$, y así obtener el Teorema 1.2 en el caso de dos factores ($k = 2$). El caso de más factores se sigue del siguiente lema (ejercicio 4.7)

**Lema 4.11.** *Sea $1 \leq H \leq X$. Para funciones $f, g : \mathbb{Z}^+ \to \mathbb{C}$ cualesquiera con $|g(n)| \leq 1$, $|f(n)| \leq 1$ para todo $n$ y soporte en $[1, X]$, se cumple la desigualdad*

$$\frac{1}{HX^2} \sum_{h \leq H} |\sum_n f(n+h)g(n)|^2 \leq \sqrt{\frac{1}{HX^2} \sum_{|h| \leq H} |\sum_n f(n+h)\overline{f(n)}|^2}.$$

*Reflexiones finales.* Como hemos visto, una vez que se tiene una cierta cota sobre sumas exponenciales en promedio (Teorema 4.9) podemos obtener el resultado de tipo "Chowla en promedio" que deseábamos (Teorema 1.2) con bastante facilidad.

Es bueno reflexionar sobre que tipo de cotas sobre sumas exponenciales necesitamos para obtener otros resultados sobre las autocorrelaciones de $\lambda$. ¿Qué pasa si deseamos saber el promedio de $\lambda(n)\lambda(n+h)\lambda(n+2h)$ para la mayor parte de valores de $h$ en un intervalo pequeño, digamos?

Muchas tales preguntas se reducen a acotar *seminormas de Gowers* $|\lambda|_{U^k}$. Ya sería un muy buen comienzo, por ejemplo, poder acotar la *segunda norma de Gowers* $|\lambda|_{U^2}$, definida por

$$|\lambda|_{U^2}^2 = \frac{1}{H^2 X} \sum_{1 \leq h_1, h_2 \leq H} \sum_{1 \leq n \leq X} \lambda(n)\lambda(n+h_1)\lambda(n+h_2)\lambda(n+h_1+h_2) \tag{4.10}$$

para todo $H = H(X) \to \infty$. En efecto, si pudiéramos mostrar que $|\lambda|_{U^2} = o(1)$, tendríamos que

$$\frac{1}{HX} \sum_{h \leq H} \sum_{n \leq X} \lambda(n)\lambda(n+h)\lambda(n+2h) = o(1)$$

despues de algunas aplicaciones de Cauchy-Schwarz. Más aún, si pudiéramos mostrar que la tercera norma de Gowers $|\lambda|_{U^3}$ es $o(1)$, podríamos deducir, de manera similar, que

$$\frac{1}{HX^2} \sum_{h \leq H} \left| \sum_{n \leq X} \lambda(n)\lambda(n+h)\lambda(n+2h) \right|^2 = o(1).$$

Acotar (4.10) por $o(1)$ resulta ser equivalente a probar la *conjetura de uniformidad de Fourier*

$$\frac{1}{hX} \int_0^X \sup_{\alpha \in \mathbb{R}/\mathbb{Z}} \left| \sum_{x < n \leq x+h} \lambda(n) e(\alpha n) \right| dx = o(1)$$

para $h = h(N) \to \infty$. Esta conjetura, planteada por primera vez en [12], parece más difícil que el Teorema 4.9, en el cual el orden de la integral y del supremo $\sup_{\alpha \in \mathbb{R}/\mathbb{Z}}$ (implícito en el Teorema 4.9) es el inverso.

*Ejercicios.*

**Ejercicio 4.5.** Sea $G$ un grupo abeliano finito.

1. Pruebe que, para todo carácter $\chi$ de $G$ y todo $g \in G$, $|\chi(g)| = 1$. (Sugerencia: muestre que $\chi(g)^{|G|} = 1$.)
2. Sea $\chi$ un carácter no trivial de $G$, es decir, un carácter tal que existe un $g \in G$ para el cual $\chi(g) \neq 1$. Muestre que
$$\psi(g) \sum_{h \in G} \psi(h) = \sum_{h \in G} \psi(gh) = \sum_{h \in G} \psi(h).$$
Concluya que $\sum_{h \in G} \psi(h) = 0$.

3. Sean $\chi$, $\chi'$ dos carácteres distintos de $G$. Muestre que $\psi = \overline{\chi}\cdot\chi'$ es un carácter no trivial. Entonces, por (2), $\sum_{g\in G}\overline{\chi}(g)\chi'(g) = 0$.
4. Muestre que hay $|G|$ carácteres distintos $\chi : G \to \mathbb{C}$. Concluya que los carácteres de $G$ forman una base ortonormal del espacio de funciones de $G$ a $\mathbb{C}$ con el producto escalar

$$\langle f_1, f_2\rangle = \frac{1}{|G|}\sum_{g\in G}\overline{f_1}(g)f_2(g).$$

**Ejercicio 4.6.**

1. Muestre que, para toda función $q$-periódica $f : \mathbb{Z} \to \mathbb{C}$ con soporte en los enteros coprimos con $q$,

$$f(n) = \sum_{\chi\in\widehat{G}_q}\widehat{f}(\chi)\chi(n),$$

donde $G_q = (\mathbb{Z}/q\mathbb{Z})^\times$, $\widehat{G}_q$ es el grupo de carácteres de $G$ (ejercicio 4.5) y

$$\widehat{f}(\chi) = \frac{1}{|G|}\sum_{m\in(\mathbb{Z}/q\mathbb{Z})^\times}\overline{\chi(m)}f(m).$$

Para este propósito, utilice el hecho que $\widehat{G}_q$ es una base ortonormal (ejercicio 4). Está claro que, si $|f(m)| \le 1$ para todo $m$, $|\widehat{f}(\chi)| \le 1$.

2. Para toda función $q$-periódica $f : \mathbb{Z} \to \mathbb{C}$,

$$f(n) = \sum_{d|q} f(n)1_{(n,q)=d} = \sum_{d|q} 1_{d|n} f_d(n/d),$$

donde $f_d : \mathbb{Z} \to \mathbb{C}$ es la función tal que $f_d(m) = f(md)$ si $(m, q/d) = 1$ y $f_d(m) = 0$ de otra manera. Concluya que para ciertos $|a_\chi| \le 1$

$$f(n) = \sum_{d|q}\sum_{\chi\in\widehat{G}_{q/d}} a_\chi 1_{d|n}\chi(n/d).$$

**Ejercicio 4.7.** Demuestre el Lema 4.11. Para hacerlo: expanda el cuadrado, meta adentro la suma en $h$, aplique Cauchy-Schwarz, expanda los cuadrados y reagrúpelos de forma que queden cuadrados de sumas en $n$. Aplique dicho lema para demostrar el Teorema 1.2 a partir del Corolario 4.10.

**Ejercicio 4.8.** En este problema vamos a concluir la estimación de

$$S_N = \sum_{p+q+b=N}\log p\log q$$

que comenzamos en el ejercicio 4.4 de §4.1. Allí vimos que para cualquier $0 < \delta < 1/2$

$$S_N = O(\delta N^2\log N) + \int_{\mathfrak{M}} F(\alpha)G^2(\alpha)e(-N\alpha)\,d\alpha$$

con $\mathfrak{M} = (-\frac{1}{N\delta}, \frac{1}{N\delta})$ los arcos mayores,

$$F(\alpha) = \sum_{b=1}^{N} e(b\alpha), \qquad\qquad G(\alpha) = \sum_{p\le N}\log p\, e(p\alpha).$$

1. Use la regla del rectángulo (2.21) para obtener la estimación

$$F(\alpha) = \frac{e(N\alpha) - 1}{2\pi i\alpha} + O(\delta^{-1}) \tag{4.11}$$

para $\alpha \in \mathfrak{M}$.

2. Escriba $G(\alpha) = \sum_{n \le N} (1_{n \text{ primo}} \log n) e(n\alpha)$. Use sumación por partes (ejercicio 2.2 de §2.1), el teorema 2.5 y la estimación (4.11) para obtener que

$$G(\alpha) = \frac{e(N\alpha) - 1}{2\pi i\alpha} + O_A\left(\frac{\delta^{-1}N}{(\log N)^A}\right)$$

para $\alpha \in \mathfrak{M}$ y $A > 0$ arbitrario.

3. Por los apartados anteriores y el cambio de variable $\alpha = t/N$, demuestre que

$$S_N = (C + O(\delta^2))N^2 + O(\delta N^2 \log N) + O_A\left(\frac{N^2}{\delta^2 (\log N)^A}\right)$$

donde $C = \int_{-\infty}^{\infty} \frac{(e(t)-1)^3 e(-t)}{(2\pi i t)^3}\, dt$. Tomando $\delta = (\log N)^{-2}$, $A = 5$, obtenemos que

$$S_N = (C + o(1))N^2.$$

4. Para demostrar que $C = 1/2$ sin dolor, verifique que el mismo razonamiento muestra que $T_N = \sum_{a+b+c=N} 1$ satisface $T_N = (C+o(1))N^2$, y luego muestre que $T_N = (1/2 + o(1))N^2$ de otra manera. Alternativamente, muestre que $C = 1/2$ como prefiera.

## 5. La autocorrelación de $\lambda$ en escala logarítmica

**5.1. Inicio y esbozo del argumento.** Quisiéramos ahora probar el Teorema 1.3. Para simplificar la notación, nos concentraremos en el caso $a_1 = a_2 = b_2 = 1$, $b_1 = 0$; es decir, probaremos que, para $w = w(x)$ tal que $w \to \infty$ cuando $x \to \infty$,

$$\sum_{\frac{x}{w} < n \le x} \frac{\lambda(n)\lambda(n+1)}{n} = o(\log w). \tag{5.1}$$

El tratamiento del caso general ($a_i$, $b_i$ arbitrarios) es prácticamente idéntico.

De hecho, probaremos (5.1) en la siguiente forma cuantitativa.

**Teorema 5.1.** *Sean $w > e^e$, $x > e^{e^e}$. Entonces*

$$\sum_{x/w < n \le x} \frac{\lambda(n)\lambda(n+1)}{n} \ll \frac{\log w}{\text{mín}(\log_3 w, \log_4 x)^{1/5}}.$$

Cuando escribimos $\log_k$, queremos decir el logaritmo iterado $k$ veces: $\log_2 x = \log\log x$, $\log_3 x = \log\log\log x$, $\log_4 x = \log\log\log\log x$.

El primer paso hacia el Teorema 5.1 consiste en usar la multiplicatividad de $\lambda$ para escribir la suma como una suma de sumas con una condición de divisibilidad.

**Lema 5.2.** *Sea $1 \le w \le x$. Sean $1 \le K_0 \le K_1 < x/w$. Entonces*

$$\sum_{\frac{x}{w} < n \le x} \frac{\lambda(n)\lambda(n+1)}{n} = \frac{1}{\ell}\left(\sum_{K_0 < p \le K_1} \sum_{\substack{\frac{x}{w} < n \le x \\ p \mid n}} \frac{\lambda(n)\lambda(n+p)}{n} + O(\log K_1)\right), \tag{5.2}$$

*donde $\ell = \sum_{K_0 < p \le K_1} p^{-1}$.*

La idea principal de la prueba es que el intervalo $x/w < n \le x$ con el peso $1/n$ es casi invariante bajo desplazamientos multiplicativos $p\cdot$, $p$ pequeño.

*Demostración.* Está claro que

$$\sum_{K_0<p\le K_1} \frac{1}{p} \sum_{\frac{x}{w}<n\le x} \frac{\lambda(n)\lambda(n+1)}{n} = \sum_{K_0<p\le K_1} \sum_{\frac{x}{w}<n\le x} \frac{\lambda(pn)\lambda(pn+p)}{pn}$$

Ahora bien, para $p \le K_1$,

$$\begin{aligned}\sum_{\frac{x}{w}<n\le x} \frac{\lambda(pn)\lambda(pn+p)}{pn} &= \sum_{\frac{x}{pw}<n\le \frac{x}{p}} \frac{\lambda(pn)\lambda(pn+p)}{pn} + O\left(\frac{1}{p}\sum_{\frac{x}{pw}<n\le\frac{x}{w}} \frac{1}{n} + \frac{1}{p}\sum_{\frac{x}{p}<n\le x} \frac{1}{n}\right)\\ &= \sum_{\substack{\frac{x}{w}<n\le x\\ p|n}} \frac{\lambda(n)\lambda(n+p)}{n} + \frac{O(\log p)}{p}.\end{aligned}$$

Aplicando el Corolario 2.6 y dividiendo por $\sum_{K_0<p\le K_1} p^{-1}$, obtenemos el resultado. □

Tras este resultado, para demostrar (5.1), será suficiente mostrar que para algún par $(K_0, K_1)$ con $\log K_1 = o(\log w \sum_{K_0<p\le K_1} \frac{1}{p})$ se cumple

$$\sum_{K_0<p\le K_1} \sum_{\substack{\frac{x}{w}<n\le x\\ p|n}} \frac{\lambda(n)\lambda(n+p)}{n} = o\left((\log w)\cdot \sum_{K_0<p\le K_1} \frac{1}{p}\right). \tag{5.3}$$

La idea principal es la siguiente. Los métodos de Matomäki y Radziwiłł (en particular, los que acabamos de ver en §4) nos bastarán para probar que

$$\sum_{K_0<p\le K_1} \frac{1}{p} \sum_{\frac{x}{w}<n\le x} \frac{\lambda(n)\lambda(n+p)}{n} = o\left((\log w)\cdot \sum_{K_0<p\le K_1} \frac{1}{p}\right). \tag{5.4}$$

Ahora bien, si vemos al hecho de ser divisible por $p$ como un evento aleatorio de probabilidad $1/p$, tiene sentido que los lados izquierdos de (5.3) y (5.4) sean aproximadamente iguales.

En verdad, que $n$ sea divisible por $p$ *es* un evento aleatorio, si tomamos $n$ al azar entre $x/w$ y $x$. El problema reside en que se trata de un evento no independiente de $\lambda(n)\lambda(n+p)$.

Ahora bien, resulta ser que – para hablar de manera aproximada – si la dependencia entre los eventos $p|n$ $(K_0 < p \le K_1)$ y $(\lambda(n), \lambda(n+1), \dots)$ $(x/w < n \le x)$ es fuerte para muchos valores de $(K_0, K_1)$, entonces existe un valor de $(K_0, K_1)$ en la cual no lo es tanto. Existe una medida de dependencia – la *información mutua*, definida en términos de la *entropía* – la cual, si bien es un tanto burda, goza de una propiedad de aditividad. Esto conlleva que se pueda tratar a la entropía como un recurso agotable; podemos pensar en ella como una sopa con una cantidad finita de lentejas, de tal manera que, si la gente se va sirviendo, eventualmente a alguien le

tendrá que tocar pocas lentejas (es decir, poca información mutua).[5] En el momento que nos toca pocas lentejas, la dependencia es débil, y sabremos proceder.

*Ejercicios.*

**Ejercicio 5.1.** Queremos ver que el resultado (5.1) no es tan fuerte como la conjetura de Chowla. Sea $(a_n)_{n\in\mathbb{N}}$ una sucesión con $|a_n| \leq 1$.

1. Demuestre que $\sum_{n\leq x} a_n = o(x)$ implica $\sum_{x/w\leq n\leq x} \frac{a_n}{n} = o(\log w)$ cuando $x \to \infty$, con $x \geq w = w(x) \to \infty$. *Sugerencia: use sumación por partes.*
2. Observe, usando la regla del rectángulo (2.21), que la sucesión $a_n = n^i$ satisface $\sum_{x/w\leq n\leq x} a_n/n = O(1) = o(\log w)$ cuando $x \to \infty$ para cualquier $x \geq w = w(x) \to \infty$, pero que $\sum_{n\leq x} a_n \neq o(x)$. Muestre que lo mismo ocurre para $a_n = \Re n^i = \cos(\log n)$.
3. Demuestre que si tuviéramos cancelación para $w = 2$, es decir $\sum_{x/2<n\leq x} \frac{a_n}{n} = o(1)$, entonces sí podríamos deducir que $\sum_{n\leq x} a_n = o(x)$. *Sugerencia: use sumación por partes.*

**Ejercicio 5.2.** El Teorema 5.1 vale en el rango $2 \leq w \leq x^{1/8}$. Usando ese resultado, deduzca que el teorema también es cierto en el rango $x^{1/8} < w \leq x$.

**5.2. Sumas y esperanzas.** En esta sección vamos a formalizar la relación entre las sumas en (5.3) y los conceptos de probabilidad e independencia que hemos comentado. Para ello, es conveniente primero partir la suma en $n$ en segmentos cortos (de longitud $H$). Esto es lo que hacemos en el siguiente resultado.

**Lema 5.3.** *Sea $H$ un número natural tal que $K_1 < H \leq x/w$. Entonces para cualquier $p \leq K_1$ tenemos*

$$\sum_{\substack{\frac{x}{w}<n\leq x\\ p|n}} \frac{\lambda(n)\lambda(n+p)}{n}$$

$$= \frac{1}{H} \sum_{\frac{x}{w}<n\leq x} \frac{1}{n} \sum_{j\leq H-p} \lambda(n+j)\lambda(n+j+p)1_{p|n+j} + O\left(\frac{\log w}{H} + \frac{1}{p}\right).$$

La idea es simplemente utilizar el hecho de que el peso $1/n$ y el intervalo $(x/w, x]$ son aproximadamente invariantes bajo pequeños desplazamientos aditivos.

*Demostración.* Para cualquier $j \leq H$, por cambio de variable

$$\sum_{\substack{\frac{x}{w}<n\leq x\\ p|n}} \frac{\lambda(n)\lambda(n+p)}{n} = \sum_{\substack{\frac{x}{w}-j<n\leq x-j\\ p|n+j}} \frac{\lambda(n+j)\lambda(n+j+p)}{n+j}.$$

Como

$$\sum_{\substack{\frac{x}{w}-j<n\leq \frac{x}{w}\\ p|n+j}} \frac{1}{n+j} + \sum_{\substack{x-j<n\leq x\\ p|n+j}} \frac{1}{n+j} = \sum_{\substack{\frac{x}{w}<n\leq \frac{x}{w}+j\\ p|n}} \frac{1}{n} + \sum_{\substack{x<n\leq x+j\\ p|n}} \frac{1}{n}$$

$$\ll \frac{1}{p}\left(1 + \log\frac{x/w+H}{x/w}\right) \ll \frac{1}{p}$$

[5]En la versión oral de estas charlas, se mencionó a un ollón de locro, pero se hizo aparente que parte de la audiencia no sabía qué era el locro.

y

$$\sum_{\substack{\frac{x}{w}<n\le x\\ p|n+j}}\left(\frac{1}{n}-\frac{1}{n+j}\right)\le\sum_{\substack{n>x/w\\ p|n+j}}\frac{j}{n(n+j)}\le 2H\sum_{\substack{n>x/w\\ p|n+j}}\frac{1}{(n+j)^2}$$
$$\le 2H\sum_{\substack{n>x/w\\ p|n}}\frac{1}{n^2}\ll\frac{H}{p^2}\cdot\frac{1}{x/wp}\le\frac{1}{p},$$

vemos que

$$(H-p)\sum_{\substack{\frac{x}{w}<n\le x\\ p|n}}\frac{\lambda(n)\lambda(n+p)}{n}$$
$$=\sum_{j\le H-p}\sum_{\frac{x}{w}<n\le x}\frac{\lambda(n+j)\lambda(n+j+p)1_{p|n+j}}{n}+O\left(\frac{1}{p}\right)\cdot(H-p).$$

Dividamos todo por $H$. Para concluír, notemos que

$$\frac{p}{H}\left|\sum_{\substack{\frac{x}{w}<n\le x\\ p|n}}\frac{\lambda(n)\lambda(n+p)}{n}\right|\le\frac{p}{H}\sum_{\substack{\frac{x}{w}<n\le x\\ p|n}}\frac{1}{n}\ll\frac{p}{H}\cdot\frac{\log w+1}{p}\le\frac{\log w}{H}+\frac{1}{p}.$$

□

Tomemos $K_0=\epsilon H/2$, $K_1=\epsilon H$, con $0<\epsilon<1$ pequeño. Nuestro objetivo – el cual nos permitirá demostrar (5.3), gracias al Lema 5.3 – será acotar de forma no trivial la suma

$$S=\frac{1}{H}\sum_{\frac{x}{w}<n\le x}\frac{1}{n}\sum_{K_0<p\le K_1}\sum_{j\le H-p}\lambda(n+j)\lambda(n+j+p)1_{p|n+j}. \tag{5.5}$$

Para ver la conexión con las probabilidades, observemos que si $N$ es la variable aleatoria que toma valores en el conjunto de enteros en el intervalo $(x/w,x]$ con probabilidad

$$\mathbb{P}(N=n)=\frac{1/n}{L}\quad\text{si}\quad n\in(x/w,x], \tag{5.6}$$

donde $L=\sum_{x/w<n\le x}\frac{1}{n}$, entonces podemos escribir

$$S=\frac{L}{H}\cdot\mathbb{E}\left(\sum_{K_0<p\le K_1}\sum_{j\le H-p}\lambda(N+j)\lambda(N+j+p)1_{N\equiv -j(p)}\right),$$

es decir, $S$ es la esperanza de una variable aleatoria que es una suma doble de variables aleatorias. Para examinar la dependencia entre la condición de divisibilidad y los términos con $\lambda$, definimos la variables aleatorias

$$X_H=(\lambda(N+1),\lambda(N+2),\dots,\lambda(N+H)),\qquad Y_H=(N \bmod p)_{K_0<p\le K_1}, \tag{5.7}$$

con $X_H$ tomando valores en $\{-1,1\}^H$ e $Y_H$ en $\Omega = \prod_{K_0<p\leq K_1} \frac{\mathbb{Z}}{p\mathbb{Z}}$. Así podemos escribir

$$S = \frac{L}{H} \cdot \mathbb{E}(F(X_H, Y_H)) \tag{5.8}$$

con $F : \{-1,1\}^H \times \Omega \to \mathbb{R}$ la función

$$F(\vec{x}, \vec{y}) = \sum_{K_0<p\leq K_1} \sum_{j\leq H-p} x_j x_{j+p} 1_{y_p \equiv -j(p)}. \tag{5.9}$$

Obtenemos el siguiente resultado. El interés en estimar la expresión en el lado izquierdo de (5.10) viene, claro está, del hecho que aparece en el lado derecho de (5.2).

**Lema 5.4.** *Sea $1 \leq w < x$. Sean $K_0 = \epsilon H/2$ y $K_1 = \epsilon H$, con $H \leq x/w$ y* máx $\left(\frac{1}{\log w}, \frac{1}{\sqrt{H}}\right) \leq \epsilon < 1$. *Entonces*

$$\sum_{K_0<p\leq K_1} \sum_{\substack{\frac{x}{w}<n\leq x \\ p|n}} \frac{\lambda(n)\lambda(n+p)}{n} = \frac{L}{H}\mathbb{E}(F(X_H, Y_H)) + O\left(\epsilon \frac{\log w}{\log H}\right), \tag{5.10}$$

*donde $L = \sum_{x/w<n\leq x} n^{-1}$ y $F$ es como en* (5.9).

*Demostración.* Por el Lema 5.3 y las estimaciones del Corolario 2.6,

(5.11)

$$\begin{aligned}\sum_{K_0<p\leq K_1} \sum_{\substack{\frac{x}{w}<n\leq x \\ p|n}} \frac{\lambda(n)\lambda(n+p)}{n} &= S + \sum_{K_0<p\leq K_1} O\left(\frac{1}{p} + \frac{\log w}{H}\right) \\ &= S + \left(\log\log K_1 - \log\log\frac{K_1}{2} + O\left(\frac{1}{\log K_1}\right)\right) + O\left(\frac{\log w}{H} \cdot \frac{K_1}{\log K_1}\right) \\ &= S + O\left(\frac{\epsilon \log w + 1}{\log K_1}\right),\end{aligned}$$

donde $S$ es como en (5.5). Gracias a $\epsilon \geq$ máx$(1/\log w, 1/\sqrt{H})$, vemos que $\epsilon \log w + 1 \leq 2\epsilon \log w$ y $\log K_1 \leq (\log H)/2$. Usamos la expresión (5.8) para $S$. □

El plan es mostrar que, para algún $H$, las variables $X_H$ e $Y_H$ son más o menos independientes, y usar este hecho para obtener $\mathbb{E}(F(X_H, Y_H)) = o(H/\log H)$, lo cual es exactamente lo necesario para poder concluír, por el Lema 5.2, que (5.1) se cumple. Para ello aprovecharemos que la función $F$ puede escribirse como una suma de variables aleatorias:

$$F(\vec{x}, \vec{y}) = \sum_{\frac{\epsilon H}{2}<p\leq \epsilon H} F_p(\vec{x}, y_p) \tag{5.12}$$

con $\vec{y} = (y_p)_p$ y $F_p(\vec{x}, \cdot) : \mathbb{Z}/p\mathbb{Z} \to \mathbb{R}$ definida por $F_p(\vec{x}, t) = \sum_{j\leq H-p,\ j\equiv -t(p)} x_j x_{j+p}$. Como

$$\|F_p\|_\infty \leq \frac{H}{p} \leq \frac{2}{\epsilon}, \tag{5.13}$$

por el teorema de los números primos vemos que

$$\|F\|_\infty \ll \frac{H}{\log H} \tag{5.14}$$

por lo que sólo necesitamos mejorar un poco esa cota para obtener nuestro objetivo para $\mathbb{E}(F(X_H, Y_H))$.

Ahora bien, tenemos el siguiente resultado.

**Lema 5.5.** *Sean $C, n \geq 1$ y $\Omega = \Omega_1 \times \ldots \times \Omega_n$ con $\Omega_m$ conjuntos finitos, y $G_m$ una función real sobre $\Omega_m$ para $1 \leq m \leq n$, con $\|G_m\|_\infty \leq C$. Definamos $G : \Omega \to \mathbb{R}$ por*

$$G(\vec{t}) = G(t_1, t_2, \ldots t_n) = G_1(t_1) + G_2(t_2) + \ldots + G_n(t_n),$$

*y sea $\overline{G}$ su promedio sobre $\Omega$. Entonces, para cualquier $0 < \mu < 1$, se cumple que*

$$|G(\vec{t}) - \overline{G}| \leq \mu \cdot Cn$$

*para todo $\vec{t} \in \Omega$ excepto para un conjunto de tamaño a lo más $2|\Omega|^{1-\frac{\mu^2/2}{\log R}}$, con $R = |\Omega|^{1/n}$.*

*Demostración.* Si consideramos una variable aleatoria $T = (T_1, \ldots, T_n)$, $T_m$ con distribución uniforme en $\Omega_m$ e independientes, el enunciado del lema equivale a decir que

$$\mathbb{P}(|G(T) - \mathbb{E}(G(T))| \geq \mu \cdot Cn) \leq 2e^{-\frac{\mu^2 n}{2}}.$$

Pero esto se deduce directamente de la desigualdad de Hoeffding, que nos dice que si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes tomando valores en el intervalo $[-C, C]$ entonces para $S = X_1 + \ldots + X_n$ se cumple que

$$\mathbb{P}(|S - \mathbb{E}(S)| \geq s) \leq 2e^{-\frac{s^2}{2C^2 n}}. \tag{5.15}$$

Un breve comentario: si bien (5.15) es de forma claramente similar al teorema central del límite, se trata de un resultado de *grandes desviaciones*, pues es válido para $s$ arbitrario, mientras que el teorema central del límite se ocupa del caso en el cual $s$ está acotado por un múltiplo constante de la desviación estándar, es decir, el caso $s \ll C\sqrt{n}$. □

Así, teniendo en cuenta (5.12) y (5.13), podemos aplicar el Lema 5.5 con $F(\vec{x}, \cdot)$ para obtener (usando el teorema de los números primos) que, para $H$ más grande que una constante y $\epsilon \geq 2/\sqrt{H}$,

$$|F(\vec{x}, \vec{y}) - \overline{F(\vec{x}, \cdot)}| \leq \mu \frac{4H}{\log H} \tag{5.16}$$

para todo $\vec{y} \in \Omega$ excepto en un conjunto $E_{\vec{x}} \subset \Omega$ que satisface

$$|E_{\vec{x}}| \leq 2|\Omega|^{1-\frac{\mu^2}{2\log H}}. \tag{5.17}$$

Esta cota, junto con (5.13), nos ayudará a estimar $\mathbb{E}(F(X_H, Y_H))$, ya que sólo tendremos que preocuparnos en mostrar que, la mayor parte del tiempo, $Y_H$ tiende a evitar un conjunto relativamente pequeño de valores $E_{X_H}$, dado por $X_H$. Claro está, como $Y_H$ está casi equidistribuída, tal aseveración se deduciría de inmediato si $X_H$ y $Y_H$ fueran variables independientes. Como veremos, bastará probar una forma muy débil de independencia, para algún $H$.

*Ejercicios.* Los siguientes ejercicios dan un ejemplo muy básico de como deducir un enunciado sobre los enteros de un enunciado probabilístico general sobre sumas de variables aleatorias.

**Ejercicio 5.3.** Pruebe la *desigualdad de Chebyshev* para una variable aleatoria $X$:

$$\mathbb{P}(|X - \mu| \geq \lambda\sigma^2) \leq \frac{1}{\lambda^2}$$

para todo $\lambda \in \mathbb{R}$, con $\mu = \mathbb{E}[X]$ la esperanza de $X$ y $\sigma^2 = \mathbb{E}[(X-\mu)^2]$ su varianza.

**Ejercicio 5.4.** En este ejercicio queremos demostrar que en $[1, x]$ casi todo entero (es decir, todos, salvo $o(x)$ de ellos) tiene $(1 + o(1)) \log\log x$ divisores primos distintos.

Podemos asumir que $x$ es entero. Sea $N$ una variable aleatoria en el espacio $\Omega = \{n \leq x\}$ tal que $\mathbb{P}(N = n) = 1/x$ para todo $n \in \Omega$. Para todo primo $p \leq D = x^{1/4}$, considere la variable aleatoria $X_p = 1_{N \equiv 0 (p)}$.

1. Demuestre que $X_p$ es una variable de Bernoulli con media $\mu_p = \frac{1}{p} + O(x^{-1})$ y por lo tanto con varianza $\sigma_p^2 = \mu_p(1-\mu_p) = \frac{1}{p}(1-\frac{1}{p}) + O(x^{-1})$.
2. Pruebe que para $p \neq q$ las variables $X_p$ y $X_q$ tienen covarianza casi nula: para $Y_p = X_p - \mu_p$ tenemos $\mathbb{E}[Y_p Y_q] \ll x^{-1}$.
3. Muestre que la variable $w = \sum_{p \leq D} X_p$ tiene media $\mu = \sum_{p \leq D} \mu_p$ y que su varianza $\sigma^2 = \mathbb{E}[(\sum_{p<D} Y_p)^2]$ satisface $\sigma^2 = \sum_{p \leq D} \sigma_p^2 + O(x^{-1/2})$. (En otras palabras, las variables $X_p$ se comportan como si fueran independientes.)
4. Usando el teorema de los números primos (o, para ser precisos, (2.18)), concluya que $\sigma^2 = \mu + O(1) = \log\log x + O(1)$.
5. Use la desigualdad de Chebyshev para demostrar el enunciado del ejercicio, observando que un entero en $[1, x]$ tiene a lo sumo 3 divisores primos mayores que $D$.

**5.3. Entropía e información mutua.** Ahora comenzamos la labor de mostrar que, para algún $H$, las variables aleatorias $X_H$ e $Y_H$ no son muy dependientes. Esto lo mediremos mediante el concepto de «información mutua», que pasamos a definir.

*5.3.1. Definiciones.* Sea $X$ una variable aleatoria con un número finito de valores posibles $x$. La *entropía* $\mathbb{H}(X)$ de $X$ es

$$\mathbb{H}(X) = -\sum_x p_x \log p_x,$$

donde $p_x$ es la probabilidad $\mathbb{P}(X = x)$ de que $X$ tome el valor $x$. La *entropía condicional* de $X$ con respecto a una variable aleatoria $Y$ (que suponemos tener también un número finito de valores, o por lo menos ser discreta) es

$$\mathbb{H}(X|Y) = \sum_y \mathbb{H}(X|Y=y)\mathbb{P}(Y=y), \tag{5.18}$$

donde, para $E$ un evento probabilístico (como $Y = y$), $\mathbb{H}(X|E)$ se define por $\mathbb{H}(X|E) = -\sum_x p_{x,E} \log p_{x,E}$, donde $p_{x,E} = \mathbb{P}(X = x|E)$.

Las siguientes propiedades básicas son fáciles de probar (ver los ejercicios). Escribimos $\mathbb{H}(X, Y)$ para denotar la entropía $\mathbb{H}((X, Y))$ de la variable aleatoria $(X, Y)$, donde $X$ e $Y$ son variables aleatorias.

1. La entropía $\mathbb{H}(X)$ y la entropía condicional $\mathbb{H}(X|Y)$ son no negativas.
2. $\mathbb{H}(X, Y) = \mathbb{H}(X|Y) + \mathbb{H}(Y) = \mathbb{H}(Y|X) + \mathbb{H}(X)$,
3. $\mathbb{H}(X|Y) \leq \mathbb{H}(X)$,

4. $\mathbb{H}(X,Y) \le \mathbb{H}(X) + \mathbb{H}(Y)$ (subaditividad de la entropía).
5. Si $X$ toma $\le N$ valores distintos, $\mathbb{H}(X) \le \log N$.

También podemos acotar con facilidad la diferencia entre las entropías de dos variables aleatorias $X$, $Y$ que toman los mismos valores con probabilidades distintas (ejercicio 5.6).

Definimos la *información mutua* $\mathbb{I}(X,Y)$:

$$\mathbb{I}(X,Y) = \mathbb{H}(X) + \mathbb{H}(Y) - \mathbb{H}(X,Y). \tag{5.19}$$

Por la subaditividad de la entropía, $\mathbb{I}(X,Y) \ge 0$. Está claro por 2. que

$$\mathbb{H}(X|Y) = \mathbb{H}(X) - \mathbb{I}(X,Y), \qquad \mathbb{H}(Y|X) = \mathbb{H}(Y) - \mathbb{I}(X,Y). \tag{5.20}$$

*5.3.2. El argumento por agotamiento de información mutua.* Consideramos ahora las variables aleatorias $X_H, Y_H$ definidas en (5.7) en términos de la variable aleatoria $N$ cuya distribución fue dada en (5.6). La meta es mostrar que existe un $H \in [H_-, H_+]$ (donde $H_-$, $H_+$ serán especificados más tarde) tal que $\mathbb{I}(X_H, Y_H)$ es pequeña en comparación con $H$.

Podemos definir para $H_1$, $H_2$ arbitrarios,

$$X_{H_1,H_1+H_2} = (\lambda(N+j))_{H_1<j\le H_1+H_2}. \tag{5.21}$$

Podemos asumir sin pérdida de generalidad que $w \le x^{1/8}$. También asumiremos que $H_1, H_2 \le x^{1/8}$. Entonces, por el ejercicio 3,

$$\mathbb{H}(X_{H_1,H_1+H_2}) = \mathbb{H}(X_{H_2}) + O\left(1/x^{5/8}\right).$$

Por la subaditividad de la entropía, deducimos que

$$\mathbb{H}(X_{H_1+H_2}) \le \mathbb{H}(X_{H_1}) + \mathbb{H}(X_{H_1,H_1+H_2}) \le \mathbb{H}(X_{H_1}) + \mathbb{H}(X_{H_2}) + O(1/x^{5/8}). \tag{5.22}$$

Vemos por el ejercicio 4 (suponiendo que $\delta \asymp (\prod_{K_0<p\le K_1} p)^{-1}$ satisface $\delta^{-1} \le x^{1/8}$) que el mismo razonamiento vale para la entropía condicional:

$$\mathbb{H}(X_{H_1,H_1+H_2}|(N+H_1 \bmod p)_{K_0<p\le K_1}) = \mathbb{H}(X_{H_2}|(N \bmod p)_{K_0<p\le K_1}) + O(1/\sqrt{x}).$$

Por el teorema de los números primos (en la forma [23, Thm. 9]),

$$\prod_{K_0<p\le K_1} p \le e^{\sum_{p\le K_1} \log p} \le e^{1{,}02K_1}. \tag{5.23}$$

Supondremos entonces que $K_1 \le (\log x)/9$.

Ahora bien, $N + H_1 \bmod p$ codifica la misma información que $N \bmod p$. Por lo tanto,

$$\mathbb{H}(X_{H_1,H_1+H_2}|(N \bmod p)_{K_0<p\le K_1}) = \mathbb{H}(X_{H_1,H_1+H_2}|(N+H_1 \bmod p)_{K_0<p\le K_1})$$

Recordamos la definición (5.7) de $Y_H$. De nuevo por subaditividad, como en (5.22), concluímos que

$$\mathbb{H}(X_{H_1+H_2}|Y_H) \le \mathbb{H}(X_{H_1}|Y_H) + \mathbb{H}(X_{H_2}|Y_H) + O(1/\sqrt{x}).$$

Iterando con $H_1 = H, 2H, 3H, \dots$ y $H_2 = H$, vemos que

$$\mathbb{H}(X_{kH}|Y_H) \le k\mathbb{H}(X_H|Y_H) + O(k/\sqrt{x})$$

para $kH \le x^{1/8}$, y en consecuencia, otra vez por subaditividad,

$$\mathbb{H}(X_{kH}) \le \mathbb{H}(X_{kH}|Y_H) + \mathbb{H}(Y_H) \le k\mathbb{H}(X_H|Y_H) + \mathbb{H}(Y_H) + O(k/\sqrt{x}),$$

así que, por (5.20),

$$\frac{\mathbb{H}(X_{kH})}{kH} \le \frac{\mathbb{H}(X_H)}{H} - \frac{\mathbb{I}(X_H, Y_H)}{H} + \frac{\mathbb{H}(Y_H)}{kH} + O\left(\frac{1}{\sqrt{x}}\right).$$

Por la propiedad 5 de la entropía y (5.23),

$$\mathbb{H}(Y_H) \le \log \left( \prod_{K_0 < p \le K_1} p \right) \le 1{,}02\epsilon H.$$

Concluímos que,

$$\frac{\mathbb{H}(X_{kH})}{kH} \le \frac{\mathbb{H}(X_H)}{H} - \frac{\mathbb{I}(X_H, Y_H)}{H} + 1{,}02\frac{\epsilon}{k} + O\left(\frac{1}{\sqrt{x}}\right). \tag{5.24}$$

He aquí el sujeto de nuestra metáfora: la tasa de entropía $\mathbb{H}(X_{H'})/H'$ (con $H' = H, kH, \dots$) es "las lentejas", y la tasa $\mathbb{I}(X_H, Y_H)/H$ es la cucharada que nos servimos. Para cuantificar que tan poca lenteja nos terminaremos sirviendo en algún momento, basta con un sencillo lema.

**Lema 5.6.** *Sea $h_1 \ge 15$ arbitrario y, para $j \ge 1$, $h_{j+1} = \lfloor 4 \log h_j \log_3 h_j \rfloor \cdot h_j$. Entonces*

$$\sum_{j=1}^{J} \frac{1}{\log h_j \log_3 h_j} \ge 100$$

*para algún $J$ cumpliendo $\log J \ll (\log_2 h_1)^2$.*

*Demostración.* Ejercicio 5.8. □

**Corolario 5.7.** *Sean $X_H$, $Y_H$ y $N$ como en (5.6) y (5.7), con $K_0 = \epsilon H/2$, $K_1 = \epsilon H$, $0 < \epsilon \le 1$ y $w \le x^{1/8}$. Sea $H_- \ge 3$. Entonces hay un $H_+ > H_-$, dependiendo sólo de $H_-$ y cumpliendo $\log_3 H_+ \le 2 \log_3 H_- + O(1)$, tal que*

$$\mathbb{I}(X_H, Y_H) \le \frac{H}{\log H \log_3 H} \tag{5.25}$$

*para algún entero $H \in [H_-, H_+]$, con tal que $x \ge \exp(H_+^9)$.*

Es fácil ver que la condición $x \ge \exp(H_+^9)$ se deduce de $\log_3 H_+ \le 2 \log_3 H_- + O(1)$ para $x$ más grande que una constante y $H_- \le \exp(\exp(\sqrt{\log_3 x}/C))$, donde $C > 0$ es otra constante.

*Demostración.* Podemos asumir sin pérdida de generalidad que $H_-$ es un entero más grande que una constante apropiada. Sea $h_1 = H_-$; definamos $h_2, h_3, \dots$ como en el Lema 5.6, y sea $k_j = \lfloor 4 \log h_j \log_3 h_j \rfloor$. Utilizando la suposición que $h_j \ge H_-$ es más grande que una constante, así como la desigualdad $k_j \le k_j h_j \le x^{1/8}$, la cual debemos asumir de todas maneras, simplificamos (5.24), obteniendo

$$\frac{\mathbb{H}(X_{h_{j+1}})}{h_{j+1}} \le \frac{\mathbb{H}(X_{h_j})}{h_j} - \frac{\mathbb{I}(X_{h_j}, Y_{h_j})}{h_j} + \frac{1}{2 \log h_j \log_3 h_j}. \tag{5.26}$$

Deducimos entonces de las propiedades (1) y (5) de la entropía que

$$\sum_{j \le J} \left( \frac{\mathbb{I}(X_{h_j}, Y_{h_j})}{h_j} - \frac{1}{2 \log h_j \log_3 h_j} \right) \le \frac{\mathbb{H}(X_{H_-})}{H_-} \le \log 2.$$

Por el Lema 5.6, vemos que hay algún $J$ dependiendo sólo de $H_- = h_1$, tal que

$$\frac{\mathbb{I}(X_{h_j}, Y_{h_j})}{h_j} - \frac{1}{2\log h_j \log_3 h_j} \leq \frac{1}{2\log h_j \log_3 h_j}$$

para algún $1 \leq j \leq J$. (¿Por qué?) Concluímos que

$$\frac{\mathbb{I}(X_{h_j}, Y_{h_j})}{h_j} \leq \frac{1}{\log h_j \log_3 h_j}.$$

Definimos $H_+ = h_j$ y obtenemos el resultado. La condición $H_+ \leq (\log x)/9$ implica $H_+ \leq x^{1/8}$, y así también la suposición $k_j h_j = h_{j+1} \leq x^{1/8}$ para todo $j \leq J-1$. □

*Ejercicios.*

**Ejercicio 5.5.** Sean $X$ e $Y$ variables aleatorias que toman un número finito de valores. Pruebe las propiedades (1)–(5) de la entropía. Sugerencias para cada propiedad:

(1) Use simplemente las definiciones de entropía y entropía condicional.
(2) De nuevo por las definiciones.
(3) Por la concavidad de $x \mapsto -x \log x$ (primera desigualdad) y por el hecho que $\log x$ es creciente (segunda desigualdad; también se deduce inmediatamente de las propiedades (1) y (2)).
(4) Use las propiedades (2) y (3).
(5) Por la concavidad de $x \mapsto \log x$.

**Ejercicio 5.6.** Sean $X$ e $Y$ variables aleatorias con valores en un conjunto $S$ de $N$ elementos. Denotemos por $\nu_X$, $\nu_Y$ sus funciones de distribución. La *distancia de variación total* $|\nu_X - \nu_Y|_{TV}$ se define como $\max_{S' \subset S} |\nu_X(S') - \nu_Y(S')|$. (Es fácil ver que es igual a $\frac{1}{2}|\nu_X - \nu_Y|_1$.)

1. Muestre que $p = \sum_{x \in S} \min(\nu_X(x), \nu_Y(x))$ es igual a $1 - |\nu_X - \nu_Y|_{TV}$. Si $p = 1$, entonces $X$ e $Y$ tienen la misma distribución, y estamos en el caso trivial. Si $p = 0$, $X$ e $Y$ tienen soportes disjuntos, y la cota que probaremos al final es muy sencilla (muéstrelo llegado el momento). Asumamos de ahora en adelante que $0 < p < 1$.
2. Sea $Z$ una variable aleatoria con función de distribución

$$\nu_Z(x) = \frac{1}{p} \min(\nu_X(x), \nu_Y(x)).$$

Sean $X'$ e $Y'$ variables aleatorias independientes con distribuciones

$$\nu_{X'}(x) = \begin{cases} \frac{\nu_X(x) - \nu_Y(x)}{1-p} & \text{si } \nu_X(x) > \nu_Y(x), \\ 0 & \text{de otra manera,} \end{cases}$$

$$\nu_{Y'}(x) = \begin{cases} \frac{\nu_Y(x) - \nu_X(x)}{1-p} & \text{si } \nu_Y(x) > \nu_X(x), \\ 0 & \text{de otra manera.} \end{cases}$$

Sea $B$ una variable que toma el valor 0 con probabilidad $p$ y el valor 1 con probabilidad $1-p$. Construyamos la variable aleatoria $C$ de la siguiente forma: si $B = 0$, $C$ toma el valor $(Z, Z)$; si $B = 1$, $C$ toma el valor $(X', Y')$. Muestre que la primera coordenada $C_1 = \pi_1(C)$ de $C$ es una variable con distribución $\nu_X$, mientras que la segunda coordenada $C_2 = \pi_2(C)$ tiene distribución $\nu_Y$. Muestre también que $\mathbb{P}(C_1 \neq C_2) = |\nu_X - \nu_Y|_{TV}$. Se dice que la variable $C$ es un *acoplamiento óptimo* de $X$ e $Y$.

3. Por las propiedades (3) de la entropía,

$$\mathbb{H}(C_1|B) \le \mathbb{H}(C_1) \le \mathbb{H}(C_1|B) + \mathbb{H}(B).$$

Sabemos que

$$\mathbb{H}(C_1|B) = p\mathbb{H}(C_1|B=0) + (1-p)\mathbb{H}(C_1|B=1) = p\mathbb{H}(Z) + (1-p)\mathbb{H}(X').$$

De la misma manera,

$$\mathbb{H}(C_2|B) = p\mathbb{H}(Z) + (1-p)\mathbb{H}(Y').$$

Por lo tanto,

$$\begin{aligned}|\mathbb{H}(X) - \mathbb{H}(Y)| &= |\mathbb{H}(C_1) - \mathbb{H}(C_2)| \\ &\le \mathbb{H}(B) + |\mathbb{H}(C_1|B) - \mathbb{H}(C_2|B)| \\ &\le \mathbb{H}(B) + (1-p)\,|\mathbb{H}(X') - \mathbb{H}(Y')| \\ &\le \mathbb{H}(B) + (1-p)\,\mín(\mathbb{H}(X'), \mathbb{H}(Y')).\end{aligned}$$

4. Concluya, por la propiedad (5) de la entropía, que

$$|\mathbb{H}(X) - \mathbb{H}(Y)| \le |\nu_X - \nu_Y|_{TV} \cdot \log N + \mathbb{H}(B).$$

En verdad, podemos dar una cota ligeramente menor: la variable $X'$ tiene soporte en un subconjunto estricto de $S$ (¿por qué?) y lo mismo es cierto de $Y'$; concluya que

$$|\mathbb{H}(X) - \mathbb{H}(Y)| \le |\nu_X - \nu_Y|_{TV} \cdot \log(N-1) + \mathbb{H}(B). \tag{5.27}$$

Aquí, por supuesto,

$$\mathbb{H}(B) = -|\nu_X - \nu_Y|_{TV} \log |\nu_X - \nu_Y|_{TV} - (1 - |\nu_X - \nu_Y|_{TV}) \log(1 - |\nu_X - \nu_Y|_{TV}).$$

**Ejercicio 5.7.** Sea $I = (x_0, x_1]$ un intervalo en $\mathbb{R}^+$. Sea $N$ una variable aleatoria que toma el valor entero $n \in I$ con probabilidad $(1/n)/L$, donde $L = \sum_{m \in I} 1/m$. Para $h \in \mathbb{Z}^+$, sea $N + h$ la variable que toma el valor $N + h$ con probabilidad $(1/n)/L$. Denote $\nu_X$ la distribución de probabilidad de una variable $X$.

1. Muestre que $|\nu_{N+h} - \nu_N|_{TV} \le h/x_0 L$.
2. Sea $f$ una función sobre $\mathbb{Z}^+$. Deduzca que

$$\left|\nu_{f(N+h)} - \nu_{f(N)}\right|_{TV} \le h/x_0 L. \tag{5.28}$$

3. Sean $X_H$ y $X_{H_1, H_1+H_2}$ como en (5.7) y (5.21). Concluya, usando (5.27) y (5.28), que, para $L = \sum_{x/w < n \le x} 1/n \ge 1$ y $1 \le H_1 \le x_0/2$, $x_0 = x/w$,

$$\begin{aligned}|\mathbb{H}(X_{H_1, H_1+H_2}) - \mathbb{H}(X_{H_2})| &\le \frac{H_1}{x_0 L} \log\left(2^{2H_2} - 1\right) + h\left(\frac{H_1}{x_0 L}\right) \\ &\le \frac{H_1}{x_0 L} H_2 \log 4 + h\left(\frac{H_1}{x_0}\right) \le \frac{H_1}{x_0}\left(H_2 \log 4 + 2 \log x_0\right),\end{aligned}$$

donde $h(\epsilon) = -\epsilon \log \epsilon - (1-\epsilon)\log(1-\epsilon) \le -2\epsilon \log \epsilon$ para $0 < \epsilon \le 1/2$. En particular, para $1 \le H_1, H_2 \le x_0^\alpha$, $\alpha > 0$,

$$|\mathbb{H}(X_{H_1, H_1+H_2}) - \mathbb{H}(X_{H_2})| \ll_\alpha \frac{1}{x_0^{1-2\alpha}}.$$

4. Sea $S \subset (x_0, x_1]$ y $\delta = \mathbb{P}(N \in S) = (\sum_{n \in S} 1/n)/L$. Muestre que la distancia de variación total entre las distribuciones de probabilidad condicional de $N + h$ y $N$ con la condición $N \in S$ es $\leq h/\delta x_0 L$. Concluya que, para $1 \leq H_1 \leq \delta x_0/2$,

$$|\mathbb{H}(X_{H_1,H_1+H_2}|N + H_1 \in S) - \mathbb{H}(X_{H_2}|N \in S)| \leq \frac{H_1}{\delta x_0} \left(H_2 \log 4 + 2 \log x_0\right),$$

con $x_0 = x/w$ y $x_1 = x$. En particular, si $1 \leq H_1, H_2, \delta^{-1} \leq x_0^{\alpha}$, $0 < \alpha < 1$,

$$|\mathbb{H}(X_{H_1,H_1+H_2}|N + H_1 \in S) - \mathbb{H}(X_{H_2}|N \in S)| \ll_{\alpha} \frac{1}{x_0^{1-3\alpha}}.$$

**Ejercicio 5.8.** Sea $h_j$ como en el Lema 5.6.

1. Muestre que $\log h_j \leq 2(j + \log h_1) \log(j + \log h_1)$ para todo $j \geq 1$.
2. Pruebe que, para $j \geq \log h_1$,
$$\frac{1}{\log h_j \log_3 h_j} \geq \frac{1/8}{j \log j \log_3 j}.$$
3. Demuestre que
$$\sum_{\log h_1 < j < (\log h_1)^{C \log_2 h_1}} \frac{1}{j \log j \log_3 j} > 800$$
para $C$ suficientemente grande, independiente de $h_1$. (Utilice, por ejemplo, un test de integrales.) Concluya que el Lema 5.6 es cierto.

**5.4. Información mutua y dependencia.** En la sección anterior hemos visto que la información mutua entre $X_H$ e $Y_H$ es pequeña para algún $H$. En esta sección vamos a ver cómo usar ese hecho para deducir que $X_H$ e $Y_H$ son más o menos independientes. Todo va a sustentarse sobre nuestra cota para la información mutua.

El esquema sería el siguiente. La variable $Y_H$ para $H$ pequeño es esencialmente uniforme. Por lo tanto, su entropía está muy cerca del máximo posible, y, como la información mutua con $X_H$ es pequeña, deducimos que la entropía $\mathbb{H}(Y_H|X_H)$ también va a estar muy cerca del máximo posible. Esto va a implicar que, con probabilidad casi 1, $X_H$ toma un valor $\vec{x}$ para el cual la entropía de la variable $Y_H$ condicionada a $X_H = \vec{x}$ está cerca del máximo posible. A su vez, eso debería implicar que la distribución de la variable $Y_H$ condicionada a $X_H = \vec{x}$ está cerca de ser uniforme, por lo que habríamos demostrado esencialmente la independencia de $X_H$ e $Y_H$.

En realidad, este esquema sólo demostrará la independencia "a efectos de esperanza", es decir,

$$\mathbb{E}(F(X_H, Y_H)) \sim \mathbb{E}(F(X_H, Y_H^*))$$

donde $Y_H^*$ es una variable con la misma distribución que $Y_H$ pero independiente de $X_H$. Empero, esta igualdad aproximada entre esperanzas es justo lo que necesitamos para acotar nuestras sumas.

Como ya hemos indicado (final de §5.2), la equidistribución de $Y_H$ condicionada a $X_H = \vec{x}$ que requerimos es bastante débil: sólo necesitamos mostrar que $Y_H$ tiende a evitar un pequeño conjunto, de tamaño acotado por la desigualdad (5.17). Debido a la forma de (5.17), el hecho que nuestra cota para la información (5.25) es mejor que la cota trivial por un factor de algo más que $\log H$ será crucial.

El primer paso es muy simple: ver que $Y_H$ es esencialmente uniforme. Sabemos por (5.7) que $Y_H$ toma valores en el conjunto

$$\Omega = \prod_{\epsilon H/2 < p \leq \epsilon H} \mathbb{Z}/p\mathbb{Z}. \tag{5.29}$$

Por el teorema chino de los restos, $\Omega$ es isomorfo a $\mathbb{Z}/M\mathbb{Z}$, donde $M = |\Omega| = \prod_{\epsilon H/2<p\leq \epsilon H} p$. Por lo tanto, si $|\Omega| < x/w$ entonces

$$\begin{aligned}\mathbb{P}(Y_H = \vec{y}) = \mathbb{P}\left(N \equiv y_p \bmod p \ \ \forall p \in \left(\frac{\epsilon H}{2}, \epsilon H\right]\right) &= \frac{1}{L} \sum_{\substack{\frac{x}{w} < n \leq x \\ n \equiv y_p (p) \forall p \in (\frac{\epsilon H}{2}, \epsilon H]}} \frac{1}{n} \\ = \frac{1}{L} \sum_{\substack{\frac{x}{w} < n \leq x \\ n \equiv y_* (|\Omega|)}} \frac{1}{n} &= \frac{1}{|\Omega|} \frac{\log w + O(\frac{1}{x/w|\Omega|})}{\log w + O(\frac{1}{x/w})} = \frac{1}{|\Omega|} + O\left(\frac{1}{x/w}\right),\end{aligned}$$

donde $L = \sum_{x/w<n\leq n} 1/n$ e $y_*$ es cualquier entero congruente a $y_p$ para cada $p \in (\epsilon H/2, \epsilon H]$. Por otra parte, por el teorema de los números primos,

$$|\Omega| = e^{\frac{\epsilon H}{2}} \left(1 + O\left((\log \epsilon H)^{-50}\right)\right) \ll e^{\frac{\epsilon H}{2}}. \tag{5.30}$$

Luego, tomando $\epsilon H < \log(x/w)$ vemos que

$$\mathbb{P}(Y_H = \vec{y}) = \frac{1}{|\Omega|}\left(1 + O\left(\frac{1}{|\Omega|}\right)\right), \tag{5.31}$$

es decir, la distribución de $Y_H$ es casi uniforme.

Deducimos mediante la cota general (5.27) que

$$\mathbb{H}(Y_H) = \left(1 + O\left(\frac{1}{|\Omega|}\right)\right) \log |\Omega|, \tag{5.32}$$

(El máximo posible sería $\log |\Omega|$.)

Apliquemos ahora el Corolario 5.7, con $H_-$ de forma que $H_+ \leq \frac{1}{9}\log x$. Obtenemos que la información mutua con $X_H$ es pequeña para cierto $H \in [H_-, H_+]$, y, por (5.20), concluímos que

$$\mathbb{H}(Y_H|X_H) \geq \mathbb{H}(Y_H) - \frac{H}{\log H \log_3 H}$$

para dicho $H$. Así, usando (5.32) y la cota (5.30), vemos que

$$\mathbb{H}(Y_H|X_H) \geq \left(1 - \frac{4}{\epsilon \log H \log_3 H}\right) \log |\Omega| \tag{5.33}$$

para $H_-$ más grande que una constante que depende sólo de $\epsilon$. En otras palabras, para algún $H \in [H_-, H_+]$, la entropía condicional $\mathbb{H}(Y_H|X_H)$ está cerca del máximo posible.

Digamos que un elemento $\vec{x} \in \{-1, 1\}^H$ es *bueno* si

$$\mathbb{H}(Y_H|X_H = \vec{x}) \geq \left(1 - \frac{4}{(\epsilon \log_3 H)^{3/4} \log H}\right) \log |\Omega|, \tag{5.34}$$

y en otro caso decimos que $\vec{x}$ es *malo*. Por la definición (5.18) de $\mathbb{H}(Y_H|X_H)$ en términos de $\mathbb{H}(Y_H|X_H = \vec{x})$, y por la desigualdad (5.33), vemos que

$$\left(1 - \frac{4}{\epsilon \log_3 H \log H}\right) \log|\Omega| \leq \sum_{\vec{x} \text{ malo}} \left(1 - \frac{4}{(\epsilon \log_3 H)^{3/4} \log H}\right) \log|\Omega| \cdot \mathbb{P}(X = \vec{x}) + \sum_{\vec{x} \text{ bueno}} \log|\Omega| \cdot \mathbb{P}(X = \vec{x})$$

ya que $\mathbb{H}(Y_H|X_H = \vec{x})$ es como máximo $\log|\Omega|$. De esta desigualdad deducimos la cota

$$\mathbb{P}(X_H \text{ malo}) \leq \frac{1}{(\epsilon \log_3 H)^{1/4}}. \tag{5.35}$$

Luego, es muy probable que $X_H$ sea igual a un valor bueno de $\vec{x}$, i.e., un valor de $\vec{x}$ tal que la entropía de la variable $Y_H$ condicionada a $X_H = \vec{x}$ estará cerca del máximo posible.

Para continuar, querríamos ver que la única distribución que está cerca de alcanzar el máximo de entropía es la uniforme. Esto sería cierto si su entropía está extremadamente cerca de dicho máximo, pero cuando hay un poco más de diferencia dicha unicidad no es cierta (ejercicio 5.9). Aun así, se puede decir que la variable no concentra mucho su masa, en el siguiente sentido.

**Lema 5.8.** *Sea $Y$ una variable aleatoria tomando valores es un conjunto $\Omega$ finito tal que $\mathbb{H}(Y) \geq (1-\delta)\log|\Omega|$, con $\frac{1}{\log|\Omega|} \leq \delta < 1$. Si un subconjunto $E \subset \Omega$ tiene tamaño*

$$|E| \leq |\Omega|^{1-M\delta} \tag{5.36}$$

*para algún $M > 0$, entonces*

$$\mathbb{P}(Y \in E) \leq \frac{2}{M}.$$

*Demostración.* Por la concavidad de la función $x \to -x \log x$, el ejercicio 5.10 nos da

$$\mathbb{H}(Y) \leq \mathbb{P}(E) \log \frac{|E|}{\mathbb{P}(E)} + \mathbb{P}(E^c) \log \frac{|E^c|}{\mathbb{P}(E^c)} \tag{5.37}$$

donde $\mathbb{P}(A) = \mathbb{P}(Y \in A)$. Escribiendo $p = \mathbb{P}(E)$ y usando $\mathbb{H}(Y) \geq (1-\delta)\log|\Omega|$, vemos que

$$(1-\delta)\log|\Omega| \leq p \log|E| + (1-p)\log|\Omega| - p\log p - (1-p)\log(1-p).$$

Ahora, usando de nuevo la concavidad con los dos últimos sumandos, así como la cota (5.36) sobre el tamaño de $E$, llegamos a

$$(1-\delta)\log|\Omega| \leq p(1-M\delta)\log|\Omega| + (1-p)\log|\Omega| + \log 2,$$

lo cual da

$$\delta(pM-1)\log|\Omega| \leq \log 2.$$

Luego, si se cumpliera $p > 2/M$, tendríamos $\delta \leq \log 2/\log|\Omega|$, en contradicción a una de las hipótesis del lema. □

El lema anterior indica que, para un $\vec{x}$ bueno, la variable $Y_H$ condicionada a $X_H = \vec{x}$ tiene una distribución que no concentra mucho su masa. Ésta es una forma muy débil de cuasiuniformidad, pero va a ser suficiente para estimar la esperanza de $F(X_H, Y_H)$, debido a (5.16) y (5.17).

**Teorema 5.9.** *Existe un $c > 0$ tal que lo siguiente se cumple. Sean $x > e^e$, $1 \le w \le x^{1/8}$ y $0 < \epsilon \le 1$. Sea $F$ como en (5.9). Para todo $3 \le H_- \le \exp(\exp(c\sqrt{\log_3 x}))$, hay un $H > H_-$, dependiendo sólo de $H_-$ y cumpliendo $\log_3 H \le 2\log_3 H_- + O(1)$, tal que, para $X_H$, $Y_H$, $N$ y $\Omega$ definidos como en (5.6), (5.7) y (5.29) con $K_0 = \epsilon H/2$ y $K_1 = \epsilon H$,*

$$\mathbb{E}(F(X_H, Y_H)) = \mathbb{E}(F(X_H, Y_H^*)) + O\left(\frac{H/\log H}{(\epsilon \log_3 H)^{1/4}}\right)$$

*para $Y_H^*$ una variable aleatoria con distribución uniforme en $\Omega$ e independiente de $X_H$.*

*Demostración.* Podemos asumir que $x$ y $\epsilon \log_3 H_-$ son más grandes que una constante, ya que si no el resultado es trivial. Aplicamos el Corolario 5.7 para obtener un $H_+$ con $\log_3 H_+ \le 2\log_3 H_- + O(1)$ y cierto $H \in [H_-, H_+]$ tal que (5.25) se cumple, i.e., tal que la información mutua entre $X_H$ e $Y_H$ es pequeña. Por la definición de esperanza,

$$\begin{aligned}\mathbb{E}(F(X_H, Y_H)) &= \sum_{\substack{\vec{x}\in\{-1,1\}^H \\ \vec{y}\in\Omega}} F(\vec{x},\vec{y})\,\mathbb{P}(X_H = \vec{x}, Y_H = \vec{y}) \\ &= \sum_{\vec{x}\in\{-1,1\}^H} \mathbb{P}(X_H = \vec{x}) \sum_{\vec{y}\in\Omega} F(\vec{x},\vec{y})\,\mathbb{P}(Y_H = \vec{y}|X_H = \vec{x}).\end{aligned}$$

Ahora dividimos el rango de $\vec{x}$ entre buenos y malos, acorde a la definición previa a (5.34). Acotamos la suma sobre los $\vec{x}$ malos fácilmente:

$$\begin{aligned}\sum_{\vec{x}\text{ malo}} \mathbb{P}(X_H = \vec{x}) \sum_{\vec{y}\in\Omega} F(\vec{x},\vec{y})\,\mathbb{P}(Y_H = \vec{y}|X_H = \vec{x}) \\ \ll \|F\|_\infty \mathbb{P}(X_H \text{ malo}) \ll \frac{H/\log H}{(\epsilon\log_3 H)^{1/4}}\end{aligned}$$

por (5.14) y (5.35). Ahora, para cada $\vec{x}$ bueno, dividimos el rango de $\vec{y}$ en $E_{\vec{x}}$ y $E_{\vec{x}}^c$, con $E_{\vec{x}}$ el conjunto de excepciones en (5.17) con $\mu^2 = (\epsilon\log_3 H)^{-1/2}$. Teniendo en cuenta (5.34), aplicamos el Lema 5.8 para obtener

$$\mathbb{P}(Y \in E_{\vec{x}}|X_H = \vec{x}) \ll \frac{1}{(\epsilon\log_3 H)^{1/4}},$$

de donde

$$\begin{aligned}\sum_{\vec{x}\text{ bueno}} \mathbb{P}(X_H = \vec{x}) \sum_{\vec{y}\in E_{\vec{x}}} F(\vec{x},\vec{y})\,\mathbb{P}(Y_H = \vec{y}|X_H = \vec{x}) \\ \ll \frac{\|F\|_\infty}{(\epsilon\log_3 H)^{1/4}} \ll \frac{H/\log H}{(\epsilon\log_3 H)^{1/4}}.\end{aligned}$$

Así, sólo nos quedan los $\vec{y}$ no excepcionales y los $\vec{x}$ buenos. En este caso usamos la estimación (5.16), y teniendo en cuenta las cotas ya obtenidas vemos que

$$\begin{aligned}\mathbb{E}(F(X_H, Y_H)) = O\left(\frac{H/\log H}{(\epsilon\log_3 H)^{1/4}}\right) \\ + \sum_{\vec{x}\text{ bueno}} \mathbb{P}(X_H = \vec{x}) \sum_{\vec{y}\notin E_{\vec{x}}} \overline{F(\vec{x},\cdot)}\,\mathbb{P}(Y_H = \vec{y}|X_H = \vec{x}).\end{aligned}$$

Ahora usamos de nuevo las cotas para $\mathbb{P}(Y \in E_{\vec{x}}|X_H = \vec{x})$, $\mathbb{P}(X_H \text{ malo})$ y $\|F\|_\infty$ para suplir los $(\vec{x}, \vec{y})$ que faltan y así obtener

$$\mathbb{E}(F(X_H, Y_H)) = O\left(\frac{H/\log H}{(\epsilon \log_3 H)^{1/4}}\right) + \sum_{\vec{x}\in\{-1,1\}^H} \mathbb{P}(X_H = \vec{x}) \sum_{\vec{y}\in\Omega} \overline{F(\vec{x},\cdot)}\, \mathbb{P}(Y_H = \vec{y}|X_H = \vec{x}).$$

Pero

$$\begin{aligned}\sum_{\vec{x}\in\{-1,1\}^H} \mathbb{P}(X_H = \vec{x}) \sum_{\vec{y}\in\Omega} \overline{F(\vec{x},\cdot)}\, \mathbb{P}(Y_H = \vec{y}|X_H = \vec{x}) &= \sum_{\vec{x}\in\{-1,1\}^H} \mathbb{P}(X_H = \vec{x})\overline{F(\vec{x},\cdot)} \\ &= \sum_{\vec{x}\in\{-1,1\}^H} \mathbb{P}(X_H = \vec{x}) \sum_{\vec{y}\in\Omega} F(\vec{x},\vec{y}) \frac{1}{|\Omega|} = \mathbb{E}(F(X_H, Y_H^*)).\end{aligned}$$

□

Ahora, por (5.12), tenemos que la esperanza del Teorema 5.9 se puede escribir como

$$\mathbb{E}(F(X_H, Y_H^*)) = \sum_{\epsilon H/2<p\le\epsilon H} \mathbb{E}(F_p(X_H, (Y_H^*)_p))$$

donde $Y_H^* = ((Y_H^*)_p)_{\epsilon H/2<p\le\epsilon H}$. Como $(Y_H^*)_p$ es independiente de $X_H$ y uniforme en $\mathbb{Z}/p\mathbb{Z}$ tenemos

$$\begin{aligned}\mathbb{E}(F_p(X_H, (Y_H^*)_p)) &= \sum_{t\in\mathbb{Z}/p\mathbb{Z}} \mathbb{E}(F_p(X_H, t))\mathbb{P}((Y_H^*)_p = t) = \frac{1}{p}\sum_{t\in\mathbb{Z}/p\mathbb{Z}} \mathbb{E}(F_p(X_H, t)) \\ &= \frac{1}{p}\sum_{t\in\mathbb{Z}/p\mathbb{Z}} \sum_{\vec{x}\in\{-1,1\}^H} F_p(\vec{x}, t)\mathbb{P}(X_H = \vec{x}) \\ &= \frac{1}{p}\sum_{\vec{x}\in\{-1,1\}^H} \mathbb{P}(X_H = \vec{x}) \sum_{j\le H-p} x_j x_{j+p} \sum_{t\in\mathbb{Z}/p\mathbb{Z}} 1_{j\equiv -t(p)}.\end{aligned}$$

Luego

$$\mathbb{E}(F(X_H, Y_H^*)) = \mathbb{E}(G(X_H))$$

con

$$G(\vec{x}) = \sum_{\epsilon H/2<p\le\epsilon H} \frac{1}{p} \sum_{j\le H-p} x_j x_{j+p}$$

y por la definición de $X_H$

$$\mathbb{E}(G(X_H)) = \sum_{\frac{x}{w}<n\le x} \frac{1}{nL} \sum_{\epsilon H/2<p\le\epsilon H} \frac{1}{p} \sum_{j\le H-p} \lambda(n+j)\lambda(n+j+p).$$

Ahora, procediendo como en la prueba del lema 5.3, podemos reescribir esta suma como

$$\frac{1}{L}\sum_{\epsilon H/2<p\le\epsilon H} \frac{1}{p}\left(H \sum_{x/w<n\le x} \frac{\lambda(n)\lambda(n+p)}{n} + O\left(H + p\log w\right)\right)$$

por lo que finalmente, por el teorema 5.9 (con $H_- = \exp(\exp(\min(\log_3 x, \log_2 w)^{1/3}))$, digamos) y el lema 5.4, obtenemos el resultado siguiente, el cual es lo que queríamos.

**Corolario 5.10.** *Sea $w \le x^{1/8}$, $w$ más grande que una constante. Sea $(\log_3 w)^{-1} \le \epsilon \le 1$. Entonces existe $H$ con $\log_3 H \gg \min(\log_4 x, \log_3 w)$ y $\log_3 H \le (3/4)\log_3 w$ tal que*

$$\sum_{\epsilon H/2 < p \le \epsilon H} \sum_{\substack{x/w < n \le x \\ p|n}} \frac{\lambda(n)\lambda(n+p)}{n} - \sum_{\epsilon H/2 < p \le \epsilon H} \frac{1}{p} \sum_{x/w < n \le x} \frac{\lambda(n)\lambda(n+p)}{n} = O\left(\epsilon + \frac{1}{(\epsilon \log_3 H)^{1/4}}\right) \cdot \frac{\log w}{\log H}. \tag{5.38}$$

La cota superior sobre $\log_3 H$ (la cual, por cierto, es de lejos más fuerte de lo que necesitamos) nos será útil cuando tengamos que aplicar el Lema 5.2 y el Corolario 5.15.

*Ejercicios.*

**Ejercicio 5.9.** Sea $\Omega$ un conjunto finito y $E \subset \Omega$ con $|E| = |\Omega|^{1-\delta} \le |\Omega|/2$. Sea $Y$ la variable aleatoria en $\Omega$ que satisface $\mathbb{P}(Y \in E) = \mathbb{P}(Y \in E^c) = 1/2$ y tal que $Y$ es uniforme en $E$ y también en $E^c$. Observe que $Y$ está muy lejos de ser uniforme en $\Omega$ pero que sin embargo su entropía es grande:

$$\mathbb{H}(Y) = (1-\delta)\log|\Omega| + O(1).$$

**Ejercicio 5.10.** Una función cóncava $f$ cumple $f((1-t)a+tb) \ge (1-t)f(a)+tf(b)$.

1. Demuestre la desigualdad
$$\frac{f(a_1)+f(a_2)+\ldots+f(a_n)}{n} \le f\left(\frac{a_1+a_2+\ldots+a_n}{n}\right) \tag{5.39}$$
para cualquier función cóncava y $n \in \mathbb{N}$.
2. Sea $f : I \to \mathbb{R}$ una función doblemente diferenciable en un intervalo $I \subset \mathbb{R}$. Asuma que $f''(t) \le 0$ para todo $t \in I$. Muestre que $f$ es cóncava.
3. Muestre que la función $x \mapsto x\log(1/x)$ es cóncava en $(0,\infty)$, y por lo tanto en $[0,\infty)$ (definiendo que el valor de $x\log(1/x)$ para $x = 0$ es 0). Usando la desigualdad (5.39), pruebe que
$$\sum_{y \in A} \mathbb{P}(Y=y)\log\frac{1}{\mathbb{P}(Y=y)} \le \mathbb{P}(Y \in A)\log\frac{|A|}{\mathbb{P}(Y \in A)}$$
para cualquier variable aleatoria $Y$ sobre $\Omega$ y subconjunto $A \subset \Omega$.
4. Demuestre la desigualdad (5.37).

**Ejercicio 5.11.** Sea $Y$ una variable aleatoria en $\Omega$ tal que $\mathbb{H}(Y) = \log|\Omega| - o(1/|\Omega|)$.

1. Sea $y \in Y$. Muestre usando (5.37) con $E = \{y\}$ que
$$\log|\Omega| - o\left(\frac{1}{|\Omega|}\right) \le -p\log p + (1-p)\log\frac{|\Omega|-1}{1-p}$$
con $p = \mathbb{P}(Y=y)$, y que, por lo tanto,
$$p\log|\Omega| + \frac{1-p}{|\Omega|} - o\left(\frac{1}{|\Omega|}\right) \le -p\log p - (1-p)\log(1-p). \tag{5.40}$$

2. Utilizando que la parte derecha de (5.40) está acotada por $\log 2$, pruebe que $p = o(1)$, y además, usando también que $p = o(1)$, pruebe que
$$t \log t + 1 - o(1) \leq t(1 + o(1))$$
con $t = p|\Omega|$.
3. Demuestre que para $t \neq 1$ tenemos $t \log t + 1 - t > 0$. Concluya que
$$p = \frac{1}{|\Omega|}(1 + o(1)),$$
es decir, $Y$ se comporta asintóticamente como una variable uniforme.

**Ejercicio 5.12.** Sea $w(n)$ el número de divisores primos distintos de $n$. Demuestre que
$$S = \frac{1}{x} \sum_{n \leq x} w(n)\lambda(n+1) = o(\log\log x) = o\left(\frac{1}{x} \sum_{n \leq x} w(n)\right)$$
usando el Teorema 2.3 y el ejercicio 5.4 de 5.2. Para ponerlo de otra manera,
$$S = \mathbb{E}(w\lambda_+) + O(1)$$
con $w$ la variable $w = \sum_{p \leq x^{1/4}} 1_{N \equiv 0 (p)}$ y $\lambda_+ = \lambda(N+1)$, donde $\mathbb{P}(N = n) = 1/x$ para todo $n \in \Omega = [1, x]$.

**Ejercicio 5.13.** Sean $w, \lambda_+$ las variables aleatorias del problema anterior.

1. $\lambda_+$ toma valores en el conjunto $\{-1, 1\}$. Use el Teorema 2.3 para demostrar que su entropía está cerca de la máxima posible, en el sentido que
$$\mathbb{H}(\lambda_+) = (1 + o(1)) \log 2.$$
2. Demuestre usando el teorema de los números primos que $w$ toma valores enteros en un intervalo $\Omega = [0, z)$, con $z \sim \frac{\log x}{\log\log x}$. Observe que la máxima entropía que $w$ podría tener es $\log |\Omega| \sim \log\log x$.
3. Use (5.37) con $E = \{w < 2 \log\log x\}$ y el ejercicio 5.4 de 5.2 para demostrar
$$\mathbb{H}(w) = o(\log |\Omega|) = o(\log\log x)$$
y que por lo tanto su entropía está lejos de la máxima posible.

**5.5. Sumas de $\lambda(n)\lambda(n+p)$, en promedio sobre $p$. Conclusión.** Por lo visto en la sección anterior, para obtener (5.1) sólo nos queda controlar sumas del tipo
$$\sum_{p \leq h} \sum_{X < n \leq 2X} \lambda(n)\lambda(n+p).$$
Estas son más complicadas que las del Corolario 4.10
$$\sum_{j \leq h} |\sum_{X < n \leq 2X} \lambda(n)\lambda(n+j)|^2$$
en el sentido de ahora sólo sumamos sobre primos, pero más sencillas ya que no tenemos el cuadrado. Vamos a ver que de nuevo es posible comprenderlas en términos de los coeficiente de Fourier de $\lambda(n)$ en intervalos cortos. La manera de hacerlo va a ser usar el método del círculo, el cual, en breve, consiste en usar la identidad
$$1_{k=0} = \int_0^1 e(k\alpha)\, d\alpha \qquad k \in \mathbb{Z} \tag{5.41}$$

para reescribir una ecuación aditiva (como $m = n + p$) de forma analítica, usando los armónicos $e(k\alpha)$.

**Lema 5.11.** *Si $|w_j| \leq 1$ tenemos que*

$$\sum_{p\leq h}\sum_{j\leq h} w_{j+p}\overline{w_j} \ll \epsilon\frac{h^2}{\log h} + \frac{h^2}{\log h}\int_{\mathfrak{M}_\epsilon}\left|\sum_{b\leq h} w_b e(b\alpha)\right| d\alpha$$

*para cualquier $\epsilon \in (0,1)$, con $\mathfrak{M}_\epsilon = \{\alpha \in [0,1] : |\sum_{p\leq h} e(p\alpha)| > \epsilon\frac{h}{\log h}\}$.*

*Demostración.* Usando la identidad (5.41) para $k = m - j - p$ tenemos

$$\sum_{p\leq h}\sum_{j\leq h}\overline{w_j}w_{j+p} = \sum_{m\leq 2h}\sum_{p\leq h}\sum_{j\leq h} w_m\overline{w_j}\int_0^1 e((m-j-p)\alpha)\,d\alpha,$$

y sacando fuera la integral y factorizando obtenemos

$$\sum_{p\leq h}\sum_{j\leq h}\overline{w_j}w_{j+p} = \int_0^1 W_{2h}(\alpha)\overline{W_h(\alpha)P_h(\alpha)}\,d\alpha,$$

con $W_d(\alpha) = \sum_{b\leq d} w_b e(b\alpha)$ y $P_d(\alpha) = \sum_{p\leq d} e(p\alpha)$. Ahora dividimos la integral entre los «arcos mayores» $\mathfrak{M}_\epsilon$ y los «arcos menores» $\mathfrak{m}_\epsilon = [0,1)\setminus\mathfrak{M}_\epsilon$. Para la parte de los arcos mayores tenemos

$$\begin{aligned}\int_{\mathfrak{M}_\epsilon} W_{2h}(\alpha)\overline{W_h(\alpha)}P_h(-\alpha)\,d\alpha&\\ \ll \frac{h^2}{\log h}\int_{\mathfrak{M}_\epsilon}|W_h(\alpha)|\,d\alpha &= \frac{h^2}{\log h}\int_{\mathfrak{M}_\epsilon}|\sum_{b\leq h} w_b e(b\alpha)|\,d\alpha.\end{aligned}$$

Para la parte de los arcos menores tenemos que

$$\left|\int_{\mathfrak{m}_\epsilon} W_{2h}(\alpha)\overline{W_h(\alpha)}P_h(-\alpha)\,d\alpha\right| \leq \frac{\epsilon h}{\log h}\int_0^1 |W_{2h}(\alpha)||W_h(\alpha)|\,d\alpha.$$

Usando la desigualdad $|ab| \leq |a|^2 + |b|^2$ vemos que

$$\int_0^1 |W_{2h}(\alpha)||W_h(\alpha)|\,d\alpha \leq \int_0^1 |W_{2h}(\alpha)|^2\,d\alpha + \int_0^1 |W_h(\alpha)|^2\,d\alpha$$

y por Parseval (ecuación (4.9))

$$\int_0^1 |W_h(\alpha)|^2\,d\alpha = \sum_{j\leq h}|w_j|^2 \leq h$$

y lo mismo para $W_{2h}$, luego obtenemos la cota $O(\epsilon h^2/\log h)$ para la parte de los arcos menores.

□

Ahora vamos a ver que la parte que queda (arcos menores) es muy pequeña, por lo que la integral sobre esa parte también va a ser pequeña.

**Lema 5.12.** *Sea $0 < \epsilon < 1$, $h > 1$. Con las definiciones del lema anterior tenemos que*

$$|\mathfrak{M}_\epsilon| \ll \frac{1}{\epsilon^4}\frac{1}{h}.$$

*Demostración.* La idea, igual que en la demostración del Lema 3.12, es controlar algún promedio de $P_h(\alpha) = \sum_{p\leq h} e(p\alpha)$ para concluir que $\mathfrak{M}_\epsilon$ es un conjunto pequeño. La media cuadrática no será suficiente (ver problema 5.14), pero sí la potencia cuarta. Tenemos que

$$\int_0^1 |P_h(\alpha)|^4\, d\alpha = \int_0^1 |P_h^2(\alpha)|^2\, d\alpha = \int_0^1 |\sum_{|j|\leq h} (\sum_{\substack{q-p=j\\ p,q\leq h}} 1) e(j\alpha)|^2\, d\alpha,$$

donde $p, q$ se mueven sobre los primos (hemos agrupado las frecuencias $e(q\alpha)\overline{e(p\alpha)} = e((q-p)\alpha)$). Ahora aplicamos Parseval (ecuación 4.9) para obtener

$$\int_0^1 |P_h(\alpha)|^4\, d\alpha = \sum_{|j|\leq h} |\sum_{\substack{q-p=j\\ p,q\leq h}} 1|^2.$$

Para $j = 0$ tenemos que la suma interior es $O(h/\log h)$ por el teorema de los números primos. Para el resto de $j$s podemos usar la cota de criba (2.31) y así obtener

$$\sum_{|j|\leq h} |\sum_{\substack{q-p=j\\ p,q\leq h}} 1|^2 \ll \frac{h^2}{(\log h)^2} + \frac{h^2}{(\log h)^4} \sum_{j=1}^{h} \prod_{p|j} \left(1 + \frac{1}{p}\right).$$

Como $\sum_{j=1}^{h} \prod(1 + 1/p) \ll h$ (ejercicio 5.15), obtenemos

$$\int_0^1 |P_h(\alpha)|^4\, d\alpha \ll \frac{h^3}{(\log h)^4},$$

de donde se deduce la cota para $|\mathfrak{M}_\epsilon|$.

□

Juntando los dos últimos lemas podemos concluir que los promedios de $\lambda(n+p)\lambda(n)$ están controlados por los coeficientes de Fourier de $\lambda(n)$ en intervalos cortos.

**Proposición 5.13.** *Sea $0 < \epsilon < 1$ y $1 \leq h \leq \epsilon X$. Entonces*

$$\sum_{p\leq h} \sum_{X<n\leq 2X} \lambda(n+p)\lambda(n) \ll \frac{\epsilon h X}{\log h} + \frac{1}{\epsilon^4 \log h} \max_{0\leq\alpha\leq 1} \int_X^{2X} \left| \sum_{x<m\leq x+h} \lambda(m) e(m\alpha) \right| dx.$$

*Demostración.* Comenzamos por la observación de que si desplazamos el intervalo de sumación en $n$ un poco la suma total casi no varía:

$$\begin{aligned} \sum_{X<n\leq 2X} \lambda(n+p)\lambda(n) &= O(j) + \sum_{X+j<n\leq 2X+j} \lambda(n+p)\lambda(n) \\ &= O(j) + \sum_{X<n\leq 2X} \lambda(n+j+p)\lambda(n+j). \end{aligned}$$

Ahora usamos esa ecuación para todo $j \leq h$, obteniendo

$$\sum_{X<n\leq 2X} \lambda(n+p)\lambda(n) = O(h) + \frac{1}{h} \sum_{j\leq h} \sum_{X<n\leq 2X} \lambda(n+j+p)\lambda(n+j).$$

Así, sumando en $p$ e intercambiando el orden de sumación, por TNP tenemos (5.42)

$$\sum_{p\leq h} \sum_{X<n\leq 2X} \lambda(n+p)\lambda(n) = O\left(\frac{h^2}{\log h}\right) + \frac{1}{h} \sum_{X<n\leq 2X} \sum_{p\leq h} \sum_{j\leq h} \lambda(n+j+p)\lambda(n+j).$$

Ahora usamos el Lema 5.11 con $w_j = \lambda(n+j)$ y llegamos a

(5.43)
$$\sum_{p\le h}\sum_{X<n\le 2X}\lambda(n+p)\lambda(n) \ll \frac{\epsilon h X}{\log h} + \frac{h}{\log h}\int_{\mathfrak{M}_\epsilon}\sum_{X<n\le 2X}|\sum_{b\le h}\lambda(n+b)e(b\alpha)|\,d\alpha.$$

Teniendo en cuenta que

$$\left|\sum_{b\le h}\lambda(n+b)e(b\alpha)\right| = \left|\sum_{n<m\le n+h}\lambda(m)e(m\alpha)\right|$$

y el Lema 5.12 obtenemos el resultado buscado. □

Usando las cotas para los coeficientes de Fourier de $\lambda(n)$ en intervalos cortos de la sección 4, vemos ahora que hay cancelación en las sumas de $\lambda(n+p)\lambda(n)$.

**Corolario 5.14.** *Sea* $\log h \le (\log X)^{1/3}$. *Entonces*

$$\sum_{p\le h}\sum_{X<n\le 2X}\lambda(n)\lambda(n+p) \ll \frac{1}{(\log h)^{\frac{1}{75}-o(1)}}\cdot\frac{hX}{\log h}.$$

*Demostración.* Por la Proposición 5.13 y el Teorema 4.9 tenemos que

$$\sum_{p\le h}\sum_{X<n<2X}\lambda(n)\lambda(n+p) \ll \frac{\epsilon h X}{\log h} + \frac{hX}{\epsilon^4\log h}\cdot\frac{1}{(\log h)^{\frac{1}{15}-o(1)}}$$

para cualquier $0<\epsilon<1$ y cualquier $1\le h\le \epsilon X$ con $\log h \le (\log X)^{1/3}$. Tomando $\epsilon = (\log h)^{-\frac{1}{75}}$ obtenemos el resultado. □

El resultado que necesitamos se deduce fácilmente del Corolario 5.14.

**Corolario 5.15.** *Sean* $1\le w\le\sqrt{x}$ *y* $1<h\le w$ *tal que* $\log h \le (\log\frac{x}{w})^{1/3}$. *Entonces*

$$\sum_{h/2<p\le h}\frac{1}{p}\sum_{x/w<n\le x}\frac{\lambda(n)\lambda(n+p)}{n} \ll \frac{1}{(\log h)^{\frac{1}{75}-o(1)}}\cdot\frac{\log w}{\log h}.$$

*Demostración.* Ejercicio 5.17. □

Llegamos a la prueba del resultado principal.

*Prueba del teorema 5.1.* Podemos suponer que $w\le x^{1/8}$ (pues podemos reducir a este caso subdividiendo el rango $x/w<n\le x$) y también que $w$ es más grande que una constante (pues, de lo contrario, la cota que debemos probar es trivial). Apliquemos el Corolario 5.10 con $\epsilon = (\log_3 H)^{-1/5}$ y el Corolario 5.15 con $h=\epsilon H$ para obtener que

$$\sum_{\epsilon H/2<p\le\epsilon H}\sum_{\substack{x/w<n\le x\\ p|n}}\frac{\lambda(n)\lambda(n+p)}{n} \ll \frac{1}{(\log_3 H)^{1/5}}\cdot\frac{\log w}{\log H}.$$

Luego aplicamos el Lema 5.2 con $K_0=\epsilon H/2$, $K_1=\epsilon H$ para concluír que

$$\sum_{\frac{x}{w}<n\le x}\frac{\lambda(n)\lambda(n+1)}{n} \ll \log\epsilon H\cdot\frac{1}{(\log_3 H)^{1/5}}\cdot\frac{\log w}{\log H} + \log\epsilon H \ll \frac{\log w}{(\log_3 H)^{1/5}}.$$

□

*Ejercicios.*

**Ejercicio 5.14.** Demuestre que

$$\int_0^1 |\sum_{p<h} e(p\alpha)|^2 \, d\alpha = \frac{h}{\log h}(1 + o(1)).$$

Deduzca que $|\mathfrak{M}_\epsilon| \ll \frac{\log h}{\epsilon^2 h}$, con $\mathfrak{M}_\epsilon$ definido como en el Lema 5.11. Observe que esta cota es peor que la obtenida en el Lema 5.12 para $1/\sqrt{\log h} < \epsilon < 1$.

**Ejercicio 5.15.** Vamos a demostrar que

$$\sum_{j \le h} \prod_{p|j} \left(1 + \frac{1}{p}\right)^C \ll_C h.$$

Para ello, pruebe las siguientes desigualdades e identidades, con $w(d) = \sum_{p:p|d} 1$:

$$\begin{aligned}\sum_{j \le h} \prod_{p|j} \left(1 + \frac{1}{p}\right)^C &\ll_C \sum_{j \le h} \prod_{p|j} \left(1 + \frac{2C}{p}\right) \\ &= \sum_{d \le h} \frac{(2C)^{w(d)}}{d} \sum_{\substack{j \le h \\ d|j}} 1 \le h \sum_{d=1}^{\infty} \frac{(2C)^{w(d)}}{d^2} \ll_C h.\end{aligned}$$

**Ejercicio 5.16.** También es factible demostrar el Lema 5.12 estudiando la suma $P(\alpha) = \sum_{p \le h} e(p\alpha)$ para todo $\alpha$ y viendo cuando es grande. Es posible ver que el conjunto $\mathfrak{M}_\epsilon$ de los valores para los cuales $P(\alpha)$ es grande consiste en los $\alpha$ cercanos a racionales con denominador pequeño. En una dirección (mostrar que los $\alpha$ que no estan cerca a tales racionales no están en $\mathfrak{M}_\epsilon$), esto no es nada fácil; se trata de la parte principal de la estrategia de Vinogradov para el problema ternario de Goldbach. Veamos como demostrar la otra dirección, por lo menos para algunos racionales de denominador pequeño.

1. Demuestre que si $\alpha = \delta/h$, $|\delta| < 1$, entonces $|P(\alpha)| = \frac{h}{\log h}(1 + o_h(1) + O(\delta))$. Observe que esto demuestra que $|\mathfrak{M}_\epsilon| \gg h$, lo cual coincide con la cota superior que se obtiene en el Lema 5.12 para $\epsilon \gg 1$.
2. Demuestre que, de todos los carácteres $\chi$ módulo 5, el único cuya función $L(s, \chi)$ tiene un polo en $s = 1$ es el carácter trivial $\chi_0$ que vale 1 para todo número no divisible por 5. (Sugerencia: para los otros carácteres $\chi \bmod 5$, utilice sumación por partes para estimar $\sum_n \chi(n) n^{-\sigma}$, $\sigma = 1 + \epsilon$, recordando que $\sum_{n \le u} \chi(n) < 5$ para todo $u$ (¿por qué?).)
3. Usando el apartado anterior y el Teorema 4.3, muestre que

$$\sum_{p \le h, p \equiv b(5)} 1 = \frac{1}{4} \frac{h}{\log h}(1 + o_h(1)) = \frac{1}{4}(1 + o_h(1)) \sum_{p \le h} 1$$

   para $b = 1, 2, 3, 4$.
4. Pruebe usando los apartados anteriores que si $\alpha = \frac{1}{5} + \frac{\delta}{h}$ con $|\delta| < 1$, entonces

$$P(\alpha) = -\frac{1}{4} \frac{h}{\log h}(1 + O(\delta) + o_h(1)).$$

**Ejercicio 5.17.** Deseamos mostrar que el Corolario 5.15 se deduce del Corolario 5.14. El procedimiento es sencillo y muy general.

1. Sea $f : \mathbb{Z}^+ \to \mathbb{C}$ arbitrario. Sea $F(t) = \sum_{p \le t} f(p)$. Muestre que
$$\sum_{\frac{h}{2} < p \le h} \frac{f(p)}{p} = \int_{h/2}^{h} \frac{F(t)}{t^2} dt + \frac{1}{h} F(h) - \frac{2}{h} F(h/2).$$
2. Sea $g : \mathbb{Z}^+ \to \mathbb{C}$ tal que $|g(n)| \le 1$ para todo $n$. Sea $G(t) = \sum_{t < n \le 2t} g(n)$. Muestre que, para $1 \le x_0 \le x_1$,
$$\sum_{x_0 < n \le x_1} \frac{g(n)}{n} = \int_{x_0}^{x_1} \frac{G(t)}{t^2} dt + O(1).$$
3. Usando (2) y el Corolario 5.14, pruebe que
$$\sum_{p \le h} \sum_{x/w < n \le x} \frac{\lambda(n)\lambda(n+p)}{n} \ll \frac{1}{(\log h)^{\frac{1}{75} - o(1)}} \frac{h \log w}{\log h} + \frac{h}{\log h} \tag{5.44}$$
para $1 \le w \le x$ y $\log h \ge (\log \frac{x}{w})^{1/3}$. Está claro que el término $h/\log h$ puede omitirse para $h \le w$.
4. Use (5.44) y el apartado (1) para deducir el Corolario 5.15. (Podemos, por cierto, suponer que $h$ es más grande que una constante, pues de lo contrario lo que queremos probar es trivial.)

## Referencias


[1] C. E. Bonferroni, *Teoria statistica delle classi e calcolo delle probabilita.*, (Pubbl. d. R. Ist. Super. di Sci. Econom. e Commerciali di Firenze. 8) Firenze: Libr. Internaz. Seeber. 62 S. (1936)., 1936.

[2] J. Bourgain, P. Sarnak, and T. Ziegler, *Disjointness of Moebius from horocycle flows*, From Fourier analysis and number theory to Radon transforms and geometry, Dev. Math., vol. 28, Springer, New York, 2013, pp. 67–83.

[3] V. Brun, *Über das Goldbachsche Gesetz und die Anzahl der Primzahlpaare.*, Arch. Math. Naturvid. **34** (1915), no. 8, 3–19 (German).

[4] J. R. Chen, *On the representation of a larger even integer as the sum of a prime and the product of at most two primes*, Sci. Sinica **16** (1973), 157–176.

[5] H. Daboussi and H. Delange, *On multiplicative arithmetical functions whose modulus does not exceed one*, J. London Math. Soc. (2) **26** (1982), no. 2, 245–264.

[6] J. Friedlander and H. Iwaniec, *Opera de cribro*, American Mathematical Society Colloquium Publications, vol. 57, American Mathematical Society, Providence, RI, 2010.

[7] H. Halberstam and H.-E. Richert, *Sieve methods*, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], London-New York, 1974, London Mathematical Society Monographs, No. 4.

[8] M. N. Huxley, *On the difference between consecutive primes*, Invent. Math. **15** (1972), 164–170.

[9] H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004.

[10] I. Kátai, *A remark on a theorem of H. Daboussi*, Acta Math. Hungar. **47** (1986), no. 1-2, 223–225.

[11] K. Matomäki and M. Radziwiłł, *Multiplicative functions in short intervals*, Ann. of Math. (2) **183** (2016), no. 3, 1015–1056.

[12] K. Matomäki, M. Radziwiłł, and T. Tao, *An averaged form of Chowla's conjecture*, Algebra Number Theory **9** (2015), no. 9, 2167–2196.

[13] ———, *Sign patterns of the Liouville and Möbius functions*, Forum Math. Sigma **4** (2016), e14, 44.

[14] K. Matomäki, M. Radziwiłł, and T. Tao, *Correlations of the von Mangoldt and higher divisor functions II. Divisor correlations in short ranges*, Mathematische Annalen (2017), 1–48.

[15] K. Matomäki, M. Radziwiłł, and T. Tao, *Correlations of the von Mangoldt and higher divisor functions I. Long shift ranges.*, Proc. Lond. Math. Soc. (3) **118** (2019), no. 2, 284–350 (English).
[16] H. L. Montgomery, *Topics in multiplicative number theory.*, vol. 227, Springer, 1971 (English).
[17] H. L. Montgomery, *Ten lectures on the interface between analytic number theory and harmonic analysis*, CBMS Regional Conference Series in Mathematics, vol. 84, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1994.
[18] H. L. Montgomery and R. C. Vaughan, *Exponential sums with multiplicative coefficients*, Invent. Math. **43** (1977), no. 1, 69–82.
[19] ———, *Multiplicative number theory. I. Classical theory*, Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, Cambridge, 2007.
[20] Y. Motohashi, *On the sum of the Möbius function in a short segment*, Proc. Japan Acad. **52** (1976), no. 9, 477–479.
[21] K. Ramachandra, *Some problems of analytic number theory*, Acta Arith. **31** (1976), no. 4, 313–324.
[22] A. A. Rényi, *On the representation of an even number as the sum of a single prime and a single almost-prime number*, Doklady Akad. Nauk SSSR (N.S.) **56** (1947), 455–458.
[23] J. B. Rosser and L. Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois J. Math. **6** (1962), 64–94.
[24] K. Soundararajan, *The Liouville function in short intervals*, Astérisque (2017), no. 390, Exp. No. 1119, 453–479, Séminaire Bourbaki. Vol. 2015/2016. Exposés 1104–1119.
[25] T. Tao, *The Erdős discrepancy problem*, Discrete Anal. (2016), Paper No. 1, 29.
[26] ———, *The logarithmically averaged Chowla and Elliott conjectures for two-point correlations*, Forum Math. Pi **4** (2016), e8, 36.

Georg-August Universität Göttingen, Mathematisches Institut, Bunsenstrasse 3-5, D-37073 Göttingen, Alemania

Université Paris Diderot, UFR de Mathématiques, case 7012, 75205 Paris CEDEX 13, Francia
*Email address*: harald.helfgott@gmail.com

Departamento de Matemáticas, Universidad Autónoma de Madrid, Madrid 28049 SPAIN
*Email address*: adrian.ubis@uam.es

Curso

# VARIEDADES ABELIANAS, UNA INTRODUCCIÓN

MARC HINDRY, MARUSIA REBOLLEDO, Y DAVID ROBERTS

Publicaciones Matemáticas del Uruguay
Volumen 18, Diciembre 2023, Páginas 287–365
S 0797-1443

# VARIEDADES ABELIANAS, UNA INTRODUCCIÓN

MARC HINDRY, MARUSIA REBOLLEDO, Y DAVID ROBERTS

Resumen. Variedades abelianas son grupos algebraicos que, al mismo tiempo, son variedades algebraicas proyectivas. El primer ejemplo es dado por curvas elípticas que son las variedades abelianas de dimensión uno. Un ejemplo histórico y muy importante es la variedad jacobiana de una curva de género $\geqslant 2$. Este curso propone una breve introducción a la rica teoría de estos objetos, esbozando tres puntos de vista: complejo analítico (toros complejos, funciones theta, formas de Riemann), geométrico algebraico (teorema del cubo, grupo de Picard, isogenías) y aritmético (teorema de Mordell-Weil, teoría de Honda-Tate, modularidad).

## Índice



Versión final: 18 de mayo de 2019.
Estas notas corresponden al curso dictado por los autores en la escuela AGRA III, Aritmética, Grupos y Análisis, del 9 al 20 de Julio de 2018 en Córdoba, Argentina.

## Introducción

Variedades abelianas son grupos algebraicos que, al mismo tiempo, son variedades algebraicas proyectivas. El primer ejemplo es dado por curvas elípticas que son las variedades abelianas de dimensión uno. Un ejemplo histórico y muy importante es la variedad jacobiana de una curva de género $\geqslant 2$. Empezamos con la exploración del caso de variedades complejas. Este caso es más concreto, pues cada variedad abeliana compleja puede ser presentada como un toro complejo $\mathbb{C}^g/\Lambda$ donde $\Lambda$ es un retículo $\cong \mathbb{Z}^{2g}$ dotado de una estructura adicional, una forma de Riemann. La segunda parte presenta la teoría del punto de vista de la geometría algebraica, es decir que se consideran variedades definidas sobre un cuerpo $K$; se demuestra que una gran parte de la geometría compleja puede ser recuperada. Como transición hacia la parte aritmética se demuestra el teorema de Mordell-Weil: el grupo $A(K)$ de los puntos de una variedad abeliana definida sobre un cuerpo de números $K$ es un grupo de tipo finito. La tercera y última parte presenta una descripción aritmética de variedades abelianas sobre un cuerpo finito $\mathbb{F}_q$ y sobre el cuerpo racional $\mathbb{Q}$.

A pesar de no dar todas las pruebas, las dos primeras partes presentan material clásico y básico, la última parte tiene un sabor distinto, presentando material contemporáneo de investigación. Otra característica de la última parte es el uso de computadores, por ejemplo del código en *Magma*. De hecho la clasificación explícita de variedades abelianas no es una cuestión puramente matemática. Esta clasificación explícita por medio de funciones $L$ es el objetivo principal de la base de datos *L-functions and modular forms database*.

La tercera parte de este curso también sirve como una introducción a la LMFDB, ya que cada sección corresponde directamente a partes particulares de la base de datos.

Se puede encontrar referencias generales sobre variedades abelianas complejas en [1, 7, 12, 14], variedades abelianas y jacobianas sobre un cuerpo cualquiera en [4, 7, 9, 10, 12, 13]. Para variedades abelianas con dimensión 1, es decir curvas elípticas en [6, 15, 16]. Se reúne material más avanzado sobre variedades abelianas en [27, 30, 35, 37, 38], información computacional y base de datos sobre curvas elípticas y variedades abelianas de dimensión 2 en [18, 19, 20, 22, 28, 34].

*Agradecimientos.* Los tres autores desean expresar su gratitud a los organizadores de la escuela y en particular a Emilio Lauret por su apoyo lingüístico. También agradecen al referí anónimo por sus observaciones precisas.

David Roberts fue apoyado por la subvención DMS-1601350 de la NSF.

## Parte 1. Variedades abelianas complejas

**Definición 0.1.** Una *variedad abeliana* es un grupo algebraico conexo que es también una variedad proyectiva.[1]

[1]La definición usual sería un grupo algebraico conexo y completo, pero una variedad proyectiva es siempre completa y, además, la recíproca es verdad para una variedad abeliana. Este hecho es no obstante no trivial y no lo queremos demostrar.

Recordamos que un grupo algebraico sobre un cuerpo $k$ es una variedad $A$ junto con aplicaciones regulares $m : A \times_k A \to A$ y $inv : A \to A$ y un elemento $e \in A(k)$ que satisfacen los axiomas de grupos usuales. Por lo tanto, definen una estructura de grupo sobre $A(\bar{k})$ con elemento neutro $e$.

**Ejemplo 0.2.** Vieron, en el curso de teoría de Galois, que las curvas algebraicas definidas por una ecuación afín de la forma $y^2 = x^3 + ax + b$ con $4a^3 + 27b^2 \neq 0$ tienen una estructura de grupos algebraicos. Así estas curvas, llamadas *curvas elípticas* son variedades abelianas de dimensión 1.

En esta primera parte, consideramos variedades abelianas definidas sobre el cuerpo $\mathbb{C}$ de los números complejos. Veremos que las variedades abelianas sobre $\mathbb{C}$ son toros complejos. Luego, vamos a examinar si todos los toros complejos son variedades abelianas complejas. Las referencias principales para esta parte son: [1, 12, 7, 14] y [9].

## 1. Toros complejos

**1.1. Variedades abelianas complejas son toros complejos.** Sea $A$ una variedad abeliana compleja. Entonces el conjunto $A(\mathbb{C})$ de los puntos complejos tiene una estructura de grupo de Lie complejo, o sea una variedad compleja donde las operaciones de grupo $m, inv$ son aplicaciones holomorfas. Este grupo de Lie es además conexo y compacto.[2]

Veremos en esta sección (Proposición 1.3) que eso implica que: 1. la ley de grupo sobre $A$ es conmutativa; 2. $A(\mathbb{C})$ es un toro complejo, es decir el cociente de un $\mathbb{C}$-espacio vectorial de dimensión finita por un retículo $\Lambda$. Referencia principal: [12].

*1.1.1. Exponencial de un grupo de Lie complejo.* Recordamos, sin prueba, algunos resultados clásicos de teoría de los grupos de Lie. Sea $T$ un grupo de Lie complejo, con elemento neutro $e$. Denotamos $V = Lie(T) = \mathrm{Tan}_e(T)$ el espacio tangente a $T$ en $e$ ; es el álgebra de Lie asociada a $T$. Es un espacio vectorial de dimensión igual a la dimensión de $T$ como variedad compleja.

Por cada vector tangente $v \in V$, hay un único morfismo $\lambda_v : \mathbb{C} \to T$ tal que $\lambda_v(0) = e$ y $(d\lambda_v)_0 : \mathrm{Tan}_0(\mathbb{C}) \to V$ manda el generador canónico $\left(\frac{\partial}{\partial t}\right)_0$ (la derivación en cero) de $\mathrm{Tan}_0(\mathbb{C})$ sobre $v$.[3]

**Definición 1.1.** La *aplicación exponencial* $\exp_T = \exp : V \to T$ es definida por $\exp(v) = \lambda_v(1)$ para todo $v \in V$.

La unicidad de $\lambda_v$ para cada $v$ permite demostrar que para cada $v \in V, s \in \mathbb{C}, t \in \mathbb{C}$, $\lambda_v(st) = \lambda_{tv}(s)$. Entonces

$$\exp(tv) = \lambda_v(t) \qquad (t \in \mathbb{C}, v \in V).$$

Una vez identificado el espacio tangente en 0 de $V$ con si mismo, $(d\exp)_0 = \mathrm{id}_V$. Por el teorema de las funciones implícitas, se deduce que la aplicación exp es un difeomorfismo local en un entorno de $0 \in V$ hacia un entorno de $e \in T$.

**Lema 1.2.** *Si $T$ es conexo, entonces $\exp(V)$ genera el grupo $T$: para cada $x \in T$, existen $v_1, \ldots, v_n$ en $V$ tales que $x = \exp(v_1) \ldots \exp(v_n)$.*

[2]En efecto, la conexidad sale de la definición de $A$ y el hecho que $A$ sea proyectiva pone sobre $A(\mathbb{C})$ una estructura de subvariedad compleja de $\mathbb{P}^n(\mathbb{C})$ compacta pues cerrada en $\mathbb{P}^n(\mathbb{C})$.

[3]Se puede pensar en $\lambda_v$ como en la geodésica sobre $T$ que parte de $e$ y tiene dirección $v$.

*Demostración.* Como exp es un difeomorfismo local en 0, $\exp(V)$ contiene un entorno abierto $U$ de $e$ en $T$, y sus traslaciones $x.U$ son entornos abiertos de cada $x \in \langle \mathrm{Im}(\exp)\rangle$. Entonces $\langle \exp(V)\rangle$ es un abierto de $T$. Como también es cerrado[4], la conexidad de $T$ implica $\langle \exp(V)\rangle = T$. □

Por la unicidad de $\lambda_v = (t \mapsto \exp_T(tv))$ se puede deducir también la siguiente propiedad: sea $F : T_1 \to T_2$ un morfismo de grupos de Lie complejos, entonces

$$F \circ \exp_{T_1} = \exp_{T_2} \circ (dF)_e, \tag{1.1}$$

es decir el siguiente diagrama conmuta:

$$\begin{array}{ccc} V_1 & \xrightarrow{(dF)_e} & V_2 \\ {\scriptstyle \exp_{T_1}}\downarrow & & \downarrow{\scriptstyle \exp_{T_2}} \\ T_1 & \xrightarrow{F} & T_2 \end{array}$$

*1.1.2. Consecuencias para un grupo de Lie complejo conexo compacto.*

**Proposición 1.3.** *Sea $T$ un grupo de Lie complejo conexo compacto y $V = \mathrm{Tan}_e(T)$. Entonces*

1. *la ley de grupo sobre $T$ es conmutativa;*
2. *$\exp = \exp_T : V \to T$ es un morfismo de grupos de Lie;*
3. *el morfismo* exp *es sobreyectivo;*[5]
4. *el núcleo de* exp *es un retículo del $\mathbb{C}-$espacio vectorial $V$ y $T$ es un toro complejo.*

Recordamos que un *retículo* de un $\mathbb{C}$-espacio vectorial $V$ de dimensión finita $g$ es un subgrupo de la forma $\mathbb{Z}e_1 \oplus \cdots \oplus \mathbb{Z}e_{2g}$ donde $e_1, \ldots, e_{2g}$ son vectores $\mathbb{R}$-linealmente independientes en $V$. Un subgrupo $\Lambda$ de $V$ es un retículo de $V$ si y sólo si $\Lambda$ es discreto y $T = V/\Lambda$ es compacto con la topología cociente[6]. Se puede dotar a tal cociente con una estructura de variedad compleja definiendo el haz de las funciones holomorfas: una función $f : U \to \mathbb{C}$ en un abierto $U$ de $T$ es *holomorfa* si y sólo si la función $\Lambda$-periódica $f \circ \pi$ es holomorfa sobre $\pi^{-1}(U)$.

Observamos que toda función holomorfa $f$ sobre $T$ es constante, pues $f \circ \pi$ es holomorfa y acotada sobre $V$. Las operaciones de grupos naturales sobre $T$ son aplicaciones holomorfas. El grupo de Lie obtenido es llamado un *toro complejo.*

Denotamos por $\mathcal{M}(T)$ el cuerpo de las funciones meromorfas de $T$.

*Demostración.* 1. Por un elemento $x \in T$, consideramos $f_x : T \to T$ el morfismo de conjugación: $f_x(y) = xyx^{-1}$ y su diferencial $(df_x)_e : V \to V$ en el neutro $e \in T$. La aplicación $T \to \mathrm{End}(V); x \mapsto (df_x)_e$ es holomorfa sobre la variedad compleja conexa compacta $T$ y a valores en el espacio de dimensión finita $\mathrm{End}(V)$, entonces es constante. En consecuencia, tenemos para todo $x \in T$, $(df_x)_e = (df_e)_e = \mathrm{id}_V$.

Se deduce de (1.1) y de lo precedente que $f_x \circ \exp_T = \exp_T \circ (df_x)_e = \exp_T$, lo que muestra que la imagen de exp está en el centro de $T$. Por conexidad de $T$, se deduce que $\exp(V) \subset Z(T)$ genera $T$ como grupo (Lema 1.2), entonces $T$ es conmutativo.

[4]porque su complementario es la unión de sus traslados, que son abiertos.

[5]o suryectivo, o exhaustivo, como el lector prefiera.

[6]es decir $U \subset T$ es abierto si $\pi^{-1}(U)$ es abierto en $V$, para $\pi : V \to T$ la proyección canónica.

2. Es consecuencia de la unicidad de $\lambda_v$: Sean $x, y$ en $V$. Como $T$ es abeliano, la aplicación $t \mapsto \exp(tx).\exp(tv)$ es un morfismo de grupos de Lie. Además su diferencial en 0 manda $\left(\frac{\partial}{\partial t}\right)_0$ sobre $x+y$, entonces $\varphi = \lambda_{x+y}$, o sea $\exp(tx).\exp(ty) = \exp(t(x+y))$ para todo $t \in \mathbb{C}, x, y \in V$. Tomando $t = 1$, obtenemos que exp es un morfismo de grupos de Lie (exp es holomorfa por definición).
3. Por 2., la imagen de exp es un subgrupo de $T$ y genera $T$, entonces es igual a $T$.
4. Por el hecho que exp es un difeomorfismo local alrededor de 0, hay un entorno $U$ de 0 tal que $U \cap \ker(\exp) = \{0\}$ (exp es localmente inyectiva). Eso demuestra que $\ker(\exp)$ es discreto. Además, por lo que precede, exp induce una aplicación $\phi : V/\Lambda \to T$ que es un isomorfismo de grupos holomorfo. Su diferencial en 0 es biyectiva, entonces $\phi$ es un isomorfismo de grupos de Lie complejos. Como $T$ es compacto, también lo es $V/\Lambda$ y así el subgrupo discreto $\Lambda$ es un retículo de $V$. Deducimos que $T \cong V/\Lambda$ es un toro complejo. □

**Corolario 1.4.** *Sea $A$ una variedad abeliana sobre $\mathbb{C}$. Entonces $A$ es un grupo abeliano y $A(\mathbb{C})$ es un toro complejo.*

En lo sucesivo, denotaremos aditivamente la ley de grupo sobre un toro complejo y 0 su elemento neutro.

**1.2. Cuando un toro complejo es una variedad abeliana.** En Subsección 1.1, demostramos que una variedad abeliana es un toro complejo. Es natural preguntarse si todos los toros complejos son variedades abelianas, es decir si admiten una inmersión holomorfa en un espacio proyectivo.

**Ejemplo 1.5.** Consideramos un toro $\mathbb{C}/\Lambda$ de dimensión 1, o sea $\Lambda$ es un retículo de $\mathbb{C}$. Denotamos por $\wp_\Lambda$ la *función de Weierstrass* definida por

$$\wp_\Lambda(z) = \frac{1}{z^2} + \sum_{\lambda \in \Lambda \setminus \{0\}} \left( \frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2} \right) \quad (z \in \mathbb{C}).$$

Entonces $u_\Lambda : z \mod \Lambda \mapsto (1 : \wp_\Lambda(z) : \wp'_\Lambda(z))$ define una inmersión holomorfa de $\mathbb{C}/\Lambda$ en $\mathbb{P}^2(\mathbb{C})$. Además la imagen de $\mathbb{C}/\Lambda$ en $\mathbb{P}^2(\mathbb{C})$ tiene ecuación $y^2 = 4x^3 - g_2(z)x - g_3(z)$ donde $g_2, g_3$ han sido definidas por ejemplo en el curso de teoría de Galois en este volumen, es decir es una curva elíptica sobre $\mathbb{C}$. Recíprocamente, para toda curva elíptica $E$ sobre $\mathbb{C}$, existe un retículo $\Lambda$ de $\mathbb{C}$ tal que la aplicación $u_\Lambda$ de antes define un isomorfismo de grupos de Lie del toro $\mathbb{C}/\Lambda$ en $E(\mathbb{C})$. Es el famoso *teorema de uniformización*. Entonces, en dimensión 1 las nociones de toros, variedades abelianas y curvas elípticas coinciden.

Pero lamentablemente, no es verdad en dimensión $> 1$. Teorema 1.8 da condiciones necesarias y suficientes para que un toro complejo de dimensión $g > 1$ sea una variedad abeliana. Daremos las líneas principales de la demostración en las subsecciones 2.3.1 y 2.3.3.

Sea $V$ un $\mathbb{C}$-espacio vectorial de dimensión $g$. Recordamos que una *forma hermitiana sobre $V$* es una aplicación $H : V \times V \to \mathbb{C}$ que es $\mathbb{C}$-bilineal en la primera variable y tal que $H(z,w) = \overline{H(w,z)}$. Para una forma hermitiana $H : V \times V \to \mathbb{C}$, denotamos $E = \Im(H) : V \times V \to \mathbb{R}$ su parte imaginaria, la cual es una forma real bilineal alternada. Dejamos la prueba al lector del siguiente hecho:

**Lema 1.6.** *La aplicación $H \mapsto E = \Im(H)$ define una correspondencia biyectiva desde el conjunto de las formas hermitianas en el conjunto de las formas reales bilineales alternadas $E$ verificando además $E(ix, iy) = E(x, y)$. La biyección inversa manda $E$ sobre $H$ definida por $H(x, y) = E(ix, y) + iE(x, y)$.*

**Definición 1.7** (Forma de Riemann)**.** Diremos que una forma hermitiana $H : V \times V \to \mathbb{C}$ es una *forma de Riemann con respecto a un retículo $\Lambda$ de $V$* si $E(\Lambda \times \Lambda) \subset \mathbb{Z}$, donde $E = \Im(H)$.

**Teorema 1.8.** *Un toro complejo $V/\Lambda$ es una variedad abeliana si y sólo si existe una forma de Riemann con respecto a $\Lambda$ que sea no degenerada.*

**Ejemplo 1.9** (Curvas elípticas)**.** Vimos en Ejemplo 1.5 que los toros de dimensión 1 son todos curvas elípticas (entonces variedades abelianas). Eso es confirmado por el teorema precedente. En efecto, sea $\Lambda = Z\lambda_1 + \mathbb{Z}\lambda_2$ con $\Im(\lambda_1/\lambda_2) > 0$ y consideramos la forma hermitiana sobre $\mathbb{C} \times \mathbb{C}$ definida por

$$H(z, w) = \frac{z\bar{w}}{\Im(\lambda_1\bar{\lambda_2})} \quad ((z, w) \in \mathbb{C} \times \mathbb{C})).$$

Se puede verificar que $H$ es una forma de Riemann no degenerada (ejercicio).

**Ejemplo 1.10.** Consideramos el toro $A_\tau = \mathbb{C}^g/(\mathbb{Z}^g + \tau\mathbb{Z}^g)$ donde $\tau \in M_g(\mathbb{C})$ es una matriz simétrica tal que $\Im(\tau)$ es definida positiva. Entonces

$$H(z, w) = {}^t z\Im(\tau)^{-1}\bar{w} \quad ((z, w) \in \mathbb{C}^g \times \mathbb{C}^g)$$

define una forma de Riemann no degenerada sobre $T_\tau$ (ejercicio). Entonces $A_\tau$ es una variedad abeliana.

**Ejemplo 1.11** (Variedades abelianas con multiplicación compleja)**.** Sea $K/\mathbb{Q}$ una *extensión CM*, es decir una extensión cuadrática totalmente imaginaria de un cuerpo de números totalmente real que denotaremos $K^+$. Denotamos $[K^+ : \mathbb{Q}] = g$ el grado de $K^+$ (de tal manera que $[K : \mathbb{Q}] = 2g$).

Decimos que un conjunto $\Phi$ de inmersiones $\varphi_k : K \hookrightarrow \mathbb{C}$ es un *tipo CM de $K$* si $\mathrm{Hom}(K, \mathbb{C})$ es la unión disjunta de $\Phi$ y $\bar{\Phi}$ donde para $\Phi = \{\varphi_1, \ldots, \varphi_g\}$, denotamos $\bar{\Phi} = \{\bar{\varphi}_1, \ldots, \bar{\varphi}_g\}$ con $\bar{\varphi}_i$ dada por la composición de $\varphi_i$ con la conjugación compleja. Un tipo CM de $K$ induce un isomorfismo $f : K \otimes_{\mathbb{Q}} \mathbb{R} \to \mathbb{C}^g$ por $x \otimes 1 \mapsto (\varphi_1(x), \ldots, \varphi_g(x))$.

Para un orden $\mathcal{O}$ de $K$, $f(\mathcal{O})$ es un retículo de $\mathbb{C}^g$. Vamos a definir una forma de Riemann no degenerada sobre el toro complejo $\mathbb{C}^g/f(\mathcal{O})$. Se puede demostrar que $K = K^+(\xi)$ con $\xi \in \mathcal{O}_K$ tal que $-\xi^2$ es un elemento totalmente positivo en $K^+$ y para todo $k \in \{1, \ldots, g\}$, $\Im(\varphi_k(\xi)) > 0$. Definimos una forma a valores reales, $\mathbb{R}$-bilineal antisimétrica por

$$E(z, w) = \sum_{k=1}^{g} \varphi_k(\xi)(\bar{z}_k w_k - z_k \bar{w}_k), \qquad (z, w \in \mathbb{C}^g). \tag{1.2}$$

La forma $E(iz, w)$ es simétrica, definida positiva. Se puede además demostrar que para todos $x, y \in K$, tenemos

$$E(f(x), f(y)) = \mathrm{Tr}_{K/\mathbb{Q}}(\xi\tilde{x}y) \tag{1.3}$$

donde $x \mapsto \tilde{x}$ es el automorfismo no trivial de $K/K^+$. Entonces $E(f(\mathcal{O}) \times f(\mathcal{O})) \subset \mathbb{Z}$. La forma de Riemann asociada es no degenerada (ejercicio).

La variedad abeliana así obtenida es dicha *a multiplicación compleja por $\mathcal{O}$, de tipo CM $(K, \Phi)$*.

**1.3. Morfismos e isogenías.** Para que una aplicación entre dos toros $f : T_1 = V_1/\Lambda_1 \to T_2 = V_2/\Lambda_2$ sea un *morfismo*, queremos que respecte las estructuras de grupos de Lie, es decir que sea una aplicación holomorfa y un morfismo de grupos. De hecho, tenemos el lema:

**Lema 1.12.** *Sean $T_1 = V_1/\Lambda_1$ y $T_2 = V_2/\Lambda_2$ dos toros complejos y $f : T_1 \to T_2$ una aplicación holomorfa. Entonces $f$ es inducida por una aplicación $\mathbb{C}$-afín $\tilde{f} : V_1 \to V_2$ tal que $\tilde{f}(\Lambda_1) \subset \Lambda_2$. Si además $f(0) = 0$, entonces $f$ es un morfismo de grupos de Lie. Su imagen es un subtoro de $T_2$ y su núcleo es un subgrupo cerrado de $T_1$, de cual la componente conexa es un subtoro de índice finito. (En el caso general, $f$ es la composición de un morfismo por una traslación).*

*Demostración.* Ver [7, Lema A.5.1.1, p. 93] o [3, Teorema I.2.3, p. 7]. □

**Definición 1.13.** Decimos que un morfismo $\varphi : T_1 \to T_2$ es una *isogenía* si es sobreyectivo y de núcleo finito. El orden de $\ker\varphi$ es llamado *grado* de $\varphi$.

*Observación* 1.14. Si $\varphi : T_1 \to T_2$ es una isogenía, entonces $\dim(T_1) = \dim(T_2)$.

**Ejemplo 1.15** (Multiplicación por un entero)**.** Sea $T = V/\Lambda$ un toro complejo de dimensión $g$ y $n \in \mathbb{Z}, n \geqslant 0$. La multiplicación por $n$ denotada $[n] = n\mathrm{id}_T \in \mathrm{End}(T)$ es una isogenía de grado $n^{2g}$. En efecto, $\ker[n] = (1/n)\Lambda/\Lambda \cong \Lambda/n\Lambda \cong (\mathbb{Z}/n\mathbb{Z})^{2g}$. Así tenemos una inyección $\mathbb{Z} \hookrightarrow \mathrm{End}(T)$.

**Ejemplo 1.16** (Dimensión 1)**.** Consideramos un toro de dimensión 1: $E = \mathbb{C}/\Lambda$ con $\Lambda$ un retículo de $\mathbb{C}$. Por Lema 1.12, el anillo $\mathrm{End}(E)$ de los endomorfismos de $E$ es isomorfo al conjunto $R$ de los números complejos $\alpha \in \mathbb{C}$ tales que $\alpha\Lambda \subset \Lambda$. Todo retículo de $\mathbb{C}$ es homotético a un retículo de la forma $\Lambda_\tau = \mathbb{Z} + \tau\mathbb{Z}$, con $\tau \in \mathbb{C}, \Im(\tau) > 0$. Así, podemos suponer, sin pérdida de generalidad, que $\Lambda = \Lambda_\tau$ para un tal $\tau$. Si $\alpha\Lambda \subset \Lambda$, entonces, en particular, existen $a, b, c, d \in \mathbb{Z}$ tales que $\alpha = a + b\tau$ y $\alpha\tau = c + d\tau$. Eso implica

$$\alpha^2 - (d + a)\alpha + (da - cb) = 0 \tag{1.4}$$

entonces $\alpha$ es un entero cuadrático. Además si $\mathbb{Z} \neq R$, $b \neq 0$ y tenemos

$$b\tau^2 + (a - d)\tau - c = 0 \tag{1.5}$$

Desde (1.4) y (1.5), deducimos que $R$ es un orden del cuerpo cuadrático imaginario $\mathbb{Q}(\tau)$.

Así, para $E$ curva elíptica sobre $\mathbb{C}$, $\mathrm{End}(E) \cong \mathbb{Z}$ o $\mathrm{End}(E)$ es un orden de un cuerpo cuadrático imaginario.

En el caso de una variedad abeliana de dimensión $> 1$, es más complicado.

**Ejemplo 1.17** (Variedad abeliana CM)**.** Sea $K/\mathbb{Q}$ una extensión CM y $A = \mathbb{C}^g/f(\mathcal{O})$ una variedad abeliana CM por un orden $\mathcal{O}$ de $K$, de tipo CM $(K, \Phi)$ (ver Ejemplo 1.11). Entonces, hay una inmersión $\mathcal{O} \hookrightarrow \mathrm{End}(A); \alpha \mapsto f_\alpha$, donde $f_\alpha$ es inducida por la multiplicación por $f(\alpha)$ en $\mathbb{C}^g$: $f_\alpha : (z_1, \ldots, z_g) \in A \mapsto (\varphi_1(\alpha)z_1, \ldots, \varphi_g(\alpha)z_g) \in A$.

Más generalmente, el anillo $\mathrm{End}(A)$ de los endomorfismos de una variedad abeliana $A$ es un orden en la $\mathbb{Q}$-álgebra de dimensión finita $\mathrm{End}_0(A) = \mathrm{End}(A) \otimes \mathbb{Q}$. Observamos que un elemento $\varphi \in \mathrm{End}(A)$ es una isogenía si y sólo si $\varphi$ es invertible en $\mathrm{End}_0(A)$. Decimos que un toro es *simple* si no tiene ningún subtoro no trivial. Se puede demostrar:

**Proposición 1.18.** *Si $A$ es simple, entonces* $\mathrm{End}_0(A)$ *es un álgebra de división (un cuerpo que puede ser no conmutativo).*

Ver Sección 4 para una descripción más avanzada de $\mathrm{End}_0(A)$.

**1.4. Descomposición salvo isogenía.** Dejamos como ejercicio la prueba del lema siguiente, lo cual demuestra que la relación de isogenía es una relación de equivalencia. Así decimos que $T_1$ y $T_2$ son *isógenos* si existe una isogenía $T_1 \to T_2$.

**Lema 1.19.** *Sean $\varphi : T_1 \to T_2$ y $\psi : T_2 \to T_3$ dos isogenías.*

1. *$\psi \circ \varphi$ es una isogenía de grado* $\deg(\psi)\deg(\varphi)$*;*
2. *existe una isogenía $\hat{\varphi} : T_2 \to T_1$ tal que $\varphi \circ \hat{\varphi} = [d]_{T_2}$ y $\hat{\varphi} \circ \varphi = [d]_{T_1}$, donde denotamos $d = \deg(\varphi)$. La isogenía $\hat{\varphi}$ es llamada* isogenía dual de $\varphi$*.*

Teorema 1.20 y Corolario 1.21 que siguen, dan la descomposición de las variedades abelianas salvo isogenía. Aquellos requieren la existencia de una forma de Riemann no degenerada sobre el toro considerado.

**Teorema 1.20** (Teorema de reducibilidad de Poincaré)**.** *Sea $A$ una variedad abeliana y $B$ una subvariedad abeliana. Entonces existe una subvariedad abeliana $C$ de $A$ tal que $B + C = A$ y $B \cap C$ es finito, es decir tal que $B \times C \to A; (b, c) \mapsto b + c$ sea una isogenía.*

*Demostración.* Ver [7, Teorema A.5.1.7, p. 95]. Denotamos $A = V/\Lambda$ y $B = V_1/\Lambda_1$ donde $V_1 \subset V$ y $\Lambda_1 = \Lambda \cap V_1$. Sea $H$ una forma de Riemann no degenerada asociada a $A$. Es natural considerar el ortogonal de $B$ para esta forma: consideramos

$$V_2 := \{v \in V : H(v, v_1) = 0;\ \text{para todo } v_1 \in V_1\}$$

y $\Lambda_2 = \Lambda \cap V_2$. Se puede demostrar que:

$$V_2 = \{v \in V : E(v, v_1) = 0 \text{ para todo } v_1 \in V_1\}$$

y $\Lambda_2 = \{v \in \Lambda : E(v, v_1) = 0;\ \text{para todo } v_1 \in V_1\} = \{v \in \Lambda : E(v, v_1) = 0;\ \text{para todo } v_1 \in \Lambda_1\}$ es un submódulo de $\Lambda$ de rango igual a $rg(\Lambda) - rg(\Lambda_1)$ porque $E$ es no degenerada y $\Lambda_1$ es un retículo (ejercicio). Entonces $\Lambda_2$ es un retículo de $V_2$ (pues de rango igual a $2\dim_{\mathbb{C}}(V_2)$) y $C := V_2/\Lambda_2$ es una subvariedad abeliana de $A$ con forma de Riemann $H|_{V_2 \times V_2}$. Como $V_1 \oplus V_2 = V$, tenemos $B + C = A$ y $B \cap C$ finito. □

**Corolario 1.21.** *Toda variedad abeliana $A$ es isógena a un producto de la forma $A_1^{n_1} \times \cdots \times A_s^{n_s}$ donde $A_1, \ldots, A_s$ son variedades abelianas simples, dos a dos no isógenas. La $\mathbb{Q}$-álgebra* $\mathrm{End}_0(A)$ *es semisimple:* $\mathrm{End}_0(A) \cong M_{n_1}(\mathrm{End}_0(A_1)) \times \cdots \times M_{n_r}(\mathrm{End}_0(A_r))$.

*Demostración.* Se deduce de Teorema 1.20 y Proposición 1.18 por inducción. Ver [7, Corolario A.5.1.8, p. 96]. □

## 2. Divisores sobre un toro, funciones theta y formas de Riemann

Esta sección es dedicada a introducir el material necesario y las ideas de la demostración de Teorema 1.8. En toda esta sección, denotamos por $T = V/\Lambda$ un toro complejo con $V$ un $\mathbb{C}$-espacio vectorial de dimensión $g$ y $\Lambda$ un retículo de $V$. La elección de una base de $V$ nos permite suponer que $V = \mathbb{C}^g$.

Deseamos determinar bajo que condiciones existe una inmersión holomorfa de $T$ en un espacio proyectivo $\mathbb{P}^n(\mathbb{C})$, es decir una aplicación holomorfa $u : T \to \mathbb{P}^n(\mathbb{C})$

que induzca un isomorfismo de variedades complejas entre $T$ y $u(T)$. Por el teorema de las funciones implícitas, una aplicación holomorfa $u$ es una inmersión si y sólo si es inyectiva y si $du$ es inyectiva en todo punto.

Es natural de considerar una aplicación $u$ que provenga de $\tilde{u} = (u_0, \dots, u_n) : V \to \mathbb{C}^{n+1}\backslash\{0\}$ con $u_0, \dots, u_n$ holomorfas y sin cero común. Una tal aplicación induce $u : V/\Lambda \to \mathbb{P}^n(\mathbb{C})$, si para todo $(z, \lambda) \in V \times \Lambda, \tilde{u}(z + \lambda) \equiv \tilde{u}(z) \in \mathbb{P}^n(\mathbb{C})$ o sea si para todo $(z, \lambda)$, existe un escalar $g_\lambda(z)$ tal que para todo $k \in \{0, \dots, n\}$, $u_k(z + \lambda) = g_\lambda(z)u_k(z)$. Eso motiva la definición de las funciones theta.

Referencias principales: [3, 7, 14].

## 2.1. Funciones theta y formas de Riemann.

*2.1.1. Funciones theta.* Para todo $t \in \mathbb{C}$, denotamos $\mathbf{e}(t) = \exp(2i\pi t)$. En lo que sigue, denotamos $V = \mathbb{C}^g$ y $\Lambda$ un retículo de $V$.

**Definición 2.1** (Función theta)**.** Una *función theta relativa a* $\Lambda$ es una función meromorfa[7] $\theta : V \to \mathbb{C}$ que satisface una ecuación de la forma

$$\theta(z + \lambda) = \mathbf{e}(f_\lambda(z)).\theta(z) \qquad ((z, \lambda) \in V \times \Lambda) \tag{2.1}$$

donde $f_\lambda : V \to \mathbb{C}$ es una función afín en $z \in V$ para todo $\lambda \in \Lambda$. En otras palabras $f_\lambda(z) = L(z, \lambda) + J(\lambda)$, con $J : \Lambda \to \mathbb{C}$ y $L : V \times \Lambda \to \mathbb{C}$ es $\mathbb{C}$-lineal en $z \in V$ para todo $\lambda \in \Lambda$. El par $(L, J)$ es llamado el *tipo* de la función theta.

*Observación* 2.2. La ecuación (2.1) determina $L$ de manera única, pero $J$ es definido a menos de agregación de un entero.

**Ejemplo 2.3.** Toda función de la forma $z \mapsto \exp(F(z))$ donde $F$ es un polinomio de grado total $\leqslant 2$ es una función theta llamada *trivial*. Más precisamente, si $F(z) = \varphi(z, z) + R(z) + S$ con $\varphi$ una forma bilineal simétrica, $R$ una forma lineal y $S$ una constante, el tipo de $\exp(F)$ es $(\frac{1}{i\pi}\varphi(z, \lambda), \frac{1}{2i\pi}(\varphi(\lambda, \lambda) + R(\lambda))$ (ejercicio).

**Ejemplo 2.4.**

1. Sea $\Lambda$ un retículo de $\mathbb{C}$. La función $\sigma : \mathbb{C} \to \mathbb{C}$ definida por:

$$\sigma(z) := z \prod_{\lambda \in \Lambda\backslash\{0\}} \left(1 - \frac{z}{\lambda}\right) \exp\left(\frac{z}{\lambda} + \frac{1}{2}\left(\frac{z}{\lambda}\right)^2\right) \tag{2.2}$$

   es una función theta relativa a $\Lambda$ llamada *función sigma de Weierstrass* (ver Ejercicio 6.1).

2. Sean $a, b$ reales. La función definida por:

$$\theta(z) = \sum_{m \in \mathbb{Z}} \exp(i\pi\tau(m + a)^2 + 2i\pi(m + a)(z + b)) \tag{2.3}$$

   es una función theta relativa a $\Lambda_\tau = \mathbb{Z} + \tau\mathbb{Z}$.

   Más generalmente, consideramos $\tau \in M_g(\mathbb{C})$ simétrica tal que $\Im(\tau)$ sea definida positiva y $\Lambda_\tau = \mathbb{Z}^g + \tau\mathbb{Z}^g$ el retículo de $\mathbb{C}^g$ como en Ejemplo 1.10. La función definida por

$$\theta(z) = \sum_{m \in \mathbb{Z}^g} \exp(i\pi {}^t m\tau m + 2i\pi {}^t mz) \tag{2.4}$$

   es una función theta relativa a $\Lambda_\tau$ llamada *función theta de Riemann* (ver Ejercicio 6.2).

[7]Cuidado: en algunas referencias, las funciones theta son definidas como holomorfas. Aquí permitimos que sean meromorfas.

*2.1.2. Forma de Riemann asociada a una función theta.* Sea $\theta$ una función theta de tipo $(L, J)$ relativamente a $\Lambda$. La relación (2.1) implica que para todos $\lambda_1, \lambda_2 \in \Lambda, z \in V$, tenemos

$$L(z, \lambda_1 + \lambda_2) - L(z + \lambda_1, \lambda_2) - L(z, \lambda_1) \equiv J(\lambda_1 + \lambda_2) - J(\lambda_1) - J(\lambda_2) \quad (\text{mód } \mathbb{Z}),$$

de lo que deducimos

$$L(z, \lambda_1 + \lambda_2) = L(z, \lambda_1) + L(z, \lambda_2) \tag{2.5}$$

$$L(\lambda_1, \lambda_2) \equiv J(\lambda_1 + \lambda_2) - J(\lambda_1) - J(\lambda_2) \quad (\text{mód } \mathbb{Z}) \tag{2.6}$$

$$L(\lambda_1, \lambda_2) \equiv L(\lambda_2, \lambda_1) \quad (\text{mód } \mathbb{Z}) \tag{2.7}$$

De (2.5), podemos extender $L$ a $\mathbb{R}$-linealidad a derecha, en una forma $L : V \times V \to \mathbb{C}$. Así la nueva forma $L$ es $\mathbb{C}$-lineal a izquierda y $\mathbb{R}$-lineal a derecha. Entonces la forma $E$ definida por

$$E(z, w) = L(z, w) - L(w, z) \quad ((z, w) \in V \times V) \tag{2.8}$$

es a valores reales, $\mathbb{R}$-bilineal alternada y, por (2.7), tiene valores enteras sobre $\Lambda \times \Lambda$. Además, para todo $(z, w) \in V \times V$, $E(iz, iw) = E(z, w)$ (ejercicio). Entonces, $E$ define una forma de Riemann (Lema 1.6): $H_\theta(z, w) = E(iz, w) + iE(z, w)$.

Así, a una función theta $\theta$ le asociamos la forma de Riemann $H_\theta$. Observamos que $H_\theta$ depende sólo de $L$ donde $(L, J)$ es el tipo de $\theta$.

Decimos que dos funciones theta $\theta_1, \theta_2$ son *equivalentes* y lo denotamos $\theta_1 \sim \theta_2$, si $\theta_1/\theta_2$ es una función theta trivial.

*Observación* 2.5. Si dos funciones theta tienen mismo tipo $(L, J)$ entonces son equivalentes.

**Proposición 2.6.**

1. *Sean dos funciones theta $\theta_1, \theta_2$, entonces $H_{\theta_1\theta_2} = H_{\theta_1} + H_{\theta_2}$.*
2. *Si $\theta$ es una función theta trivial entonces $H_\theta = 0$. Si $\theta_1, \theta_2$ son dos funciones theta equivalentes, entonces $H_{\theta_1} = H_{\theta_2}$.*

*Demostración.*

1. Para $i = 1, 2$, denotamos por $(L_i, J_i)$ el tipo de $\theta_i$. Entonces, $\theta_1\theta_2$ es una función theta de tipo $(L_1 + L_2, J_1 + J_2)$. Se deduce 1.
2. Desde Ejemplo 2.3, deducimos que si $\theta$ es trivial entonces $L = \frac{1}{i\pi}\varphi(z, w)$ es simétrica (con las notaciones del ejemplo), lo que implica $E_\theta = 0$. Así $H_\theta = 0$. De eso y de 1., también se deduce que si $\theta_1 \sim \theta_2$ entonces $H_{\theta_1} = H_{\theta_2}$.

□

*2.1.3. Función theta normalizada.* En la clase de equivalencia de una función theta hay una función theta particular que llamamos *normalizada*.

**Lema 2.7.** *Sea $\theta$ una función theta y $H = H_\theta$ la forma de Riemann asociada. Entonces existe una función theta $\widetilde{\theta}$ equivalente a $\theta$ tal que*

$$\widetilde{\theta}(z + \lambda) = \mathbf{e}\left(\frac{1}{2i}H(z, \lambda) + \frac{1}{4i}H(\lambda, \lambda) + K(\lambda)\right)\widetilde{\theta}(z) \qquad ((z, \lambda) \in V \times \Lambda) \tag{2.9}$$

*donde $K : \Lambda \to \mathbb{R}$ y verifica*

$$K(\lambda + \mu) - K(\lambda) - K(\mu) \equiv \frac{1}{2}E(\lambda, \mu) \quad (\text{mód } \mathbb{Z}). \tag{2.10}$$

*Además, existe $c > 0$ tal que, para todo $z \in V$,*

$$|\widetilde{\theta}(z)| \leqslant c. \exp\left(\frac{\pi}{2}H(z, z)\right). \tag{2.11}$$

La función $\tilde{\theta}$ es llamada *función theta normalizada asociada a* $\theta$ (o a $H$). La función $\psi$ definida por $\psi(z) = \mathbf{e}(K(z))$ verifica

$$\psi(\lambda+\mu) = \psi(\lambda)\,\psi(\mu)\,\mathbf{e}\left(\frac{1}{2}E(\lambda,\mu)\right) \tag{2.12}$$

y es llamada *semi-carácter asociado a la forma de Riemann* $H$.

*Demostración.* Sea $(L, J)$ el tipo de $\theta$ y $H = H_\theta$. Toda función $\theta_1$ en la clase de equivalencia de $\theta$ tiene un tipo $(L_1, J_1)$ donde $L_1(z,w) = L(z,w) + \frac{1}{i\pi}\varphi(z,w)$ con $\varphi$ una forma bilineal simétrica (ver Ejemplo 2.3). Deducimos que $L_1(z,w) - L_1(w,z) = L(z,w) - L(w,z) = E(z,w)$ (lo que implica $H_{\theta_1} = H$ como en Proposición 2.6). Recíprocamente, si $L_1$ es tal que

$$L_1(z,w) - L_1(w,z) = E(z,w) \tag{2.13}$$

entonces $L_1 - L$ es una forma bilineal simétrica. Con $\varphi := i\pi(L_1 - L)$ y para toda forma lineal $R$, la función $\theta_1 = \theta_{L_1,R}$ definida por $\theta_1(z) = \exp(\varphi(z,z) + R(z))\theta(z)$ tiene tipo $(L_1, J_1)$ con $J_1(\lambda) \equiv J(\lambda) + \frac{1}{2i\pi}(\varphi(\lambda,\lambda) + R(\lambda))$ (mód $\mathbb{Z}$). Como $E = \Im(H)$, la forma bilineal $L_1(z,w) = \frac{1}{2i}H(z,w)$ verifica (2.13). La función theta asociada $\theta_R(z) = \theta_{\frac{1}{2i}H(z,w),R}$ es de tipo $\left(\frac{1}{2i}H(z,\lambda), \frac{1}{4i}H(\lambda,\lambda) + K_R(\lambda)\right)$ con

$$K_R(\lambda) \equiv J(\lambda) - \frac{1}{2}L(\lambda,\lambda) + \frac{1}{2i\pi}R(\lambda) \quad (\text{mód } \mathbb{Z}).$$

Desde (2.6) se deduce que para toda forma lineal $R$, $K_R$ verifica (2.10) (ejercicio).

Basta ahora elegir $R$ tal que $K_R = K_0 + \frac{1}{2i\pi}R$ sea a valores reales. Desde (2.6), podemos suponer que $\Im(K_0)$ es $\mathbb{Z}$-lineal y extenderlo $\mathbb{R}$-linealmente a $V$. Entonces $R(z) = 2\pi(\Im(K_0(z)) - i\Im(K_0(iz)))$ define una forma lineal tal que $K_R(\lambda) \in \mathbb{R}$ para todo $\lambda \in \Lambda$.

Para terminar, (2.11) se deduce del hecho que la función $|\widetilde{\theta}(z)|\exp(-\frac{\pi}{2}H(z,z))$ es $\Lambda$-periódica y continua, entonces cotada. □

**Corolario 2.8.** *Si $\theta$ es entera, entonces $H_\theta$ es positiva (es decir $H_\theta(z,z) \geqslant 0$ para todo $z \in V$). Además, para todo $z_0 \in V$, la función holomorfa $z \mapsto \theta(z_0 + z)$ es constante sobre el núcleo $N = \{z \in V; H_\theta(z,w) = 0, \forall w \in V\}$ de $H_\theta$.*

*Demostración.* Consideramos $\theta$ holomorfa y $\widetilde{\theta}$ la función theta normalizada asociada. Supongamos que existe $z_0$ tal que $H_\theta(z_0,z_0) < 0$. Entonces (2.11) implica que la función holomorfa $t \in \mathbb{C} \mapsto (\widetilde{\theta}(tz_0))$ tiende hacia 0 cuando $|t|$ tiende hacia el infinito. Entonces por el teorema de Liouville, para todo $t \in \mathbb{C}$, $\theta(tz_0) = 0$. Como por continuidad $H(z,z) < 0$ en un vecino $U$ de $z_0$, el argumento precedente aplicado a todo $z \in U$ implicaría que $\theta$ es idénticamente cero. Deducimos que $H$ es positiva por contradicción.

Si $z \in N$ entonces

$$|\widetilde{\theta}(z_0+z)| \leqslant c\exp\left(\frac{\pi}{2}H(z_0+z,z_0+z)\right) = c\exp\left(\frac{\pi}{2}H(z_0,z_0)\right).$$

Aplicando de nuevo el teorema de Liouville a la función holomorfa $z \mapsto \widetilde{\theta}(z_0+z) - \widetilde{\theta}(z_0)$ nos da el segundo resultado. □

En particular, obtenemos la recíproca de Proposición 2.6:

**Corolario 2.9.** *La forma de Riemann $H_\theta$ es cero si y sólo si $\theta$ es trivial.*

**2.2. Divisores.** Recordamos:

**Definición 2.10** (Divisores)**.** Sea $X$ una variedad compleja conexa.

1. Sea $(U_\alpha, f_\alpha)_\alpha$ una familia donde $(U_\alpha)_\alpha$ es un recubrimiento de $X$ y $f_\alpha$ son funciones meromorfas sobre $U_\alpha$ no idénticamente cero sobre ninguna componente conexa de $U_\alpha$. Decimos que una tal familia es *admisible* si para todos $\alpha, \beta$, sobre $U_\alpha \cap U_\beta$, $f_\alpha/f_\beta$ es holomorfa y no se anula. Dos tales familias admisibles son equivalentes si su unión todavía es admisible.
2. Un *divisor (de Cartier) sobre* $X$ es una clase de equivalencia de una familia admisible $(U_\alpha, f_\alpha)$.
3. Decimos que un divisor $D$ es *efectivo* si puede ser descrito por una familia $(U_\alpha, f_\alpha)$ con $f_\alpha$ holomorfa sobre $U_\alpha$ para todo $\alpha$.
4. Si $D$ es dado por $(U_\alpha, f_\alpha)$ entonces la familia $(U_\alpha, 1/f_\alpha)$ define un divisor que depende sólo de $D$ y es denotado $-D$. Si $D'$ es un divisor dado por $(U'_\alpha, f'_\alpha)$ entonces $(U_\alpha \cap U'_\beta, f_\alpha.f'_\beta)$ define un divisor que depende sólo de $D$ y $D'$, denotado por $D + D'$.
5. Un divisor es *principal* si está dado por $(X, f)$ con $f$ meromorfa sobre $X$. Decimos que dos divisores son *linealmente equivalentes* si $D - D'$ es principal. Lo denotamos $D \sim D'$.

El conjunto de los divisores sobre $X$ es un grupo abeliano que denotamos $\mathrm{Div}(X)$. Ver [3, 14].

Consideramos ahora $X = T = V/\Lambda$ un toro complejo, como antes. La proyección $\pi : V \to T$ define una aplicación $\pi^* : \mathrm{Div}(T) \to \mathrm{Div}(V)$, cuya imagen es constituida por los divisores $\Lambda$-*periódicos*, es decir los divisores $D'$ tales que $t_\lambda^* D' = D'$ para todo $\lambda \in \Lambda$, donde $t_\lambda$ es la traslación[8] por $\lambda$.

Observamos que si $\theta$ es una función theta relativamente a $\Lambda$, el divisor $(\theta) \in \mathrm{Div}(V)$ es $\Lambda$-periódico, entonces hay un divisor $D_\theta \in \mathrm{Div}(T)$ tal que $\pi^*(D_\theta) = (\theta)$. Si $(U_i)_{i\in I}$ es un recubrimiento de $V$ constituido de abiertos $\Lambda$-pequeños[9], entonces $D_\theta$ puede ser descrito[10] por la familia $(\pi(U_i), \theta \circ (\pi|_{U_i})^{-1}))_{i\in I}$.

**Teorema 2.11** (Poincaré)**.** *Para todo $D \in \mathrm{Div}(T)$, existe $\theta$ una función theta relativamente a $\Lambda$ tal que $D_\theta = D$. Además, si $D$ es efectivo, la función $\theta$ es holomorfa.*

*Demostración.* Ver [3, p. 43] o [17, Teorema 18, p. 35]. □

Dejamos como ejercicio la proposición siguiente:

**Proposición 2.12.** *El divisor $D_\theta$ es trivial si y sólo si $\theta$ es una función theta trivial. Entonces, la aplicación $\theta \mapsto D_\theta$ define un isomorfismo de grupos*

$$\{\textit{funciones theta}\}/\{\textit{funciones triviales}\} \cong \mathrm{Div}(T).$$

De Corolarios 2.8 y 2.9 y de Teorema 2.11, tenemos el siguiente resultado.

**Corolario 2.13.** *La aplicación que a un divisor $D \in \mathrm{Div}(T)$ asocia la forma de Riemann $H_D := H_\theta$ donde $\pi^*(D) = (\theta)$ está bien definida y es un morfismo de grupos de $\mathrm{Div}(T)$ en el grupo $\mathcal{R}(T)$ de las formas de Riemann sobre $T$. Si $D$ es efectivo, $\theta$ es entera entonces $H_D$ es positiva.*

[8]Si $D'$ es dado por $(U_\alpha, f_\alpha)$ entonces $t_\lambda^* D'$ es dado por $(U_\alpha + \lambda, f_\alpha(z - \lambda))$.

[9]Un abierto $U$ es $\Lambda$-*pequeño* si no encuentra ningún de sus traslados por $\Lambda$.

[10]Esta familia es admisible porque desde la ecuación de $\theta$, para todo $\lambda$, $\theta(z + \lambda)/\theta(z)$ no tiene cero ni polo sobre $U_i$ para todo $i$.

**2.3. Esbozo de prueba del criterio en Teorema 1.8.** Si $\theta_0, \dots, \theta_n$ son funciones theta holomorfas sobre $V = \mathbb{C}^g$ relativamente a $\Lambda$, de mismo tipo y sin cero común, pues que satisfacen a la misma ecuación (2.1), la aplicación $(\theta_0, \dots, \theta_n) : V \to \mathbb{C}^{n+1}\backslash\{0\}$ induce una aplicación holomorfa de $V/\Lambda$ en $\mathbb{P}^n(\mathbb{C})$ que denotaremos $(\theta_0 : \cdots : \theta_n)$.

**Lema 2.14.** *Si $u : T = V/\Lambda \longrightarrow \mathbb{P}^n(\mathbb{C})$ es una aplicación holomorfa, entonces existen $\theta_0, \dots, \theta_n$ funciones theta enteras normalizadas de mismo tipo tales que $u = (\theta_0 : \cdots : \theta_n)$.*

*Demostración.* Denotamos por $(x_0 : \cdots : x_n)$ las coordenadas en $\mathbb{P}^n(\mathbb{C})$. Salvo de una permutación de índices, podemos suponer que $u(T)$ no es contenida en el hiperplano $\mathcal{H}_0 = (x_0 = 0)$. Entonces el pullback por $u$ de $H_0$ define un divisor efectivo[11] $D$ sobre $T$.

Denotamos $\theta_0$ la función theta normalizada asociada a $D$ por Proposición 2.11. Como $D$ es efectivo, $\theta_0$ es una función entera. Denotamos $\tilde{u} : V \to \mathbb{C}^{n+1}\backslash\{0\}$ induciendo $u$ y consideramos las funciones theta $\theta_j := \left(\frac{x_j \circ \tilde{u}}{x_0 \circ \tilde{u}}\right)\theta_0$ para $j = 1, \dots, n$. Pues como $\frac{x_j \circ \tilde{u}}{x_0 \circ \tilde{u}}$ es $\Lambda$-periódica, esas funciones theta son de mismo tipo que $\theta_0$ (en particular son normalizadas de forma de Riemann $H_{\theta_0}$). Además, son enteras, no tienen cero común y $u(z) = (\theta_0(z) : \cdots : \theta_n(z))$. □

*2.3.1. Condición necesaria de Teorema 1.8.* Supongamos que un toro complejo $T$ es una variedad abeliana, i.e. que existe una inmersión holomorfa $u : T \to \mathbb{P}^n(\mathbb{C})$. Sean $\theta_0, \dots, \theta_n$ tales que $u = (\theta_0 : \cdots : \theta_n)$ como en Lema 2.14. La forma de Riemann $H$ asociada a estas funciones theta enteras equivalentes es positiva (Corolario 2.8). Como cada función $\theta_j$ es constante sobre los conjuntos $z_0 + N$ ($z_0 \in V$), el hecho que $u$ sea una inmersión fuerza al núcleo $N$ de $H$ a ser trivial (¡la inmersión tiene que separar los puntos!). En conclusión la forma de Riemann $H$ asociada a $D$ es no degenerada.

*2.3.2. Teorema de Riemann-Roch.* Para terminar la prueba de Teorema 1.8, queremos construir una inmersión holomorfa de un toro $T$ dotado de una forma de Riemann no degenerada en un espacio proyectivo. Por Lema 2.14, sabemos que tales inmersiones son dadas por funciones theta enteras de mismo tipo. En esta subsección, examinamos al espacio vectorial $Th(L, J)$ de las funciones theta *holomorfas* de tipo dado $(L, J)$.

Decimos que $(L, J)$ es un *tipo* si $L : V \times \Lambda \to \mathbb{C}$ es $\mathbb{C}$-lineal a izquierda, $J : \Lambda \to \mathbb{C}$ y verifican las propiedades (2.5),(2.6) y (2.7). A un tipo $(L, J)$, extendiendo $L$ a $V \times V$ por $\mathbb{R}$-linealidad a derecha, podemos asociar una forma $\mathbb{R}$-bilineal alternada $E$ por (2.8). La forma $H$ definida por $H(z, w) = E(iz, w) + iE(z, w)$ es una forma de Riemann (ver Subsección 2.1.2).

*Observación* 2.15. Recíprocamente si $H$ es una forma de Riemann y $E = \Im(H)$, considerando $L(z, \lambda) = \frac{1}{2i}H(z, \lambda)$ y $J(\lambda) = \frac{1}{4i}H(\lambda, \lambda) + \frac{1}{4}E(\lambda, \lambda)$, entonces $(L, J)$ es un tipo de forma de Riemann asociada $H$ (ejercicio).

Recordamos que para una forma $\mathbb{R}$-bilineal alternada $E$ a valores enteras sobre un $\mathbb{Z}$-módulo libre $\Lambda$ de rango $2g$ que es no degenerada, existe una base $(\omega_1, \dots, \omega_{2g})$

[11] El hiperplano $\mathcal{H}_0$ puede ser descrito por la familia $(U_i, x_0/x_i)_{0 \leqslant i \leqslant n}$ donde $U_i = \mathbb{P}^n(\mathbb{C})\backslash(x_i = 0)$ y $D = u^*\mathcal{H}_0$.

dicha *base simpléctica* (o *base de Frobenius*) de $\Lambda$ en la cual la matriz de $E$ es de la forma

$$\begin{pmatrix} 0 & \Delta \\ -\Delta & 0 \end{pmatrix}$$

donde $\Delta = Diag(d_1, \ldots, d_g)$ con $d_1, \ldots, d_g$ enteros $> 0$ tales que $d_1 \mid \cdots \mid d_g$. Deducimos que $\det(E) > 0$. El *pfaffiano de* $E$ es el entero: $\mathrm{Pf}(E) = \sqrt{\det(E)} = d_1 \ldots d_g$.

**Teorema 2.16** (Teorema de Riemann-Roch para las variedades abelianas)**.** *Sea $(L, J)$ un tipo. Supongamos que la forma de Riemann asociada $H$ es no degenerada, y denotamos $E = \Im(H)$. Entonces, el espacio vectorial de las funciones theta holomorfas de tipo $(L, J)$ tiene dimensión*

$$\dim_{\mathbb{C}}(Th(L, J)) = \mathrm{Pf}(E) > 0.$$

*Demostración.* Ver [17, Teorema 24, p. 45], [7, Teorema A.5.3.3, p. 104], [8, Teoremas 2.2 y 2.3, pp. 11–12]. □

*Observación* 2.17. Sea $D$ un divisor efectivo sobre $T$ y $\theta$ una función theta entera tal que $(\theta) = \pi^*(D)$. Denotamos $(L, J)$ el tipo de $\theta$. La aplicación $\vartheta \mapsto \vartheta/\theta$ define un isomorfismo entre $Th(L, J)$ y el espacio vectorial $\mathcal{L}(D) = \{f \in \mathcal{M}(T); D + (f) \geqslant 0\} \cup \{0\}$. Esto explica el nombre de Teorema 2.16.

*2.3.3. Final de la prueba de Teorema 1.8.* Supongamos que el toro $T$ es dotado de una forma de Riemann no degenerada $H$.

De Observación 2.15, hay un tipo $(L, J)$ de forma de Riemann $H$ y por Teorema 2.16, $\dim_{\mathbb{C}}(Th(L, J)) > 0$. Denotamos $(\theta_0, \ldots, \theta_n)$ una base de $Th(L, J)$ y $D = D_{\theta_0} (= D_{\theta_i}$ para todo $i)$. Obtenemos una aplicación holomorfa $\Phi_D = (\theta_0 : \cdots : \theta_n) : T \to \mathbb{P}^n(\mathbb{C})$.

**Definición 2.18.** Decimos que un divisor $D$ es *muy amplio* si $\Phi_D$ es una inmersión holomorfa. Decimos que $D$ es *amplio* si un múltiplo positivo de $D$ es muy amplio.

El fin de la prueba se deduce del siguiente resultado:

**Teorema 2.19** (Lefschetz)**.** *Sea $D$ un divisor sobre $T$ con forma de Riemann asociada $H_D$ no degenerada. Entonces $3D$ es muy amplio, es decir, $3D$ define una inmersión holomorfa $\Phi_{3D} : T \to \mathbb{P}^n(\mathbb{C})$.*

*Demostración.* Ver [7, Teorema A.5.3.6, p. 105]. □

Observamos que, con Sección 2.3.1, el Teorema de Lefschetz demuestra:

**Corolario 2.20.** *Un divisor $D$ sobre $T$ es amplio si y sólo si $H_D$ es una forma de Riemann no degenerada.*

## 3. Teorema de Appell-Humbert y variedad abeliana dual

**3.1. Teorema de Appell-Humbert.** Sea $A = V/\Lambda$ una variedad abeliana.

Recordamos que $\mathcal{R}(A)$ es el grupo de las formas de Riemann sobre $A$. Para $H \in \mathcal{R}(A)$ decimos que una función $\psi : V \to U(1)$ es un semi-carácter asociado a $H$ si verifica (2.12) y denotamos

$$\mathcal{P}(A) = \{(H, \psi); H \in \mathcal{R}(A), \psi \text{ semi-carácter asociado a } H\}.$$

El conjunto $\mathcal{P}(A)$ es un grupo por la ley $(H_1, \psi_1).(H_2, \psi_2) = (H_1 + H_2, \psi_1\psi_2)$. Por Lema 2.7 y Teorema 2.11, a un divisor $D \in \mathrm{Div}(A)$ podemos asociarle una

función theta y entonces una función theta normalizada y un par $(H, \psi) \in \mathcal{P}(A)$. Consideramos la aplicación $\Psi : \mathrm{Div}(A) \to \mathcal{P}(A); D \mapsto (H_D, \psi_D)$ así definida.

Denotamos por $\mathrm{Pic}(A)$ el cociente de $\mathrm{Div}(A)$ por el subgrupo $\mathrm{Princ}(A)$ de los divisores principales, y $\mathrm{Pic}^0(A)$ el cociente por $\mathrm{Princ}(A)$ del subgrupo de los divisores $D$ tales que $H_D = 0$. El *grupo de Néron-Severi* es el cociente $NS(A) = \mathrm{Pic}(A)/\mathrm{Pic}^0(A)$. Observamos que los semi-caracteres para la forma de Riemann cero son exactamente los elementos del dual de Pontryagin $\mathrm{Hom}(\Lambda, U(1))$ de $\Lambda$, donde $U(1) = \{z \in \mathbb{C}, |z| = 1\}$.

**Teorema 3.1** (Appell-Humbert)**.** *La aplicación* $\Psi : \mathrm{Div}(A) \to \mathcal{P}(A); (D \mapsto (H_D, \psi_D))$ *es un morfismo de grupos que induce un isomorfismo* $\mathrm{Pic}(A) \to \mathcal{P}(A)$*, por lo cual* $\mathrm{Pic}^0(A)$ *se identifica a* $\mathrm{Hom}(A, U(1))$ *y que induce* $NS(A) \cong \mathcal{R}(A)$*.*

*Demostración.* Dejamos al lector la verificación que $\Psi$ es un morfismo de grupos. Sea $D \in \mathrm{Div}(A)$ tal que $(H_D, \psi_D)$ sea trivial. Entonces la función normalizada $\widetilde{\theta}$ asociada a $D$ es $\Lambda$-periódica, lo que implica que $D$ es principal. Y también vale la recíproca. Entonces $\Psi$ induce un morfismo inyectivo $\mathrm{Pic}(A) \to \mathcal{P}(A)$ que denotaremos todavía $\Psi$.

Demostramos que $\Psi$ es sobreyectivo. Sea $(H, \psi) \in \mathcal{P}(A)$. Podemos escribir $H$ como diferencia de dos formas de Riemann definidas positivas: $H = H_1 - H_2$ y para cada $H_i$, por Teorema 2.16, existe una función theta holomorfa normalizada $\theta_i$ de forma de Riemann $H_i$. La función theta meromorfa $\theta = \theta_1/\theta_2$ es normalizada y tiene forma de Riemann $H$. Denotamos por $\alpha$ su semi-carácter. Entonces $\psi/\alpha \in \mathrm{Hom}(\Lambda, U(1))$. Consideramos la función $\vartheta$ definida por $\vartheta(z) = \frac{\psi(z)}{\alpha(z)}\theta(z)$, extendiendo $\psi/\alpha$ en una función sobre $V$ por $\mathbb{R}$-linealidad. La función $\vartheta$ es una función theta normalizada de forma de Riemann $H$ y semi-carácter $\psi$. Así, $\Psi(D_\vartheta) = (H, \psi)$.

Deducimos que $\Psi : \mathrm{Pic}(A) \to \mathcal{P}(A)$ es un isomorfismo de grupos. Además, $H_D = 0$ si y sólo los semi-caracteres asociados a $H_D$ son los elementos de $\mathrm{Hom}(\Lambda, U(1))$. Entonces $\Psi(\mathrm{Pic}^0(A))$ se identifica a $\mathrm{Hom}(\Lambda, U(1))$. Deducimos el resultado. □

**3.2. Variedad abeliana dual.** Denotamos por $\bar{V}^*$ el conjunto de las formas $\mathbb{C}$-antilineales sobre $V$ (es decir las formas $\ell : V \to \mathbb{C}$ tales que $\ell(\alpha z) = \bar{\alpha}\ell(z)$ para todo $\alpha \in \mathbb{C}, z \in V$). La aplicación $\ell \mapsto \Im(\ell)$ define un isomorfismo entre el $\mathbb{R}$-espacio vectorial definido por $\bar{V}^*$ y $\mathrm{Hom}_{\mathbb{R}}(V, \mathbb{R})$ (ejercicio). Entonces la forma $\mathbb{R}$-bilineal $\langle , \rangle : \bar{V}^* \times V \to \mathbb{R}$ definida por $\langle \ell, v \rangle := \Im(\ell(v))$, es no degenerada. Eso implica que $\widehat{\Lambda} := \{\ell \in \bar{V}^*; \langle \ell, \Lambda \rangle \subset \mathbb{Z}\}$ es un retículo de $\bar{V}^*$ (ejercicio).

**Definición 3.2.** Llamamos a $\widehat{\Lambda}$ el *retículo dual de* $\Lambda$ y al toro de dimensión $g$

$$\widehat{A} := \bar{V}^*/\widehat{\Lambda}$$

el *toro dual de* $A$.

*Observación* 3.3. La aplicación $\bar{V}^* \to \mathrm{Hom}(\Lambda, U(1)); \ell \mapsto e(\langle \ell, . \rangle)$ es un morfismo sobreyectivo, de núcleo $\widehat{\Lambda}$, induciendo un isomorfismo $f : \bar{V}^*/\widehat{\Lambda} \cong \mathrm{Hom}(\Lambda, U(1))$.

Supongamos ahora que $A$ es una variedad abeliana y sea $D \in \mathrm{Div}(A)$. Consideramos la aplicación $\varphi_D : A \to \mathrm{Pic}(A); a \mapsto [t_a^* D - D]$, donde $t_a : x \mapsto x + a$ es la traslación por $a$ en $A$.

**Proposición 3.4.**

1. *la imagen de* $\varphi_D$ *está en* $\mathrm{Pic}^0(A)$ *;*

2. *la aplicación $\varphi_D$ depende sólo de la clase de $D$ en $NS(A) = \mathrm{Pic}(A)/\mathrm{Pic}^0(A)$;*
3. *si $D$ es amplio, entonces $\varphi_D : A \to \mathrm{Pic}^0(A) = \widehat{A}$ es una isogenía de grado* $\det(E) = \mathrm{Pf}(E)^2$.

*Demostración.* 1. Sea $a \in V$ y seguimos denotando a $a$ su imagen en $A = V/\Lambda$. Sea $\theta$ una función theta asociada a $D$, entonces $\pi^*(t_a^* D) = (\theta_a)$ donde $\theta_a(z) = \theta(z + a)$. Entonces $\pi^*(t_a^* D - D) = (\theta_a/\theta)$, es decir $t_a^* D - D = D_{\theta_a/\theta}$. Un cálculo muestra que la ecuación de la función theta normalizada equivalente a $\theta_a/\theta$ tiene multiplicador $e(\frac{1}{2i}H(a,\lambda)) = e(E(a,\lambda))$, donde $H = H_D$ y $E = \mathrm{Im}(H)$. Esto demuestra que la función de Riemann asociada a $\theta_a/\theta$ entonces a $t_a^* D - D$ es cero.

2. Si $D$ es tal que $H_D = 0$, entonces la función theta normalizada asociada a $\theta_a/\theta$ tiene multiplicador cero. Así, es $\Lambda$-periódica y deducimos que $t_a^* D - D$ es principal.
3. Observamos que con Teorema 3.1 podemos ver la aplicación $\varphi_D$ a través del diagrama

$$\begin{array}{ccc} A & \xrightarrow{\varphi_D} & \mathrm{Pic}^0(A) \\ & \searrow^{\psi} & \downarrow \cong \\ & & \mathrm{Hom}(\Lambda, U(1)) \end{array}$$

donde la aplicación vertical asocia a un divisor $D'$ el semi-carácter de la función theta normalizada asociada, entonces, como lo hemos visto en 1., el morfismo $\psi$ es dado por $\psi(a) = (\lambda \mapsto e(E(a,\lambda))$. En el caso donde $E$ es no degenerada, $\psi$ es sobreyectivo, entonces $\varphi_D$ es una isogenía. Su núcleo es $\{z \in V; E(z,\lambda) \in \mathbb{Z} \text{ para todo } \lambda \in \Lambda\}/\Lambda$. Considerando una base simpléctica de $\Lambda$ (ver Subsección 2.3.2), se puede demostrar que es un grupo finito de orden $\mathrm{Pf}(E)^2$.

□

**Corolario 3.5.** *Si $A$ es una variedad abeliana, entonces $\widehat{A}$ es también una variedad abeliana llamada* variedad abeliana dual de $A$.

*Demostración.* Sea $H$ una forma de Riemann no degenerada sobre $A$ y $D$ un divisor amplio asociado. Por Observación 3.3, Teorema 3.1 y la prueba de Proposición 3.4, 3., tenemos el diagrama conmutativo

$$\begin{array}{ccc} A = V/\Lambda & \xrightarrow{a \mapsto E(a,.)} & \bar{V}^*/\widehat{\Lambda} \\ \downarrow \varphi_D & & \downarrow \ell \mapsto e(\langle \ell, . \rangle) \\ \mathrm{Pic}^0(A) & \xrightarrow{\cong} & \mathrm{Hom}(A, U(1)) \end{array}$$

Como $E$ es no degenerada la aplicación $\varphi_H : a \mapsto E(a,.)$ es un isomorfismo de $V$ con $\bar{V}^*$ que manda $\Lambda$ sobre $\widehat{\Lambda}$. Consideramos la forma hermitiana $H^*$ sobre $\bar{V}^*$ definida por $H^*(z,w) := H(\varphi_H^{-1}(z), \varphi_H^{-1}(w))$. Desde Proposición 3.4,3., el núcleo de $\varphi_H$ es finito, entonces $\varphi_H^{-1}(\widehat{\Lambda})/\Lambda$ es finito. Deducimos que un múltiplo de $H^*$ es una forma de Riemann y es no degenerada porque $H$ lo es. □

**Proposición 3.6.** *1. Tenemos $\widehat{\widehat{A}} = A$.*

2. *Un morfismo de toros* $f : A_1 \to A_2$ *induce un morfismo* dual $\hat{f} : \hat{A}_2 \to \hat{A}_1$ *y* $\hat{\hat{f}} = f$.
3. *El functor* $\hat{\cdot}$ *de la categoría de los toros es exacto.*

*Demostración.* Dejamos la prueba en ejercicio (o ver [1, §2.4]). □

### 3.3. Polarización.

**Definición 3.7.** Sea $A$ una variedad abeliana. Una *polarización* sobre $A$ es el dato de la clase de un divisor amplio en $NS(A)$, o de manera equivalente, es el dato de una forma de Riemann $H$ no degenerada. Digamos que $(A, H)$ es una variedad abeliana *polarizada*. Una polarización $H$ es *principal* si $\mathrm{Pf}(\Im(H)) = 1$ y decimos en este caso que $(A, H)$ es *principalmente polarizada.*

Por Proposición 3.4, una polarización $[D]$ define una isogenía $\varphi_D : A \to \hat{A}$ de grado $\mathrm{Pf}(\Im(H_D))$. En particular, si la polarización es principal, la correspondiente isogenía es un isomorfismo.

Recíprocamente, una isogenía $\varphi : A \to \hat{A}$ es un $\varphi_D$ con $D$ un divisor amplio si y sólo si $\varphi$ es inducida por una aplicación analítica $\Phi : V \to \bar{V}^*$ tal que la forma $H_\Phi : V \times V \to \mathbb{C}$ definida por $H(z, w) = \Phi(z)(w)$ es una forma hermitiana definida positiva.

Un morfismo (resp. una isogenía) $f : (A, H) \to (A, H')$ de variedades abelianas polarizadas es un morfismo (resp. una isogenía) $f : A \to A'$ tal que $[f^* D'] = [D]$ (donde $H' = H_{D'}$ y $H = H_D$), es decir si $f : A = V/\Lambda \to A' = V'/\Lambda'$ proviene de una aplicación $F : V \to V'$ tal que $H'(F(z), F(w)) = H(z, w)$ para todos $z, w$ en $V$.

**Ejemplo 3.8** (Dimensión 1)**.** Toda curva elíptica es principalmente polarizada. En efecto, si $\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$ con $\Im(\omega_1/\omega_2) > 0$, la forma real alternada $E = \Im(H)$ asociada a la forma de Riemann $H(z, w) = \frac{z\bar{w}}{\Im(\omega_1 \bar{\omega}_2)}$ tiene matriz $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ en la base $(\omega_1, \omega_2)$ de $\Lambda$.

**Ejemplo 3.9.** Con las notaciones de Ejemplo 1.10, $A_\tau$ es principalmente polarizada.

En dimensión $> 1$, un importante ejemplo de variedad principalmente polarizada es dado por las jacobianas de curvas (ver Parte 2), pero no toda variedad abeliana es principalmente polarizada, aún si tenemos el resultado siguiente:

**Proposición 3.10.** *Toda variedad abeliana polarizada es isógena a una variedad abeliana principalmente polarizada.*

*Demostración.* Sea $(A, H)$ una variedad abeliana polarizada, $A = V/\Lambda$. Consideramos una base simpléctica $(\omega_1, \ldots, \omega_{2g}$ de $\Lambda$ con respecto a la forma alternada no degenerada $E = \Im(H)$. Con las notaciones de Subsección 2.3.2, consideramos el nuevo retículo de $V$ dado por $\Lambda' = \mathbb{Z}\frac{1}{d_1}\omega_1 + \mathbb{Z}\frac{1}{d_g}\omega_g + \mathbb{Z}\omega_{g+1} + \ldots \mathbb{Z}\omega_{2g}$. Entonces, $E$ toma todavía valores enteros sobre $\Lambda' \times \Lambda'$ y tiene matriz de determinante 1 en la base precedente de $\Lambda'$. Entonces el toro $V/\Lambda'$ es una variedad abeliana principalmente polarizada e isógena a $A$. □

## 4. Endomorfismos de las variedades abelianas

Sea $(A, H)$ una variedad abeliana polarizada. La isogenía $\varphi = \varphi_H$ define un elemento invertible en $\mathrm{End}_0(A)$.

**Definición 4.1** (Involución de Rosati)**.** Para $u \in \mathrm{End}(A)$ consideramos

$$u^\dagger := \varphi^{-1}\hat{u}\varphi \in End_0(A).$$

Se puede ver que

$$(u^\dagger)^\dagger = u, \quad (u+v)^\dagger = u^\dagger + v^\dagger \quad (u \circ v)^\dagger = v^\dagger \circ u^\dagger.$$

La anti-involución inducida sobre $\mathrm{End}_0(A)$ es llamada *involución de Rosati.*

Para $u \in \mathrm{End}_0(A)$ denotamos $\mathrm{Tr}(u)$ la traza del endomorfismo real de $V$ inducido por $u$.

**Teorema 4.2.** *La aplicación* $(u, v) \mapsto \mathrm{Tr}(u^\dagger \circ v)$ *es una forma bilineal simétrica definida positiva y racional sobre* $\mathrm{End}_0(A)$.

Así, si $(A, H)$ es una variedad polarizada simple, $B := \mathrm{End}_0(A)$ es un álgebra de división de rango finito sobre $\mathbb{Q}$, dotada de una anti-involución $\dagger$ tal que $\mathrm{Tr}(u^\dagger u) > 0$ para todo $u \neq 0$. Tales álgebras de división han sido clasificados por Albert en 1930.

Sea $K$ el centro de $B$ y $K_0$ el subcuerpo de los elementos fijos por $\dagger_K$. Como $B \otimes_K \bar{K}$ es un álgebra de matrices, la dimensión de $B$ sobre $K$ es un cuadrado. Ademas, pues $\dagger$ es una anti-involución, $[K : K_0] \leqslant 2$. Denotamos

$$[B : K] = d^2, \quad [K : \mathbb{Q}] = e \leqslant 2[K_0 : \mathbb{Q}].$$

**Teorema 4.3** (Clasificación de Albert)**.** *El cuerpo* $K_0$ *es un cuerpo de números algebraico totalmente real y el par* $(B, \dagger)$ *es de uno de los tipos siguientes:*

**Tipo I:** $B = K = K_0$ *(*$d = 1$*) y* $\dagger = \mathrm{id}$.

**Tipo II:** $K = K_0$ *y* $B$ *es un álgebra de cuaterniones indefinida*[12] *sobre* $K$ *(*$d = 2$*). La involución*$\dagger$ *es de la forma* $x^\dagger = ax^*a^{-1}$ *donde* $.^*$ *es la involución usual de* $B$ *y* $a \in B$ *un elemento tal que* $a^2 \in K$ *con* $a^2$ *totalmente negativo.*

**Tipo III:** $K = K_0$ *y* $B$ *es un álgebra de cuaterniones definida*[13] *sobre* $K$ *(*$d = 2$*). En este caso* $x^\dagger = x^*$ *es la involución usual sobre* $B$.

**Tipo IV:** $[K : K_0] = 2$ *y* $K$ *es un cuerpo CM*[14]*. En el caso donde* $K = B$*,* $A$ *es una variedad abeliana CM por* $K$.

*Además, para todos los tipos, tenemos la restricción de dimensión:* $ed^2 \mid 2g$*. En particular, para los tipos II y III tenemos* $2e \mid g$*. Para el tipo I, tenemos* $e \mid g$.

*Observación* 4.4. Con estas restricciones respectadas, para cada uno de este tipo, existe una variedad abeliana con el álgebra de endomorfismos correspondiente, a menos de dos excepciones para los tipos III y IV.

Para más detalles el lector podrá consultar [12, p. 186] o [1, §5.5 y cap. 9].

## 5. Espacios de móduli

**5.1. Matriz de periodos y condiciones de Riemann.** Consideramos $V$ un $\mathbb{C}$-espacio vectorial de dimensión $g$ y $\Lambda$ un retículo en $V$. Sea $\underline{e} = (e_1, \ldots, e_g)$ una $\mathbb{C}$-base de $V$ y $(\omega_1, \ldots, \omega_{2g})$ una $\mathbb{Z}$-base de $\Lambda$. La matriz $\Pi$ de los elementos $\omega_i$ en la base $\underline{e}$ es llamada *matriz de periodos.* Una vez fijadas la base $\underline{e}$ y $\underline{\omega}$, identificamos $V$ a $\mathbb{C}^g$ y $\Lambda$ a $\Pi\mathbb{Z}^{2g}$ y entonces $V/\Lambda \cong \mathbb{C}^{2g}/\Pi\mathbb{Z}^{2g}$.

[12] o sea $B \otimes_K \mathbb{R} \cong M_2(\mathbb{R})$ para toda inmersión $K \hookrightarrow \mathbb{R}$.

[13] o sea $B \otimes_K \mathbb{R}$ es el cuerpo de los cuaterniones para toda inmersión $K \hookrightarrow \mathbb{R}$.

[14] es decir, por definición, una extensión cuadrática totalmente imaginaria de un cuerpo de números totalmente real

**Teorema 5.1.** *El toro $T = V/\Lambda$ es una variedad abeliana si y sólo si existe una matriz $J \in M_2(\mathbb{Z})$ no degenerada alternada verificando las* condiciones de Riemann*:*

$$\Pi J^{-1}\,{}^t\Pi = 0 \tag{5.1}$$

$$i\Pi J^{-1}\,{}^t\overline{\Pi} > 0 \tag{5.2}$$

Presentamos la prueba del teorema precedente como un ejercicio (Ejercicio 6.6). Para una prueba completa (y entonces una solución al ejercicio), el lector puede referirse a [1, §4.2]. En Ejercicio 6.6, aparece que si $V/\Lambda$ es una variedad abeliana, entonces la matriz $J$ es la matriz en la base $\underline{\omega}$ de una forma alternada no degenerada con respecto a $\Lambda$ (es decir $J = (E(\omega_i, \omega_j))_{1 \leqslant i,j \leqslant 2g}$).

Por ejemplo si $A = V/\Lambda$ es principalmente polarizada y si elegimos para $\underline{\omega}$ una base simpléctica relativamente a la polarización, entonces $J = \left(\begin{smallmatrix} 0 & I_g \\ -I_g & 0 \end{smallmatrix}\right)$ y, con $\Pi = (\Pi_1\,\Pi_2)$ donde $\Pi_1, \Pi_2$ están en $M_g(\mathbb{C})$, las condiciones de Riemann se convierten en:

$$\Pi_2\,{}^t\Pi_1 = \Pi_1\,{}^t\Pi_2 \tag{5.3}$$

$$i\Pi_2\,{}^t\overline{\Pi_1} - i\Pi_1\,{}^t\overline{\Pi_2} > 0. \tag{5.4}$$

Se puede demostrar que $\Pi_2$ es invertible y un cambio de base manda la matriz de periodos $(\Pi_1\,\Pi_2)$ sobre $(\tau\ I_g)$ donde $\tau = \Pi_2^{-1}\Pi_1$ (ver Ejercicio 6.9, 1)). Las relaciones de Riemann dicen entonces que $\tau$ es simétrica de parte imaginaria $\Im(\tau)$ definida positiva. Este observación es el punto de inicio de la construcción del espacio de móduli para las variedades abelianas complejas como sigue en Subsección 5.2 y en Ejercicio 6.9.

**5.2. Espacios de móduli.** En esta subsección, nos interesamos en clasificar las variedades abelianas principalmente polarizadas salvo isomorfismo.

**Ejemplo 5.2** (Dimensión 1)**.** Consideramos el *semi-plano de Poincaré* $\mathcal{H} = \{z \in \mathbb{C}; \Im(z) > 0\}$. El grupo $\mathrm{SL}_2(\mathbb{R})$ actúa sobre $\mathcal{H}$ por homografías: $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right).z = \frac{az+b}{cz+d}$. El cociente $\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}$ tiene estructura de superficie de Riemann. La aplicación $\tau \mapsto \mathbb{C}/\Lambda_\tau$ define una correspondencia biyectiva entre los puntos de $\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}$ y las clases de isomorfismo de curvas elípticas sobre $\mathbb{C}$ (ver [15, ap. C, §13]).

Vamos a enunciar un teorema análogo en dimensión $g \geqslant 1$. Denotamos por $\mathcal{A}_g(\mathbb{C})$ al conjunto de las clases de isomorfismos de variedades abelianas complejas principalmente polarizadas. Observamos que $\mathcal{A}_1(\mathbb{C})$ es simplemente el conjunto de las clases de isomorfismo de curvas elípticas complejas, pues son todas principalmente polarizadas. Denotamos por $\mathcal{H}_g$ el *semi-espacio de Siegel*, es decir el conjunto de las matrices $\tau \in M_g(\mathbb{C})$ simétricas tales que $\Im(\tau)$ sea definida positiva (lo que denotamos $\Im(\tau) > 0$ por brevedad).

Denotamos por $\mathrm{Sp}_{2g}(K) = \{M \in \mathrm{GL}_{2g}(K); MJ^tM = J\}$ donde

$$J = \begin{pmatrix} 0 & I_g \\ -I_g & 0 \end{pmatrix}.$$

Sea $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{Sp}_{2g}(\mathbb{R})$. Entonces

$$\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right).\tau := (a\tau + b)(c\tau + d)^{-1} \tag{5.5}$$

define una acción (a la izquierda) de $\mathrm{Sp}_{2g}(\mathbb{R})$ sobre $\mathcal{H}_g$ (ver Ejercicio 6.8).

**Teorema 5.3.** *El cociente* $\mathrm{Sp}_{2g}(\mathbb{Z})\backslash\mathcal{H}_g$ *tiene una estructura de variedad analítica compleja. La aplicación* $\tau \mapsto A_\tau = \mathbb{C}^g/(\mathbb{Z}^g + \tau\mathbb{Z}^g)$ *define una biyección de* $\mathcal{H}_g$ *hasta* $\mathcal{A}_g(\mathbb{C})$.

Ejercicio 6.9 demuestra la biyección. Ver [1, cap. 8] o [3, §VII.1] para la prueba completa, en un contexto aún más general.

## 6. Ejercicios

*Ejercicio* 6.1. Sea $\Lambda$ un retículo de $\mathbb{C}$. Consideramos la función sigma de Weierstrass definida por

$$\sigma(z) := z \prod_{\lambda\in\Lambda\setminus\{0\}} \left(1 - \frac{z}{\lambda}\right) \exp\left(\frac{z}{\lambda} + \frac{1}{2}\left(\frac{z}{\lambda}\right)^2\right). \tag{6.1}$$

1. Demostrar que $\left(\frac{\sigma'}{\sigma}\right)' = -\wp_\Lambda$ donde la función $\wp$ de Weierstrass es definida en Ejemplo 1.5.
2. Verificar que $\wp_\Lambda$ es $\Lambda$-periódica.
3. Deducir que $\sigma$ es una función theta relativa a $\Lambda$.

*Ejercicio* 6.2. Sea $\tau \in M_g(\mathbb{C})$ simétrica tal que $\Im(\tau)$ sea definida positiva. Consideramos la función definida por

$$\theta(z) = \sum_{m\in\mathbb{Z}^g} \exp(i\pi\,{}^t m\tau m + 2i\pi\,{}^t mz) \tag{6.2}$$

1. Demostrar que para todo $m, \ell, k$ en $\mathbb{Z}^g$ y $z \in \mathbb{C}^g$, tenemos

$${}^t m\tau m + 2\,{}^t m(z + \ell + \tau k) = {}^t(m + k)\tau(m + k) + 2\,{}^t(m + k)z + 2\,{}^t m\ell - 2\,{}^t kz - {}^t k\tau k.$$

2. Deducir que para todo $z \in \mathbb{C}^g$ y todo $\ell, k$ en $\mathbb{Z}^g$,

$$\theta(z + \ell + \tau k) = \theta(z)\exp(-2i\pi\,{}^t kz - i\pi\,{}^t k\tau k)$$

y que $\theta$ es una función theta relativa a $\Lambda_\tau = \mathbb{Z}^g + \tau\mathbb{Z}^g$.

*Ejercicio* 6.3. Demostrar que toda función theta que no se anula es una función theta trivial.

*Ejercicio* 6.4. Consideramos el toro $A_\tau = \mathbb{C}^g/(\mathbb{Z}^g + \tau\mathbb{Z}^g)$ donde $\tau \in M_g(\mathbb{C})$ es una matriz simétrica tal que $\Im(\tau)$ es definida positiva (es decir un elemento de $\mathfrak{h}_g$). Demostrar que

$$H(z, w) = {}^t z\Im(\tau)^{-1}\bar{w} \qquad ((z, w) \in \mathbb{C}^g \times \mathbb{C}^g)$$

define una forma de Riemann no degenerada que induce una polarización principal sobre $A_\tau$. [Indicación: verificar que $E(m+\tau n, h+\tau\ell) = {}^t nh - {}^t m\ell$ donde $E = \Im(H)$.]

*Ejercicio* 6.5. Demostrar que una matriz $\Pi \in M_{g,2g}(\mathbb{C})$ es una matriz de periodos de un toro complejo si y sólo si la matriz por bloques $\left(\begin{array}{c}\Pi\\ \hline \overline{\Pi}\end{array}\right)$ es invertible. *(Para una solución, ver* [1, Proposition 1.1.2, p. 9]*).*

*Ejercicio* 6.6 (Condiciones de Riemann). Adoptamos las notaciones de Subsección 5.1: sea $\Pi$ una matriz de periodos de un toro complejo $T = V/\Lambda$ con respecto a una base $\underline{e}$ de $V$ y una base $\underline{\omega}$ de $\Lambda$.

1. Consideramos una forma alternada no degenerada $E : \Lambda \times \Lambda \to \mathbb{Z}$, la extendemos a $\mathbb{C}^g = \Lambda \otimes \mathbb{R}$ y consideramos $H : \mathbb{C}^g \times \mathbb{C}^g \to \mathbb{C}$ definida por $H(z,w) = E(iz,w) + iE(z,w)$. Denotamos por $J$ la matriz de $E$ con respecto a la base $(\omega_1, \ldots, \omega_{2g})$ de $\Lambda$ es decir $J = (E(\omega_i, \omega_j))_{1 \leqslant i,j \leqslant 2g}$.
   *a*) Verificar que de la definición de $J$, tenemos que para todos $x, y$ en $\mathbb{R}^{2g}$,
   $$E(\Pi x, \Pi y) = {}^t x J y.$$
   *b*) Consideramos la matriz por bloques
   $$L := \begin{pmatrix} \Pi \\ \overline{\Pi} \end{pmatrix}^{-1} \begin{pmatrix} iI_g & 0_g \\ 0_g & -iI_g \end{pmatrix} \begin{pmatrix} \Pi \\ \overline{\Pi} \end{pmatrix}.$$
   Verificar que $\Pi L = i\Pi$.
   *c*) Deducir que $E(iz, iw) = E(z,w)$ para todos $z, w$ en $\mathbb{C}^g$ si y sólo si ${}^t L J L = J$. (Indicación: como $\Lambda \otimes \mathbb{R} = V = \mathbb{C}^g$, se puede escribir $z = \Pi x$ y $w = \Pi y$ con $x, y$ en $\mathbb{R}^{2g}$).
   *d*) Concluir que $H$ es una forma hermitiana si y sólo si $\Pi J^{-1}\, {}^t\Pi = 0$.
   *e*) Demostrar que si $H$ es una forma hermitiana, entonces tiene matriz $2i(\overline{\Pi} J^{-1}\, {}^t\Pi)^{-1}$. Deducir que $H$ es una forma hermitiana definida positiva si y sólo si $i\Pi J^{-1}\, {}^t\overline{\Pi} > 0$.
2. Usar las preguntas precedentes para demostrar Teorema 5.1.

*Ejercicio* 6.7. Sean $A = V/\Lambda$ y $B = V'/\Lambda'$ dos variedades abelianas de dimensión respectiva $g$ y $g'$. Sea $\Pi \in M_{g,2g}(\mathbb{C})$ (resp. $\Pi' \in M_{g',2g'}(\mathbb{C})$) una matriz de periodos de $A$ (resp. $B$) con respecto a la elección de bases de $V$ y $\Lambda$ (resp. $V', \Lambda'$). Hacemos las identificaciones $A = \mathbb{C}^g/\Pi\mathbb{Z}^{2g}$ y $B = \mathbb{C}^{g'}/\Pi'\mathbb{Z}^{2g'}$. Supongamos que $f : A \to B$ es un morfismo de variedades abelianas. Denotamos por $F : \mathbb{C}^g \to \mathbb{C}^g$ el único isomorfismo tal que $F(\Lambda) \subset \Lambda'$ y $F$ induce $f$. Denotamos por $M \in M_{g,g'}(\mathbb{C})$ la matriz de $F$ y $R = M_{2g,2g'}(\mathbb{Z})$ la matriz de $F|_{\Lambda_\tau}$ con respecto a las bases precedentes. Demostrar que $M\Pi = \Pi' R$.

*Ejercicio* 6.8. Sea $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{Sp}_{2g}(\mathbb{R})$.

1. Demostrar que
   $$\overline{{}^t(c\tau + d)}(a\tau + b) - \overline{{}^t(a\tau + b)}(c\tau + d) = \tau - \overline{\tau} = 2i\Im(\tau).$$
2. Con lo que precede, demostrar que si $v$ es tal que $(c\tau + d)v = 0$ entonces ${}^t\overline{v}\Im(\tau)v = 0$.
3. Deducir que $(c\tau + d)$ es invertible.
4. Denotamos $\tau' := (a\tau + b)(c\tau + d)^{-1}$. Demostrar que
   $$\overline{{}^t(c\tau + d)}(\tau' - {}^t\tau')(c\tau + d) = 0 \tag{6.3}$$
   $$\overline{{}^t(c\tau + d)}(\tau' - {}^t\overline{\tau'})(c\tau + d) = 2i\Im(\tau). \tag{6.4}$$
5. Deducir que (5.5) define una acción de $\mathrm{Sp}_{2g}(\mathbb{R})$ sobre $\mathcal{H}_g$.

*Ejercicio* 6.9.

1. Sea $A = V/\Lambda$ una variedad principalmente polarizada y $\underline{\omega}$ una base simpléctica relativamente a la polarización. Denotamos por $\Pi = (\Pi_1 \Pi_2)$ con $\Pi_k \in M_g(\mathbb{C})$ $(k = 1, 2)$, la matriz de periodos de $\underline{\omega}$ en una base $\underline{e}$ de $V$ e identificamos $A$ con $\mathbb{C}^g/\Pi\mathbb{Z}^{2g}$ como en subsección 5.1. Recordamos que $\Pi_1, \Pi_2$ verifican las condiciones de Riemann (5.3) y (5.4).

*a*) Demostrar que $(\omega_1, \ldots, \omega_g)$ es una $\mathbb{C}$-base del $\mathbb{C}$-espacio vectorial $V$. (Indicación: considerar $W$ el $\mathbb{R}$ espacio vectorial generado por $\omega_1, \ldots, \omega_g$ y demostrar que $W \oplus iW = V$.) Deducir que $\Pi_1$ y $\Pi_2$ son invertibles.
*b*) Demostrar que $\tau := \Pi_2^{-1}\Pi_1 \in \mathcal{H}_g$.
*c*) Demostrar que la multiplicación por la matriz $\Pi_2^{-1} \in \mathrm{GL}_g(\mathbb{C})$ induce un isomorfismo de variedades abelianas polarizadas desde $\mathbb{C}^g/\Pi\mathbb{Z}^{2g} \cong A$ hasta $A_\tau = \mathbb{C}^g/\Lambda_\tau = \mathbb{C}^g/(\tau\, I_g)\mathbb{Z}^{2g}$.

2. Sea $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{Sp}_{2g}(\mathbb{Z})$, $\tau \in \mathcal{H}_g$ y $\tau' = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right).\tau = (a\tau + b)(c\tau + d)^{-1} \in \mathcal{H}_g$. Demostrar que

$$(\tau'\, I_g) = {}^t(c\tau + d)^{-1}(\tau\, I_g)\, {}^t\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Deducir que la multiplicación por ${}^t(c\tau + d)^{-1}$ induce un isomorfismo de variedades abelianas (principalmente) polarizadas desde $A_\tau = \mathbb{C}^g/\Lambda_\tau$ hasta $A_{\tau'} = \mathbb{C}^g/\Lambda_{\tau'}$.

3. Sean $\tau, \tau'$ en $\mathcal{H}_g$ tales que $A_\tau$ y $A_{\tau'}$ son isomorfas como variedades abelianas polarizadas. Denotamos por $F : \mathbb{C}^g \to \mathbb{C}^g$ el único isomorfismo tal que $F(\Lambda_{\tau'}) \subset \Lambda_\tau$ y $F$ induce $f : A_{\tau'} \xrightarrow{\cong} A_\tau$. Denotamos por $M \in \mathrm{GL}_g(\mathbb{C})$ la matriz de $F$ en la base canónica de $\mathbb{C}^g$ y $R \in M_{2g}(\mathbb{Z})$ la matriz de $F|_{\Lambda_\tau}$ en las bases simplécticas dadas por las matrices de periodos $(\tau\, I_g)$ y $(\tau'\, I_g)$ respectivamente. Recordamos que $M(\tau\, I_g) = (\tau'\, I_g)R$.

Demostrar que ${}^tR$ está en $\mathrm{Sp}_{2g}(\mathbb{Z})$ y que $\tau' = {}^tR.\tau$.

## Parte 2. Variedades abelianas: Geometría

Por varias razones queremos poder utilizar variedades abelianas sobre cualquier cuerpo $K$. Si la característica de $K$ es cero, podemos en parte utilizar el *principio de Lefschetz*. Si $A$ es definida sobre $K$ entonces es definida sobre un subcuerpo $K_0 \subset K$, que es de tipo finito sobre $\mathbb{Q}$, y se puede considerar una inyección $K_0 \hookrightarrow \mathbb{C}$ y considerar $A$ como una variedad abeliana compleja. Sin embargo este principio es inaplicable cuando la característica de $K$ es positiva, por ejemplo cuando $K$ es un cuerpo finito. Además cuando, por ejemplo, $K$ es un cuerpo de números, queremos guardar las propiedades aritméticas, es decir que $A(K)$ es un grupo (lo que no es obvio si se mira $A(K)$ como un subconjunto de $A(\mathbb{C})$) y, por ejemplo, considerar la acción del grupo de Galois $G_K := \mathrm{Gal}(\bar{K}/K)$ sobre $A(\bar{K})$. Veremos que se puede recuperar algebraicamente casi toda la geometría compleja como dualidad, formas de Riemann, con estructuras más ricas.

**Aviso.** *Esta parte requiere algún entendimiento del vocabulario básico de geometría algebraica: variedades, cuerpo de funciones de una variedad, morfismos, dimensión, puntos lisos y singulares, divisores (Weil, Cartier) y fibrados (de línea) tal como están presentados por ejemplo en los dos primeros capítulos de* [5] *o la parte A de* [7]. *En la segunda sección de esta parte damos una breve descripción sobre las nociones de divisores y fibrados.*

### 7. Grupos algebraicos

Repetimos en el contexto de la geometría algebraica la definición vista en el inicio del curso.

**Definición 7.1.** Un *grupo algebraico* sobre un cuerpo $K$ es una variedad algebraica $G$ junto con morfismos definidos sobre $K$, multiplicación $m_G : G \times_K G \to G$, inversión $\mathrm{inv}_G : G \to G$ y un elemento $e \in G(K)$ que satisfacen los axiomas de grupos usuales.

Una *variedad abeliana* definida sobre un cuerpo $K$ es un grupo algebraico sobre el cuerpo $K$ que, además, es una variedad proyectiva.

Observamos que la estructura de grupo algebraico produce aplicaciones naturales:

- traslaciones por un elemento $x \in G$ que denotamos $t_x : G \to G$ (cuando $G$ no es conmutativo, por supuesto, hay dos tipos: traslaciones a la derecha y a la izquierda); la aplicación $t_x$ es biyectiva con inverso $t_{\mathrm{inv}_G(x)}$.
- La "multiplicación por $[n]$" es definida inductivamente por $[0](x) = e_G$, $[1](x) = x$, $[-1](x) = \mathrm{inv}_G(x)$ y finalmente la relación de recurrencia $[n](x) = m_G(x, [n-1](x))$. Observamos que $[n]$ es un homomorfismo sólo cuando $G$ es conmutativo. Sin embargo en todos casos la diferencial $d[n]_{e_G} : \mathrm{Tan}_{e_G}(G) \to \mathrm{Tan}_{e_G}(G)$ es simplemente la multiplicación por $n$, así observamos que, cuando $n$ es coprimo con la característica del cuerpo $K$, la aplicación $[n]_G : G \to G$ define un morfismo finito separable y en particular sobreyectivo.

**Ejemplo 7.2.** Es fácil dar ejemplos de variedades afines con una ley de grupo.

1. (Grupo $\mathbb{G}_a$) La línea afín $G := \mathbb{A}^1$ con la adición $\mathbb{A}^1 \times \mathbb{A}^1 \to \mathbb{A}^1$, el elemento $0 \in \mathbb{A}^1(K)$ y la aplicación $\mathrm{inv}(x) = -x$ es un grupo algebraico afín.
2. (Grupo $\mathbb{G}_m$) La línea afín pinchada $G := \mathbb{A}^1 \backslash \{0\}$ con la multiplicación $G \times G \to G$ definida por $(x, y) \mapsto xy$, el elemento $1 \in G(K)$ y la aplicación $\mathrm{inv}(x) = x^{-1}$ es un grupo algebraico afín.
3. (Grupo $\mathrm{GL}_n$) La variedad afín de las matrices de tamaño $n \times n$ con determinante no nulo $G := \mathbb{A}^{n^2} \backslash \{\det = 0\}$ con la multiplicación de matrices $G \times G \to G$, el elemento identidad $I_n \in G(K)$ y la aplicación $\mathrm{inv}(M) = M^{-1}$ es un grupo algebraico afín. Otros ejemplos pueden ser dados como subgrupo de $\mathrm{GL}_n$: grupo especial $\mathrm{SL}_n$, grupo simpléctico, grupo ortogonal, etc. Debido a su importancia para las variedades abelianas, detallamos el ejemplo del grupo de las *similitudes simplécticas*. Denotamos $J = J_g = \begin{pmatrix} 0 & I_g \\ -I_g & 0 \end{pmatrix}$ la matriz antisimétrica de tamaño $2g \times 2g$ y definimos

$$\mathrm{GSp}_{2g} := \left\{ M \in \mathrm{GL}_{2g} \mid \exists \mu = \mu(M) \in \mathbb{G}_m, \text{ tal que } {}^t MJM = \mu J \right\}$$

Este grupo puede ser colocado en una sucesión exacta

$$0 \to \mathrm{Sp}_{2g} \to \mathrm{GSp}_{2g} \xrightarrow{\mu} \mathbb{G}_m \to 0,$$

donde $\mathrm{Sp}_{2g}$ es el subgrupo de isometrías simplécticas (que cumplen $\mu(M) = 1$).

**Ejemplo 7.3.** Es más difícil construir ejemplos de grupos algebraicos proyectivos. El primer ejemplo de grupo algebraico proyectivo es una curva elíptica, o sea, una curva de género 1 con un punto marcado. Veremos que una curva de género $g \geqslant 2$ corresponde a una variedad abeliana de dimensión $g$, su jacobiana.

1. (curvas elípticas [6, 15, 16]) Se puede representar como una cúbica plana; damos la ecuación cuando la característica del cuerpo $K$ es diferente de 2 y 3:

$$E = \{(x : y : z) \in \mathbb{P}^2 \mid zy^2 = x^3 + axz^2 + bz^3\}$$

con la condición para que la curva sea lisa $\Delta := 4a^3 - 27b^2 \neq 0$. El elemento neutro es el "punto en el infinito" $(0:1:0)$ el inverso es dado por $[-1](x:y:z) = (x:-y:z)$ y se puede describir la adición con la regla: $P+Q+R=0$ si y sólo si $P$, $Q$, $R$ estén alineados.

2. El producto de dos variedades abelianas es claramente una variedad abeliana. En particular, si $E_1,\ldots,E_g$ son curvas elípticas, el producto $E_1 \times \cdots \times E_g$ es una variedad abeliana de dimensión $g$.
3. (Jacobianas de dimensión 2 [7, 13]) Veremos que se puede asociar a cada curva de género $g$ una variedad abeliana (un grupo algebraico proyectivo) de dimensión $g$, llamada *jacobiana*. La descripción en el caso $g=2$ puede ser dada concretamente. Una curva de género 2 es siempre hiperelíptica, es decir existe un morfismo finito de grado dos $\pi : C \to \mathbb{P}^1$ con una involución canónica $\iota : C \to C$ tal que $\pi \circ \iota = \pi$ (si la curva es dada por una ecuación $y^2 = f(x)$ la involución canónica es simplemente $\iota(x,y) = (x,-y)$). Consideramos la superficie $X = C \times C/\mathcal{S}_2$ cociente de $C \times C$ por el grupo $\mathcal{S}_2$ generado por $\sigma(P_1,P_2) = (P_2,P_1)$. Los puntos de la superficie $X$ pueden identificarse con divisores efectivos de grado 2 sobre $C$; la superficie contiene la curva $L = \{[(P,\iota(P))] \mid P \in C\}$ que es isomorfa a $\mathbb{P}^1$ y se puede contraer[15] en un punto $\pi : X \to J$; más precisamente si $0 \in J$ es el punto tal que $\pi(L) = \{0\}$, la aplicación $\pi$ es un isomorfismo de $X \backslash L$ sobre $J \backslash \{0\}$ que manda $L$ sobre el punto 0. La variedad algebraica $J$ es una variedad abeliana, se puede definir una ley de grupo así. Escogemos 0 como elemento neutro y denotamos $D_0 = [(P,\iota(P))]$ un divisor que lo representa, para $D_1$, $D_2$ divisores efectivos de grado 2 existe un divisor efectivo $D_3$ tal que $D_1 + D_2 \sim D_3 + D_0$ y se define $[D_1] + [D_2] := [D_3]$; en general $D_3$ es único (salvo cuando $D_3 \sim D_0$). El inverso se obtiene con $\mathrm{inv}([D]) = [\iota(D)]$.

**Lema 7.4** (Lema de rigidez)**.** *Sea $X$ variedad proyectiva, $Y$, $Z$ variedades algebraicas y $f : X \times Y \to Z$ un morfismo. Si $f$ es constante sobre un trozo $X \times \{y_0\}$, entonces es constante sobre todo trozo $X \times \{y\}$. Si además $f$ es constante sobre un trozo $\{x_0\} \times Y$, entonces $f$ es constante.*

*Demostración.* Ver [7, Lema A.7.1.1, p. 119] o [12, p. 43]. □

Observamos que la proyectividad de $X$ es esencial para el lema. Por ejemplo la aplicación $f : \mathbb{A}^1 \times \mathbb{A}^1 \to \mathbb{A}^1$ definida por $f(x,y) = xy$ es constante sobre $\mathbb{A}^1 \times \{0\}$ y $\{0\} \times \mathbb{A}^1$ pero no es constante. La primera consecuencia es el siguiente teorema que corresponde, sobre $\mathbb{C}$, a la Proposición 1.3.

**Teorema 7.5.** *Una variedad abeliana es un grupo algebraico conmutativo. Más generalmente un morfismo $\phi : A \to B$ entre dos variedades abelianas que cumple que $\phi(e_A) = e_B$ es un homomorfismo.*

*Demostración.* Sea $\phi : A \to B$, introducimos $f(x,y) = \phi(xy)\,\mathrm{inv}_B(\phi(y))\,\mathrm{inv}_B(\phi(x))$. Observamos que $f(e_A,y) = \phi(e_A y)\,\mathrm{inv}_B(\phi(y))\,\mathrm{inv}_B(\phi(e_A)) = \phi(y)\,\mathrm{inv}_B(\phi(y)) = e_B$ e igualmente $f(x,e_A) = e_B$. El lema de rigidez implica que $f(x,y) = e_B$ y entonces que $\phi$ es un homomorfismo: $\phi(xy) = \phi(x)\phi(y)$. Aplicando esto a $\phi = \mathrm{inv}_A$, vemos que para todos $x,y$ tenemos $\mathrm{inv}_A(xy) = \mathrm{inv}_A(x)\,\mathrm{inv}_A(y)$, lo que es posible sólo si $A$ es conmutativo. □

[15]Para verificar este punto, invocamos el criterio de Castelnuovo ([5, Teorema 5.7, p. 414]) y verificamos que $L$ es una recta $\cong \mathbb{P}^1$ con auto-intersección $L \cdot L = -1$, ver ejercicio 1.

**Teorema 7.6** (Weil)**.** *Sean $X$ una variedad lisa y $A$ una variedad abeliana, sean $U$ un subconjunto abierto no vacío (denso) de $X$ y $\phi : U \to A$ un morfismo. Entonces se puede extender $\phi$ a un morfismo de $X$ hacia $A$.*

*Demostración.* Ver [7, Corolario A.7.1.4, p. 120]. □

La importancia de este resultado viene del hecho que una inclusión de cuerpos de funciones $i : K(A) \hookrightarrow K(X)$ induce automáticamente un morfismo $f : X \to A$ tal que $f^* = i$, y no sólo una aplicación racional.

## 8. Divisores de Weil y Cartier, fibrados de línea

*En esta sección describimos sucintamente y sin demostraciones las varias nociones de divisores y fibrados (de línea) en geometría algebraica.*

*8.0.1. Divisores de Weil.* Sea $X$ una variedad algebraica, un *divisor de Weil* $D$ es una combinación lineal de hipersuperficies con coeficientes enteros. En otros términos se puede escribir:

$$D = \sum_Z n_Z Z,$$

donde $Z$ recorre las subvariedades de codimensión 1 de $X$ y $n_Z \in \mathbb{Z}$, con la condición que los coeficientes $n_Z$ sean casi todos nulos. Se define la adición de dos divisores, sumando los coeficientes. El divisor $D$ es *efectivo* o *positivo* si para todos los coeficientes $n_Y \geqslant 0$. Se escribe $D_1 \geqslant D_2$ cuando $D_1 - D_2 \geqslant 0$. Los divisores de Weil de una variedad algebraica forman un grupo denotado $\mathrm{Div}(X)$.

Sea $f : X \to Y$ un morfismo dominante de variedades algebraicas, es decir que la imagen $f(X)$ no está contenida en ninguna hipersuperficie de $Y$. Para una hipersuperficie $Z$ en $Y$, la imagen recíproca $Z' = f^{-1}(Z) = \{x \in X \mid f(x) \in Z\}$ es una hipersuperficie de $X$. Se define entonces $f^*Z = dZ'$ donde $d$ es la multiplicidad. Observamos que la aplicación

$$f^* : \mathrm{Div}(Y) \to \mathrm{Div}(X),$$

que es un homomorfismo de grupos, es bien definida solo cuando $f$ es dominante.

Cuando $f \in K(X)^\times$ es una función racional (no nula), se puede definir el *divisor de la función* $f$ como la diferencia de sus ceros con sus polos, o sea:

$$\mathrm{d}iv(f) := \sum_Y \mathrm{ord}_Y(f) Y,$$

donde $\mathrm{ord}_Y(f)$ es el orden de anulación de $f$ según $Y$ (es positivo si $f$ se anula sobre $Y$ y es negativo si $f$ tiene un polo sobre $Y$). Tenemos

$$\mathrm{d}iv(fg) = \mathrm{d}iv(f) + \mathrm{d}iv(g)$$

y así los divisores $\mathrm{d}iv(f)$ forman un subgrupo $P(X)$ llamado el subgrupo de los divisores principales. Se nota $\mathrm{C}l(X)$ el cociente $\mathrm{Div}(X)/P(X)$.

**Ejemplo 8.1.** Escogemos $X = \mathbb{P}^n$, una hipersuperficie $Y \subset \mathbb{P}^n$ es definida por una ecuación $F = 0$, donde $F$ es homogéneo de grado $d$. Se obtiene un homomorfismo *grado* de $\mathrm{Div}(\mathbb{P}^n)$ hacia $\mathbb{Z}$ que manda $Y = Z(F)$ sobre $d = \deg(F)$. Si $Y'$ es definida por un polinomio $F'$ de grado $d$, observamos que la función racional $f := F'/F$ tiene como divisor $\mathrm{d}iv(f) = Y' - Y$. Concluimos que $\mathrm{C}l(\mathbb{P}^n) \cong \mathbb{Z}$.

*8.0.2. Divisores de Cartier.* Para definir un *divisor de Cartier* sobre $X$ escogemos un recubrimiento de abiertos $\{U_i\}_{i\in I}$ con funciones racionales $f_i \in K(U_i)^\times$ tales que $f_i f_j^{-1}$ no tenga polos ni ceros en $U_i \cap U_j$; o sea, tenemos que $f_i f_j^{-1} \in \mathcal{O}(U_i \cap U_j)^\times$. Identificamos dos datos $\{(U_i, f_i)\}_{i\in I}$ y $\{(V_j, g_j)\}_{j\in J}$ cuando $f_i g_j^{-1}$ es regular sin ceros sobre $U_i \cap V_j$; un *divisor de Cartier* es un tal dato $\{(U_i, f_i)\}_{i\in I}$ módulo esta equivalencia.

La adición de dos divisores de Cartier es definida por

$$\{(U_i, f_i)\}_{i\in I} + \{(V_j, g_j)\}_{j\in J} = \{(U_i \cap V_j, f_i g_j)\}_{(i,j)\in I\times J}$$

Un divisor de Cartier $D = \{(U_i, f_i)\}_{i\in I}$ es *efectivo* si las funciones $f_i$ son regulares sobre $U_i$; el *soporte* $supp(D)$ de $D$ es la unión de los ceros y polos de las funciones $f_i$. Un divisor de Cartier es *principal* si es de la forma $div(f) = (X, f)$. Los divisores principales forman un subgrupo de los divisores de Cartier. El grupo de clases de divisores de Cartier es denotado $\mathrm{C}aCl(X)$.

Cuando $f : X \to Y$ es un morfismo, definimos la imagen recíproca de $D = \{(U_i, f_i)\}_{i\in I}$ como

$$f^* D = \{(f^{-1}(U_i), f_i \circ f)\}_{i\in I},$$

en particular si $D = (Y, g)$, tenemos $f^*D = (X, g \circ f)$. Esta definición solo es correcta cuando $f(X)$ no está contenido en $supp(D)$ pero observamos que se puede fácilmente remplazar $D$ por $D' \sim D$ tal que $D$ cumpla la condición; de hecho basta remplazar $D = \{(U_i, f_i)\}_{i\in I}$ por $D' = \{(U_i, f_i f_j^{-1})\}_{i\in I}$ para mover el soporte fuera de $U_j$ . Así se puede definir

$$f^* : \mathrm{C}aCl(Y) \to \mathrm{C}aCl(X).$$

Resumimos esto de la manera siguiente, $\mathrm{C}aCl$ es un funtor contravariante, pues tenemos claramente que si $f_1 : X \to Y$ y $f_2 : Z \to X$ son morfismos, $(f_1 \circ f_2)^* = f_2^* \circ f_1^*$.

**Ejemplo 8.2.** Escogemos $X = \mathbb{P}^n$, si $D = \{(U_i, f_i)\}_{i\in I}$ es un divisor de Cartier efectivo, podemos suponer que cada $U_i$ es contenido en uno de los abiertos canónicos $V_j = \{x \in \mathbb{P}^n \mid x_j \neq 0\}$ y entonces podemos ver $f_i$ como un polinomio en $x_0/x_j, \ldots, x_n/x_j$. Se demuestra fácilmente que la condición de recubrimiento implica que los $f_i$ son las deshomogeneizaciones de un polinomio homogéneo $F$. Concluimos nuevamente que $\mathrm{C}aCl(\mathbb{P}^n) \cong \mathbb{Z}$. Si $\phi : \mathbb{P}^m \to \mathbb{P}^n$ es un morfismo dado por polinomios homogéneos $(F_0, \ldots, F_n)$ de grado $d$, la aplicación $\phi^* : \mathbb{Z} \cong \mathrm{C}aCl(\mathbb{P}^n) \to \mathrm{C}aCl(\mathbb{P}^m) \cong \mathbb{Z}$ es dada por $n \mapsto dn$ (Ejercicio).

*8.0.3. Fibrados de línea.* Como sólo usaremos "fibrados de línea", después de un tiempo hablaremos simplemente de "fibrados".

Un *fibrado* de línea sobre $X$ es una familia continua de líneas parametrizada por $X$. Más precisamente es un morfismo $p : E \to X$ tal que

- La fibra $E_x := p^{-1}\{x\}$ es un espacio vectorial de dimensión 1.
- La fibración es localmente trivial, es decir, que se puede recubrir $X$ con abiertos $U_i$ sobre los cuales $E_{U_i} = p^{-1}U_i$ es trivial, existe isomorfismos $\phi_i$ que transforman $p$ en la primera proyección $p_1 : U_1 \times \mathbb{A}^1 \to U_i$; en diagrama

$$\begin{array}{ccc} E_{U_i} & \xrightarrow{\phi_i} & U_i \times \mathbb{A}^1 \\ {\scriptstyle p}\downarrow & \swarrow_{p_1} & \\ U_i & & \end{array}$$

Sea $p : E \to X$ y $p' : E' \to X'$ dos fibrados de línea. Un homomorfismo de fibrados de línea es un morfismo $\phi : E \to E'$, combinado con un morfismo $\bar{\phi} : X \to X'$, tal que $\phi_x : E_x \to E'_{\phi(x)}$ sea lineal, y tal que $\phi$, $\bar{\phi}$ conmuten con los morfismos definiendo los fibrados, es decir que el siguiente diagrama sea conmutativo

$$\begin{array}{ccc} E & \xrightarrow{\phi} & E' \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle p'} \\ X & \xrightarrow{\bar{\phi}} & X' \end{array}$$

Generalmente se identifican dos fibrados isomorfos.

**Ejemplo 8.3.** La variedad $X \times \mathbb{A}^1$ con la primera proyección es un fibrado llamado *fibrado trivial*. Consideramos la aplicación cociente $\pi : \mathbb{A}^{n+1}\backslash\{0\} \to \mathbb{P}^n$ que define $\mathbb{P}^n$, podemos construir un fibrado sobre $\mathbb{P}^n$

$$E = \{(x, v) \in \mathbb{P}^n \times \mathbb{A}^{n+1} \mid v = 0 \text{ o } \pi(v) = x\} \to \mathbb{P}^n.$$

Este fibrado no es trivial, en la notación clásica de Serre es el fibrado $\mathcal{O}(-1)$.

Una *sección* de un fibrado $p : E \to X$ es un morfismo $s : X \to E$ tal que $p \circ s = id_X$. El conjunto de las secciones de un fibrado forman un espacio vectorial denotado $\Gamma(X, E)$ (o también $H^0(X, E)$). La propiedad más importante es que, para una variedad proyectiva $X$, el espacio $\Gamma(X, E)$ es de dimensión finita. Cuando $\Gamma(X, E) \neq \{0\}$ es de dimensión $n + 1$, se puede definir una aplicación racional $\Phi_E : X \cdots \to \mathbb{P}\Gamma(X, E) \cong \mathbb{P}^n$ por la fórmula $\Phi_E(x) = (s_0(x), \ldots, s_n(x))$, donde los $s_i$ forman una base de $\Gamma(X, E)$.

La *suma* de dos fibrados de línea $E$ y $E'$ sobre $X$ es definida como el morfismo $E'' \to X$ tal que las fibras sean el producto tensorial de las dos fibras $(E'')_x = E_x \otimes E'_x$. La imagen recíproca por un morfismo $\phi : Y \to X$ de un fibrado es el producto fibrado $f^*E = Y \times_X E$; al nivel de conjuntos, se puede describir como $\{(y, v) \in Y \times E \mid f(y) = p(v)\}$; las fibras de $p' : f^*E \to Y$ son tales que $(\phi^*E)_y = E_{\phi(y)}$. Se puede describir la construcción con el diagrama

$$\begin{array}{ccc} f^*E & \longrightarrow & E \\ {\scriptstyle p'}\downarrow & & \downarrow{\scriptstyle p} \\ Y & \xrightarrow{f} & X \end{array}$$

*Notación* 8.4. Para denotar la suma de dos fibrados $\mathcal{L}$ y $\mathcal{M}$ sobre $X$ utilizaremos *dos* notaciones: $\mathcal{L} \otimes \mathcal{M}$ o $\mathcal{L} + \mathcal{M}$. De la misma manera denotamos $\mathcal{L}^{\otimes n}$ o $\mathcal{L}^n$ o $n\mathcal{L}$ la suma de $\mathcal{L}$ con si mismo $n$ veces.

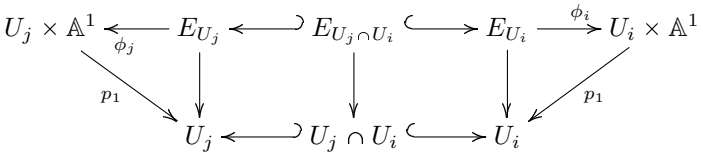
El *dual* de un fibrado $p : E \to X$ es el fibrado $\check{p} : \check{E} \to X$ tal que $\check{E}_x$ sea el dual (como espacio vectorial) de $E_x$.

Las clases de isomorfismo de fibrados de línea sobre $X$, con la suma (producto tensorial) forman un grupo llamado el *grupo de Picard* de $X$ y denotado $\mathrm{Pic}(X)$; el inverso de $E$ es su dual. La asociación $X \mapsto \mathrm{Pic}(X)$ es un funtor contravariante, a cada morfismo $f : Y \to X$, corresponde un homomorfismo de grupos $f^* : \mathrm{Pic}(X) \to \mathrm{Pic}(Y)$, y tenemos también $(f_1 \circ f_2)^* = f_2^* \circ f_1^*$.

Comparación de los grupos $\mathrm{C}l(X)$, $\mathrm{C}aCl(X)$ y $\mathrm{Pic}(X)$. Estos no son idénticos en general pero están estrechamente vinculados y, en muchos casos, isomorfos.

Describimos una correspondencia entre (clases de) divisores de Cartier y de Weil. A un divisor de Cartier $D = \{(U_i, f_i)\}_{i\in I}$ se asocia la familia de divisores de Weil principales $D_{U_i} = div(f_i)$ y se observa que la condición $f_i f_j^{-1} \in \mathcal{O}(U_i \cap U_j)^\times$ permite pegar estos divisores de Weil en un divisor de Weil global sobre $X$, que por tanto no es necesariamente principal. Se obtiene así una aplicación $\lambda : CaCl(X) \to Cl(X)$. Esta aplicación es siempre un homomorfismo pero no precisa ser ni inyectiva, ni sobreyectiva en general. Por ejemplo, cuando $X$ es singular, puede existir hipersuperficies $Y$ en $X$, pasando por un punto $x$, que no pueden ser definidas en una vecindad de $x$ por una ecuación.

Describimos la correspondencia entre fibrados y (clases de) divisores de Cartier. Se puede reconstruir un fibrado pegando sus pedazos $E_{U_i} \cong U_i \times \mathbb{A}^1$ a través de los diagramas




es decir consideramos $\phi_j \circ \phi_i^{-1} : (U_j \cap U_i) \times \mathbb{A}^1 \to (U_i \cap U_j) \times \mathbb{A}^1$ que debe escribirse $(x, v) \mapsto (x, f_{ij}(x)v)$ con una función $f_{ij}$ regular e invertible sobre $U_j \cap U_i$ (tales $f_{ij}$ se llaman *funciones de transición* del fibrado $E$). Vemos que un divisor de Cartier $D = \{(U_i, f_i)\}_{i\in I}$ define un fibrado, escogiendo como funciones de transición $f_{ij} := f_i f_j^{-1}$. Recíprocamente, si $s : X \to E$ es una sección, $s$ proporciona $U_j \xrightarrow{s} E_{U_i} \xrightarrow{\phi_i} U_i \times \mathbb{A}^1$ que tiene la forma $x \mapsto (x, f_i(x)$. Entonces una sección de $E$ define una familia $(U_i, f_i)$ y un divisor de Cartier. Se obtiene así una aplicación $\kappa : CaCl(X) \to \mathrm{Pic}(X)$.

**Teorema 8.5.** *Sea $X$ una variedad (irreducible), el homomorfismo $\kappa : CaCl(X) \to \mathrm{Pic}(X)$ es un isomorfismo. Sea $X$ una variedad (irreducible) lisa, el homomorfismo $\lambda : CaCl(X) \to Cl(X)$ es un isomorfismo.*

En el caso de una curva proyectiva lisa $C$, tenemos la aplicación *grado* que asocia a un divisor $D = \sum_{P\in C} n_P P$ su grado $\deg(D) = \sum_{P\in C} n_P$ y se obtiene una sucesión exacta

$$0 \longrightarrow \mathrm{Pic}^0(C) \longrightarrow \mathrm{Pic}(C) \xrightarrow{\deg} \mathbb{Z} \longrightarrow 0 \tag{8.1}$$

Cuando $C = \mathbb{P}^1$ tenemos $\mathrm{Pic}^0(C) = 0$ y $\mathrm{Pic}(C) = \mathbb{Z}$, pero cuando $C$ no es una curva racional, el grupo $\mathrm{Pic}^0(C)$ no es trivial y además tiene una rica estructura, precisamente de variedad abeliana de dimensión el género de $C$. Esta sucesión exacta se generaliza a variedades de dimensión mayor de la forma siguiente. Si $X$ es una variedad lisa proyectiva tenemos una sucesión exacta análoga

$$0 \longrightarrow \mathrm{Pic}^0(X) \longrightarrow \mathrm{Pic}(X) \longrightarrow \mathrm{NS}(X) \longrightarrow 0 \tag{8.2}$$

donde $\mathrm{Pic}^0(X)$ es el subgrupo de los fibrados (o clases de divisores) que se puede deformar algebraicamente en el fibrado trivial, y $\mathrm{NS}(X)$ es el grupo de Néron-Severi. Dos características permanecen: el grupo $\mathrm{NS}(X)$ es un grupo de tipo finito (pero no es necesariamente isomorfo a $\mathbb{Z}$) y el grupo $\mathrm{Pic}^0(X)$ es el grupo de puntos de una variedad abeliana.

Terminamos este preliminar geométrico con una generalización del teorema clásico de Bézout *dos curvas planas proyectivas de grado $d_1$ y $d_2$ se intersectan en $d_1 d_2$ puntos (contados con multiplicidades).*

**Teorema 8.6.** *Sea $X$ una variedad lisa proyectiva de dimensión $n$. Existe una única aplicación multilineal simétrica:*

$$\begin{array}{ccc} \mathrm{Pic}(X) \times \cdots \times \mathrm{Pic}(X) & \to & \mathbb{Z} \\ (\mathcal{L}_1, \ldots, \mathcal{L}_n) & \mapsto & (\mathcal{L}_1 \cdots \mathcal{L}_n) \end{array}$$

*tal que, si $\mathcal{L}_i$ corresponde a hipersuperficies $Y_i$ que se intersectan transversalmente, el número $(\mathcal{L}_1 \cdots \mathcal{L}_n)$ es igual al número de puntos de $Y_1 \cap \cdots \cap Y_n$.*

## 9. Fibrados de línea sobre variedades abelianas

Empezamos con dos resultados generales describiendo fibrados de línea sobre productos de variedades proyectivas. Como sólo usaremos "fibrados de línea", después de un tiempo hablaremos simplemente de "fibrados".

**Teorema 9.1** (Teorema del subibaja)**.** *Sean $X$, $Y$ variedades y sea $\mathcal{L}$ un fibrado de línea sobre $X \times Y$. Denotamos $p_1$, $p_2$ las dos proyecciones de $X \times Y$ y para cada $x \in X$ (resp. $y \in Y$), denotamos $i_x(y) = (x, y)$ (resp. $j_y(x) = (x, y)$). Suponemos que para cada $x \in X$, tenemos $i_x^* \mathcal{L}$ trivial, entonces existe $\mathcal{M}$, un fibrado de línea sobre $X$ tal que $\mathcal{L} = p_1^* \mathcal{M}$. Si, además, existe $y_0 \in Y$ tal que $j_y^* \mathcal{L}$ sea trivial, entonces $\mathcal{L}$ es trivial.*

*Demostración.* Ver [7, Lema A.7.2.3, p. 123] o [12, Corolario 6, p. 54]. □

**Teorema 9.2** (Teorema del cubo abstracto)**.** *Sean $X$, $Y$, $Z$ tres variedades proyectivas y $x_0$, $y_0$, $z_0$ puntos sobre ellas. Sea $\mathcal{L}$ un fibrado sobre $X \times Y \times Z$ con la propiedad de tornarse trivial cuando se le restringe a $\{x_0\} \times Y \times Z$, $X \times \{y_0\} \times Z$, $X \times Y \times \{z_0\}$, entonces $\mathcal{L}$ es trivial sobre $X \times Y \times Z$.*

*Demostración.* Ver [12, p. 55]. □

Se deduce fácilmente el teorema siguiente

**Teorema 9.3** (Teorema del cubo para variedades abelianas)**.** *Sea $A$ una variedad abeliana y $\mathcal{L}$ un fibrado sobre $A$. Para cada subconjunto $I \subset \{1, 2, 3\}$ denotamos $s_I(x_1, x_2, x_3) = \sum_{i \in I} x_i$. El siguiente fibrado es trivial sobre $A \times A \times A$:*

$$\sum_{I \neq \emptyset} (-1)^{|I|} s_I^* \mathcal{L} = 0.$$

*Demostración.* Llamamos $\mathrm{Cubo}(\mathcal{L})$ al miembro izquierdo de la última igualdad. Aplicamos Teorema 9.2 mostrando que $\mathrm{Cubo}(\mathcal{L})$ restringido a $A \times A \times \{0\}$ es trivial; notando simetrías tendremos también que $\mathrm{Cubo}(\mathcal{L})$ es trivial sobre los dos trozos $A \times \{0\} \times A$ y $\{0\} \times A \times A$. Si denotamos $i(x, y) = (x, y, 0)$ tenemos $s_{123} \circ i = s_{12} \circ i$ y también $s_{23} \circ i = s_2 \circ i$, $s_{12} \circ i = s_1 \circ i$ y $s_3 \circ i = 0$; así la fórmula deseada cumple

$$i^*(\mathrm{Cubo}(\mathcal{L})) = i^*(s_{12}^* \mathcal{L} - s_{12}^* \mathcal{L} - s_2^* \mathcal{L} - s_1^* \mathcal{L} + s_1^* \mathcal{L} + s_2^* \mathcal{L}) = 0.$$

□

**Corolario 9.4.** *Sea $f$, $g$, $h$ tres morfismos de $X$ hacia una variedad abeliana $A$ y $\mathcal{L}$ un fibrado sobre $A$. El siguiente fibrado es trivial sobre $X$:*

$$(f+g+h)^* \mathcal{L} - (f+g)^* \mathcal{L} - (g+h)^* \mathcal{L} - (f+h)^* \mathcal{L} + f^* \mathcal{L} + g^* \mathcal{L} + f^* \mathcal{L} = 0.$$

*Demostración.* Considerando $(f,g,h): X \to A^3$, la igualdad anterior es equivalente a la fórmula $(f,g,h)^*(\mathrm{Cubo}(\mathcal{L})) = 0$. □

Aplicando Corolario 9.4 con $X = A$ y las aplicaciones $f = [n]$, $g = [1] = id_A$, $h = [-1]$, obtenemos

$$[n]^*\mathcal{L} - [n+1]^*\mathcal{L} - [0]^*\mathcal{L} - [n-1]^*\mathcal{L} + [n]^*\mathcal{L} + \mathcal{L} + [-1]^*\mathcal{L} = 0.$$

Por inducción deducimos

**Lema 9.5** (Mumford)**.** *Sea $\mathcal{L}$ un fibrado sobre una variedad abeliana $A$. Tenemos:*

$$[n]^*\mathcal{L} = \frac{n^2+n}{2}\mathcal{L} + \frac{n^2-n}{2}[-1]^*\mathcal{L}. \tag{9.1}$$

*En particular si $\mathcal{L}$ es simétrico (es decir $[-1]^*\mathcal{L} = \mathcal{L}$) tendremos*

$$[n]^*\mathcal{L} = \mathcal{L}^{n^2}. \tag{9.2}$$

*Si $\mathcal{L}$ es antisimétrico (es decir $[-1]^*\mathcal{L} = \mathcal{L}^{-1}$) tendremos*

$$[n]^*\mathcal{L} = \mathcal{L}^{n}. \tag{9.3}$$

Para el corolario siguiente utilizamos la noción de número de intersección de $n$ divisores (o fibrados) sobre una variedad proyectiva de dimensión $n$.

**Corolario 9.6.** *Sea $A$ una variedad abeliana de dimensión $g$; la multiplicación $[n]_A$ es un morfismo finito de grado $n^{2g}$.*

*Demostración.* Escogemos un fibrado amplio y simétrico $\mathcal{L}$, el número de intersección (ver Teorema 8.6) $(\mathcal{L})^g := (\mathcal{L}\cdot\cdots\cdot\mathcal{L}) > 0$ y calculamos $([n]^*\mathcal{L})^g = (n^2\mathcal{L})^g = n^{2g}(\mathcal{L})^g = \deg([n])(\mathcal{L})^g$. Como el fibrado $\mathcal{L}$ es amplio, $(\mathcal{L})^g > 0$ y concluimos. □

Con respecto a traslaciones, la propiedad más importante es la siguiente.

**Teorema 9.7** (Teorema del cuadrado)**.** *Sea $A$ una variedad abeliana y $\mathcal{L}$ un fibrado sobre $A$. La aplicación $\phi_{\mathcal{L}} : A \to \mathrm{Pic}(A)$ definida por $\phi_{\mathcal{L}}(x) := t_x^*\mathcal{L} \otimes \mathcal{L}^{-1}$ es un homomorfismo de grupos.*

*Demostración.* Sigue de aplicar Corolario 9.4 con $f(x) = x$, $g(x) = a$ y $h(x) = b$. □

**Definición 9.8.** El grupo $\mathrm{Pic}^0(A)$ es el subgrupo de los $\mathcal{L} \in \mathrm{Pic}(A)$ tales que $\phi_{\mathcal{L}} = 0$.

Observamos que, utilizando el teorema del cuadrado, vemos que $t_x^*\mathcal{L} \otimes \mathcal{L}^{-1} \in \mathrm{Pic}^0(A)$. En particular el homomorfismo $\phi_{\mathcal{L}}$ toma valores en $\mathrm{Pic}^0(A)$.

**Proposición 9.9.** *Un fibrado $\mathcal{L}$ es antisimétrico, i.e. $[-1]^*\mathcal{L} = \mathcal{L}^{-1}$, si y sólo si*

$$s_{1,2}^*\mathcal{L} = p_1^*\mathcal{L} + p_2^*\mathcal{L} \text{ en } \mathrm{Pic}(A\times A), \tag{9.4}$$

*si y sólo si $\mathcal{L} \in \mathrm{Pic}^0(A)$, es decir si $K(\mathcal{L}) = A$.*

*Un fibrado $\mathcal{L}$ es simétrico, i.e. $[-1]^*\mathcal{L} = \mathcal{L}$, si y sólo si*

$$s_{1,2}^*\mathcal{L} + d_{1,2}^*\mathcal{L} = 2p_1^*\mathcal{L} + 2p_2^*\mathcal{L} \text{ en } \mathrm{Pic}(A\times A)$$

*donde se usó las notaciones $s_{1,2}(x_1,x_2) = x_1 + x_2$, $d_{1,2}(x_1,x_2) = x_1 - x_2$ y $p_i(x_1,x_2) = x_i$.*

*Demostración.* Ver [7, Proposiciones A.7.3.2 y A.7.3.3, p. 129]. □

Observamos que, sobre el cuerpo $\mathbb{C}$, las fórmulas precedentes se pueden demostrar fácilmente, representando un fibrado a través del teorema de Appell-Humbert (teorema 3.1).

## 10. Polarización, isogenía, variedad dual

**10.1. Isogenías.** Recordamos que, en el "mundo" de la característica $p$, una extensión finita de cuerpos $L/K$ se descompone en una parte separable y una parte inseparable. De la misma manera un morfismo finito $\phi : X \to Y$ se descompone en una parte separable y una parte inseparable y tenemos $\deg \phi = \deg_{sep} \phi . \deg_{insep} \phi$. La definición siguiente corresponde, sobre $\mathbb{C}$, a la Definición 1.13.

**Definición 10.1.** Una *isogenía* $\alpha : A \to B$ entre dos variedades abelianas es un homomorfismo que cumple:

- El núcleo de $\alpha$ es finito.
- El homomorfismo $\alpha$ es sobreyectivo.
- Tenemos $\dim A = \dim B$.

**Definición 10.2.** El *grado* de una isogenía $\alpha : A \to B$ es su grado como morfismo finito, es decir $\deg(\alpha) = [K(A) : \alpha^*(K(B))]$. Cuando la isogenía es separable tenemos $\deg(\alpha) = |(\ker \alpha)(\bar{K})|$; en el caso general, si $p^e$ es el grado de inseparabilidad de la extensión $K(A)/\alpha^*(K(B))$, tenemos $\deg(\alpha) = p^e|(\ker \alpha)(\bar{K})|$.

De hecho dos de las tres propiedades implican la tercera. El principal ejemplo de isogenía es la multiplicación por un entero $n \neq 0$, pero otro ejemplo clave es el llamado *Frobenius* que sólo existe en característica $p$.

**Definición 10.3.** Sea $X$ una variedad (proyectiva) definida sobre un cuerpo $K$ de característica $p$. El *Frobenius* de $X$ es el morfismo definido en coordenadas por

$$\mathrm{Frob}_X(x_0, \ldots, x_n) := (x_0^p, \ldots, x_n^p)$$

Su imagen es una variedad también definida sobre $K$ y denotada $X^{(p)}$.

*Observación* 10.4. La definición no depende de las coordenadas; además si $X$ es definida sobre el cuerpo finito $\mathbb{F}_p$, tenemos $X^{(p)} = X$. El Frobenius es el ejemplo tipo de morfismo inseparable, es decir que la extensión $K(X)/\mathrm{Frob}_X^*(K(X^{(p)}))$ es una extensión finita *puramente inseparable* (de grado $p^{\dim X}$). Observamos que la diferencial $(d\,\mathrm{Frob}_X)_x : \mathrm{Tan}_x(X) \to \mathrm{Tan}_{\mathrm{Frob}_X(x)}(X^{(p)})$ es la aplicación nula.

**Teorema 10.5.** *Sea $A$ una variedad abeliana de dimensión $g$ sobre un cuerpo $K$. Para todo $n \neq 0$ la multiplicación $[n] = [n]_A$ es una isogenía de grado $\deg[n] = n^{2g}$.*

- *Si $\mathrm{car}(K) = 0$ o $\mathrm{car}(K) = p$ no divide $n$, entonces $\ker[n](\bar{K}) \cong (\mathbb{Z}/n\mathbb{Z})^{2g}$.*
- *Si $\mathrm{car}(K) = p$, existe $r = r_A \in [0, g]$ tal que $\ker[p^m](\bar{K}) \cong (\mathbb{Z}/p^m\mathbb{Z})^r$. El entero $r_A$ se llama el p-rango de $A$.*

*Demostración.* Utilizaremos el siguiente lema elemental de grupos (ejercicio).

"*Un grupo conmutativo de cardinal $n^r$ y tal que para todo $m$ divisor de $n$, el cardinal de los elementos cancelados por $m$ es igual a $m^r$ es necesariamente isomorfo a $(\mathbb{Z}/n\mathbb{Z})^r$*"

Esta observación es suficiente cuando $p = \mathrm{car}(K)$ no divide $n$, porque entonces la diferencial es una aplicación inyectiva y la isogenía es separable. Cuando $n = p = \mathrm{car}(K)$ la diferencial es nula y se puede deducir que $[p]$ se factoriza a través del Frobenius, es decir existe una otra isogenía [16] $V : A^{(p)} \to A$ tal que $[p] = V \circ \mathrm{Frob}_A$. De hecho el homomorfismo de cuerpos $K(A) \to K(A)$ dado por $f \mapsto f \circ [p]$ tiene imagen contenida en $K(A)^p = K(A^{(p)})$ y obtenemos una inyección $K(A) \to K(A^{(p)})$

[16] La letra "$V$" es tradicional y corresponde a la palabra alemana *Verschiebung*.

que corresponde a un aplicación racional $V : A^{(p)} \to A$ tal que $V \circ \mathrm{Frob} = [p]$. Además el Teorema 7.6 nos dice que $V$ es un morfismo. Tenemos $p^{2g} = \deg[p] = \deg V \deg \mathrm{Frob}_A = p^g \deg V$, así $\deg V = p^g$. Suponemos que $\deg_{\mathrm{insep}} V = p^s$, con $s \in [0, g]$ entonces $\deg_{\mathrm{insep}}[p] = p^{s+g}$ y $\ker[p](\bar{K})$ tiene $p^{g-s}$ elementos. El teorema sigue con $r = g - s$. □

*Notación* 10.6. Denotaremos $A[n]$ al grupo finito $\ker[n]_A(\bar{K})$ de puntos de torsión cancelados por $n$.

El lema siguiente contiene el hecho que, como sobre $\mathbb{C}$ (Lema 1.19) la relación de isogenía es simétrica y también que una isogenía es "invertible después de tensorizar por $\mathbb{Q}$".

**Lema 10.7.** *Sea $\phi : A \to B$ una isogenía de grado $d$ entre variedades abelianas definidas sobre $K$. Entonces existe otra isogenía $\hat{\phi} : B \to A$ tal que $\hat{\phi} \circ \phi = [d]_A$ y $\phi \circ \hat{\phi} = [d]_B$.*

*Demostración.* Damos la prueba cuando la característica de $K$ no divide $d$. Tenemos claramente en este caso $\ker \phi \subset A[d]$. La imagen del homomorfismo de cuerpos $K(A) \to K(A)$ dado por $f \mapsto f \circ [d]$ se puede identificar con $K(A)^{\ker[d]}$ (el subcuerpo fijado por los elementos de $\ker[d]$ actuando por traslaciones). De la misma manera, el homomorfismo de cuerpos $K(B) \to K(A)$ dado por $h \mapsto h \circ \phi$ permite identificar $K(B)$ con $K(A)^{\ker \phi}$. Observamos que $K(A) \cong K(A)^{\ker[d]} \subset K(A)^{\ker \phi} \cong K(B)$. Ahora esta aplicación corresponde a una inyección $h \mapsto h \circ \hat{\phi}$ por una aplicación racional $\hat{\phi} : B \to A$ ; además $\hat{\phi}$ es un morfismo gracias a Teorema 7.6. Por construcción tenemos $\hat{\phi} \circ \phi = [d]_A$ entonces $\phi \circ \hat{\phi} \circ \phi = \phi \circ [d]_A = [d]_B \circ \phi$. Así, como $\phi$ es sobreyectiva, vemos que $\phi \circ \hat{\phi} = [d]_B$. □

**Definición 10.8.** Si $\mathcal{L}$ es un fibrado de línea sobre una variedad abeliana $A$, denotamos $K(\mathcal{L})$ el núcleo del homomorfismo $\phi_{\mathcal{L}} : A \to \mathrm{Pic}^0(A)$ dado por $\phi_{\mathcal{L}}(x) = t_x^* \mathcal{L} \otimes \mathcal{L}^{-1}$.

**Teorema 10.9.** *Sea $\mathcal{L}$ un fibrado de línea, amplio sobre una variedad abeliana $A$, el grupo $K(\mathcal{L})$ es finito.*

*Demostración.* Ver [7, Teorema A.7.2.10, p. 127] o [12, p. 77]. □

**10.2. Variedad abeliana dual y polarizaciones.** Enunciamos ahora la versión algebraica de la variedad abeliana dual (ver Subsección 3.2 sobre $\mathbb{C}$); se trata de formalizar la idea que el grupo $\mathrm{Pic}^0(A)$ es el grupo de puntos de una variedad abeliana.

**Definición 10.10.** Una variedad abeliana *dual* de $A$ es una variedad abeliana $B$ con un fibrado (llamado fibrado de Poincaré) $\mathcal{P} \in \mathrm{Pic}(A \times B)$ que verifica que los dos homomorfismos:

$$\begin{array}{ccc} B & \longrightarrow & \mathrm{Pic}^0(A) \\ b & \longmapsto & j_b^* \mathcal{P} \end{array} \quad \text{y} \quad \begin{array}{ccc} A & \longrightarrow & \mathrm{Pic}^0(B) \\ a & \longmapsto & i_a^* \mathcal{P} \end{array}$$

son biyecciones (donde $j_b(a) = (a, b) = i_a(b)$).

**Teorema 10.11.** *La variedad abeliana dual de $A$ existe y es única (salvo isomorfismo); es denotada $\widehat{A}$.*

*Demostración.* Ver [12, §III.13]. □

Cuando tenemos un fibrado $\mathcal{L}$ amplio sobre $A$, se puede construir $\widehat{A}$ como el cociente $A/K(\mathcal{L})$. En el caso $K(\mathcal{L}) = 0$, que corresponde a una *polarización principal*, tenemos $A \cong \widehat{A}$ y el divisor (fibrado) de Poincaré se puede describir como $\mathcal{P} = s_{12}^*\mathcal{L} - p_1^*\mathcal{L} - p_2^*\mathcal{L}$.

Para merecer el nombre de dual, tenemos la propiedad que $\widehat{\widehat{A}} \cong A$. En general $\widehat{A}$ no es isomorfa a $A$ pero hay isogenías particulares llamadas *polarizaciones* $\lambda : A \to \widehat{A}$ que son de la forma $\lambda = \phi_{\mathcal{L}}$ para un fibrado amplio $\mathcal{L}$. Se puede demostrar que $\lambda$ es simétrica en el sentido que $\check{\lambda} = \lambda$ (donde $\check{\lambda}$ es definida en la línea siguiente).

Identificando $\widehat{A}$ y $\mathrm{Pic}^0(A)$, se puede definir el dual de un homomorfismo $\alpha : A \to B$ como la composición de los homomorfismos:

$$\check{\alpha} : \widehat{B} \cong \mathrm{Pic}^0(B) \xrightarrow{\alpha^*} \mathrm{Pic}^0(A) \cong \widehat{A}.$$

Cuando $\alpha$ es una isogenía, $\check{\alpha}$ es también una isogenía (¡Cuidado! No es exactamente la misma isogenía *dual* que la isogenía definida sobre el cuerpo $\mathbb{C}$ en la primera parte).

Podemos ahora demostrar la versión algebraica del teorema de reducibilidad de Poincaré (sobre $\mathbb{C}$ ver Teorema 1.20).

**Teorema 10.12** (Teorema de reducibilidad de Poincaré)**.** *Si $B$ es una subvariedad abeliana de $A$ definida sobre $K$, existe $C$ una subvariedad abeliana de $A$ también definida sobre $K$ tal que $B \cap C$ es finito y $s(b,c) = b + c$ define una isogenía $s : B \times C \to A$.*

*Demostración.* Sea $i : B \hookrightarrow A$ la inyección; escogemos $\mathcal{L}$ amplio sobre $A$ y consideramos la isogenía $\phi_{\mathcal{L}} : A \to \widehat{A}$. Definimos $C$ como el componente conexo del núcleo de $\check{\imath} \circ \phi_{\mathcal{L}}$. Tenemos $\dim C = \dim(\ker \check{\imath}) \geqslant \dim \widehat{A} - \dim \widehat{B} = \dim A - \dim B$. Si $x \in B \cap C$ entonces $0 = \check{\imath} \circ \phi_{\mathcal{L}}(x) = \check{\imath}(t_x^*\mathcal{L} \otimes \mathcal{L}^{-1})$; si denotamos $\mathcal{L}_B$ la restricción de $\mathcal{L}$ a $B$ o sea $i^*(\mathcal{L})$ entonces tenemos $t_x^*\mathcal{L}_B \otimes \mathcal{L}_B^{-1} = 0$ que se puede traducir por $x \in K(\mathcal{L}_B)$. Observamos que $\mathcal{L}_B$ es amplio sobre $B$ y entonces $K(\mathcal{L}_B)$ es finito (por Teorema 10.9). Terminamos concluyendo que $s : B \times C \to A$ tiene un núcleo finito y en consecuencia $\dim s(B \times C) = \dim B + \dim C \geqslant \dim A$; así tenemos igualdad y $s$ es sobreyectiva. □

## 11. Representaciones de Galois

En esta sección denotamos $G_K := \mathrm{Gal}(\bar{K}/K)$ el grupo de Galois absoluto de un cuerpo $K$ (es decir, cuando $\mathrm{car}(K) = 0$ denotamos $\bar{K}$ la clausura algebraica de $K$, y cuando $\mathrm{car}(K) = p$, denotamos $\bar{K}$ la clausura algebraica separable de $K$). Este grupo, para digamos $K$ cuerpo de números, es demasiado grande para ser controlado en su totalidad y se le estudia a través de sus representaciones.

**Definición 11.1.** Sea $G_i$ una familia de grupos (resp. módulos, resp. anillos) con homomorfismos $\psi_i : G_{i+1} \to G_i$. El *límite proyectivo* es el grupo (resp. módulo, resp. anillo)

$$\varprojlim G_i := \left\{ (g_i)_i \in \prod_i G_i \mid \forall i,\ \psi_i(g_{i+1}) = g_i \right\}$$

Utilizaremos los siguientes ejemplos claves:

- El anillo de los enteros $p$-ádicos se obtiene como

$$\mathbb{Z}_p := \varprojlim \mathbb{Z}/p^n\mathbb{Z}$$

donde los morfismos son las proyecciones $\psi_n : \mathbb{Z}/p^{n+1}\mathbb{Z} \to \mathbb{Z}/p^n\mathbb{Z}$ que mandan la clase de $x$ módulo $p^{n+1}$ sobre su clase módulo $p^n$.
- El *módulo de Tate* de una variedad abeliana que se define como

$$T_p(A) = \varprojlim A[p^n],$$

donde $\psi_n : A[p^{n+1}] \to A[p^n]$ es la multiplicación por $p$. Como grupo tenemos $T_p(A) \cong \mathbb{Z}_p^{2\dim A}$, mientras $p \neq \mathrm{car}(K)$.
- Si $\mu_{\ell^n}$ es el grupo de las raíces $\ell^n$-ésimas de la unidad (para $\ell$ distinto de la característica), podemos definir análogamente

$$T_\ell(\mathbb{G}_m) := \varprojlim \mu_{\ell^n}$$

Para cada $n$ coprimo con la característica de $K$, el grupo $G_K$ actúa sobre el grupo $\ker[n](\bar{K})$ a través de un cociente finito (de hecho a través de $\mathrm{Gal}(K(A[n])/K)$), así obtenemos la representación

$$\rho_{A,n} : G_K \to \mathrm{GL}(2g, \mathbb{Z}/n\mathbb{Z}).$$

Tomando límites inductivos sobre $\ell^n$ (donde $\ell$ es primo distinto a $p = \mathrm{car}(K)$) obtenemos

$$\rho_{A,\ell^\infty} : G_K \to \mathrm{GL}(T_\ell(A)) \cong \mathrm{GL}(2g, \mathbb{Z}_\ell)$$

Sea $a \in A[m]$ y $\check{a} \in \hat{A}[m]$, escogemos $D$ divisor sobre $A$ tal que la clase de $D$ sea $\check{a} \in \mathrm{Pic}^0(A)$, entonces existe $f \in K(A)^\times$ tal que $\mathrm{div}(f) = mD$. Tomando imágenes por $[m]^*$ vemos que $\mathrm{div}(f \circ [m]) = [m]^* \mathrm{div}(f) = m([m]^* D) = m(mD + \mathrm{div}(h)) = m\,\mathrm{div}(fh)$, es decir que, ajustando constantes, existe $g \in K(A)^\times$ tal que $f \circ [m] = g^m$. Esta observación nos permite definir:

$$e_m : A[m] \times \hat{A}[m] \longrightarrow \mu_m, \quad \text{por} \quad e_m(a, \check{a}) = \frac{g(x+a)}{g(x)}, \tag{11.1}$$

observando que $e_m(a,\check{a})^m = \left(\frac{g(x+a)}{g(x)}\right)^m = \frac{f\circ[m](x+a)}{f\circ[m](x)} = 1$ y entonces $\frac{g(x+a)}{g(x)}$ es constante (independiente de $x$) y es una raíz $m$-ésima de la unidad. Las aplicaciones $e_m$ verifican las propiedades siguientes

**Teorema 11.2** (Emparejamiento[17] de Weil)**.** *Las aplicaciones $e_m : A[m] \times \hat{A}[m] \longrightarrow \mu_m$ son bilineales y cumplen:*

1. *El emparejamiento $e_m$ es no degenerado (es decir de núcleo trivial a la derecha e izquierda).*
2. *(Compatibilidad) Tenemos la relación $e_n(ma, m\check{a}) = (e_{mn}(a,\check{a}))^m$, lo que permite extender los $e_{\ell^n}$ a un emparejamiento*

$$e_{\ell^\infty} : T_\ell(A) \times T_\ell(\hat{A}) \to T_\ell(\mathbb{G}_m).$$

3. *(Galois equivariancia) Sea $\sigma \in G_K$ entonces*

$$e_m(\sigma(a), \sigma(\check{a})) = \sigma(e_m(a,\check{a})).$$

[17]En inglés *pairing*; en francés *accouplement*.

*4. Sea $\mathcal{L}$ un fibrado sobre $A$, entonces el emparejamiento*

$$e^{\mathcal{L}} : A[m] \times A[m] \to \mu_m, \qquad (a,b) \mapsto e_m(a, \phi_{\mathcal{L}}(b))$$

*es antisimétrico.*

*Demostración.* Ver [12, pp. 185–186]. □

*Observación* 11.3. Como el emparejamiento es Galois equivariante, vemos que las representaciones $\rho_n$ o $\rho_{\ell^\infty}$, a priori con valores en $\mathrm{GL}_{2g}$, toman sus valores en $\mathrm{GSp}_{2g}$, el grupo de las similitudes simplécticas.

Este emparejamiento nos permite "reconstruir" las formas de Riemann en un cuadro algebraico.

**Definición 11.4.** Sea $\mathcal{L}$ un fibrado sobre $A$, definimos un emparejamiento

$$e_\ell^{\mathcal{L}} : T_\ell(A) \times T_\ell(A) \longrightarrow \mathbb{Z}_\ell \tag{11.2}$$

por la fórmula

$$e_\ell^{\mathcal{L}}(x,y) = e_{\ell^\infty}(x, \phi_{\mathcal{L}}(y)). \tag{11.3}$$

**Relación con las formas de Riemann sobre un toro complejo.** Hemos visto que una variedad abeliana compleja de dimensión $g$ se puede ver como un toro $A(\mathbb{C}) = \mathbb{C}^g/\Lambda$ y que, a cada fibrado $\mathcal{L}$ (o divisor) amplio, corresponde una forma de Riemann que podemos describir como una aplicación bilineal antisimétrica:

$$E_{\mathcal{L}} : \Lambda \times \Lambda \longrightarrow \mathbb{Z}$$

En verdad, no conseguimos por completo recuperar algebraicamente $E_{\mathcal{L}}$, pero notando que, para una variedad abeliana compleja $A = \mathbb{C}^g/\Lambda$, tenemos $A[n] = \Lambda/n\Lambda$ y entonces

$$\varprojlim A[n] = \varprojlim \Lambda/n\Lambda = \Lambda \otimes_{\mathbb{Z}} \varprojlim \mathbb{Z}/n\mathbb{Z}, \quad \text{y} \quad T_\ell(A) = \varprojlim A[\ell^n] = \Lambda \otimes_{\mathbb{Z}} \mathbb{Z}_\ell.$$

Así tenemos $T_\ell(A) \cong \Lambda \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$ y finalmente podemos identificar

$$E_{\mathcal{L},\mathbb{Z}_\ell} : (\Lambda \otimes_{\mathbb{Z}} \mathbb{Z}_\ell) \times (\Lambda \otimes_{\mathbb{Z}} \mathbb{Z}_\ell) \longrightarrow \mathbb{Z}_\ell$$

con el emparejamiento de Weil (11.2).

La acción del anillo $\mathrm{End}(A)$ sobre $T_\ell(A)$ nos da una inyección $\mathrm{End}(A) \to \mathrm{End}_{\mathbb{Z}_\ell}(T_\ell(A))$ (ejercicio). Es un poco más sutil ver que eso induce una inyección $\mathrm{End}(A) \otimes \mathbb{Z}_\ell \to \mathrm{End}(T_\ell(A)$ entonces $\dim \mathrm{End}(A) \leqslant 4(\dim A)^2$.

La involución de Rosati puede ser definida en este contexto: escogemos un fibrado amplio $\mathcal{L}$ amplio sobre $A$ y la polarización asociada $\phi_{\mathcal{L}} : A \to \widehat{A}$ y obtenemos

$$\alpha \in \mathrm{End}(A) \otimes \mathbb{Q} \mapsto \alpha^\dagger := \phi_{\mathcal{L}}^{-1} \circ \check{\alpha} \circ \phi_{\mathcal{L}} \in \mathrm{End}(A) \otimes \mathbb{Q}$$

Observamos que si $\phi_{\mathcal{L}}$ no es una polarización principal entonces $\phi_{\mathcal{L}}^{-1}$ sólo existe después de tensorizar con $\mathbb{Q}$.

Terminamos esta sección con un resultado mucho más difícil,

**Teorema 11.5** (Tate, Zarhin, Faltings [27, 37, 40])**.** *Sea $K$ un cuerpo finito, un cuerpo de números o un cuerpo de tipo finito sobre ellos, y sean $A$, $B$ variedades abelianas sobre $K$ y $p$ un primo distinto de la característica de $K$. El homomorfismo*

$$\mathrm{Hom}(A,B) \otimes \mathbb{Z}_p \longrightarrow \mathrm{Hom}_{\mathbb{Z}_p[G_K]}(T_p(A), T_p(B)) \tag{11.4}$$

*es un isomorfismo.*

## 12. Curvas y jacobianas

El ejemplo histórico de variedad abeliana es la jacobiana de una curva: el grupo de clases de divisores de grado cero $\mathrm{Pic}^0(C)$ tiene una estructura de grupo algebraico proyectivo. Admitiendo este hecho se pueden describir algunas propiedades de esta variedad abeliana que denotamos $J_C$ y que se llama *jacobiana* de $C$. Escogiendo un punto "origen" $P_0$ tenemos el morfismo

$$j = j_{P_0} : C \to J_C, \quad \text{dado por } P \mapsto [(P) - (P_0)]$$

que se puede extender a un morfismo

$$j_r = j_{r,P_0} : C^r \to J_C, \quad \text{dado por } (P_1, \ldots, P_r) \mapsto [\sum_{i=1}^{r} (P_i) - r(P_0)].$$

Con la ayuda del teorema de Riemann-Roch (para la curva $C$), se puede ver que, cuando $g = g(C) \geqslant 1$, el morfismo $j$ es una inmersión y que $W_r(C) := j_r(C^r)$ es una subvariedad de dimensión $\min(r, g)$. En particular tenemos el siguiente resultado clásico.

**Teorema 12.1.** *La jacobiana de una curva $C$ de género $g$ es una variedad abeliana de dimensión $g$; esta variedad abeliana está dotada de un divisor canónico (salvo traslaciones) $\Theta_C := W_{g-1} = j_{g-1}(C^{g-1})$, que induce una polarización principal sobre $J_C$.*

*Demostración.* Ver [7, Teorema A.8.1.1, pp. 134–135]. □

Sobre los complejos, la construcción clásica parece diferente. Si $X$ es una curva lisa proyectiva definida sobre $\mathbb{C}$, entonces $X(\mathbb{C})$ es una superficie[18] de Riemann compacta. El espacio vectorial de las 1-formas diferenciales holomorfas $H^0(X(\mathbb{C}), \Omega^1_X)$ es de dimensión $g$, el grupo de homología singular (o de Betti) se denota $H^1(X(\mathbb{C}), \mathbb{Z}) \cong \mathbb{Z}^{2g}$ y los dos están vinculados por la integración

$$H^0(X(\mathbb{C}), \Omega^1_X) \times H^1(X(\mathbb{C}), \mathbb{Z}) \to \mathbb{C}, \quad (\omega, [\gamma]) \mapsto \int_\gamma \omega.$$

Esto nos permite ver a $H^1(X(\mathbb{C}), \mathbb{Z}) \cong \mathbb{Z}^{2g}$ como un retículo en el espacio dual $H^0(X(\mathbb{C}), \Omega^1_X)^* \cong \mathbb{C}^g$. Las *relaciones de Riemann* (ver [7, §A.6]) entre los periodos permite demostrar

**Teorema 12.2.** *El toro complejo $H^0(X(\mathbb{C}), \Omega^1_X)^*/H^1(X(\mathbb{C}), \mathbb{Z})$ es una variedad abeliana, y la forma de Riemann canónicamente asociada induce una polarización principal.*

Denotamos $J_X(\mathbb{C})$ este toro complejo; el vínculo con la presentación algebraica es dada por el teorema de Abel-Jacobi: escogiendo un punto $P_0 \in X(\mathbb{C})$, se puede definir una inyección

$$j = j_{P_0} : X(\mathbb{C}) \to J_X(\mathbb{C}) \quad \text{dada por } j_{P_0}(P)(\omega) = \int_{P_0}^{P} \omega \mod H^1(X(\mathbb{C}), \mathbb{Z})$$

y extenderla a divisores.

[18] El choque entre las palabras "curva" (objeto algebraico de dimensión 1 sobre $\mathbb{C}$) y "superficie" (objeto topológico de dimensión 2 sobre $\mathbb{R}$) es histórico e inevitable.

**Teorema 12.3** (Teorema de Abel-Jacobi)**.** *Consideramos el morfismo $j$ del grupo de los divisores de grado nulo* $\mathrm{Div}^0(X)$ *hacia* $J_X(\mathbb{C})$*, entonces $j$ es sobreyectivo y el núcleo es compuesto por los divisores principales* $\mathrm{div}(f)$*. En particular, se puede identificar* $J_X(\mathbb{C})$ *y* $\mathrm{Pic}^0(X)(\mathbb{C})$*.*

En otra dirección, cuando la curva es definida sobre un cuerpo finito, hay una relación interesante entre el número de puntos sobre $C$ y sobre $J_C$.

**Teorema 12.4** (Weil)**.** *Sea $C/\mathbb{F}_q$ una curva lisa proyectiva de género $g$. Existe enteros algebraicos $\alpha_1, \dots \alpha_{2g}$ tales que:*

1. *El conjunto de los $\alpha_i$ es estable por Galois y cada $\alpha_i$ verifica $|\alpha_i| = \sqrt{q}$; así el conjunto es permutado por $\alpha \mapsto q/\alpha$.*
2. *Para cada $m \geqslant 1$, tenemos $|C(\mathbb{F}_{q^m})| = q^m + 1 - \sum_{i=1}^{2g} \alpha_i^m$.*
3. *Tenemos $|J_C(\mathbb{F}_q)| = \prod_{i=1}^{2g}(\alpha_i - 1)$.*

Ver [7, ejercicio A.8.11, p. 150].

## 13. Alturas de Néron-Tate y Teorema de Mordell-Weil

**13.1. Buena reducción, criterio de Néron-Ogg-Shafarevich.** Sea $K$ un cuerpo de números y $v$ una plaza finita, que corresponde a un ideal primo $\mathfrak{p}_v$ y tiene cuerpo residual $\mathbb{F}_v := \mathcal{O}_K/\mathfrak{p}_v$. Podemos definir la *reducción* de puntos módulo $\mathfrak{p}_v$ o módulo $v$ como

$$\mathrm{red}_v : \mathbb{P}^n(K) \to \mathbb{P}^n(\mathbb{F}_v), \qquad (x_0 : \cdots : x_n) \mapsto (\tilde{x}_0 : \cdots : \tilde{x}_n)$$

donde $\tilde{x}$ denota la imagen en $\mathbb{F}_v$ de un elemento $\mathfrak{p}_v$-entero de $K$, y se escoge coordenadas $x_i$, quienes son $\mathfrak{p}_v$-enteras tales que una de ellas sea una $\mathfrak{p}_v$-unidad.

Se puede también dar una primera noción ingenua de reducción de una variedad proyectiva. Si $X \subset \mathbb{P}^n$ es definida por un ideal $I_X \in K[x_0, \dots, x_n]$, se define $\mathcal{I}_X = I_X \cap \mathcal{O}_K[x_0, \dots, x_n]$ y finalmente

$$\tilde{\mathcal{I}}_X = \left\{ \tilde{F} \mid F \in \mathcal{I}_X \right\} \quad \text{y} \quad \tilde{X} = \left\{ x \in \mathbb{P}^n_{\mathbb{F}_v} \mid \forall \tilde{F} \in \mathcal{I}_X,\ \tilde{F}(P) = 0 \right\}$$

Con esta definición es más o menos claro que la aplicación de reducción de puntos es compatible, es decir que nos proporciona la aplicación

$$\mathrm{red}_v : X(K) \to \tilde{X}(\mathbb{F}_v). \tag{13.1}$$

**Definición 13.1.** Diremos que $X$ tiene *buena reducción* en $v$ si la reducción $\tilde{X}$ es lisa.

El defecto de esta definición es que la noción depende de la inmersión $X \hookrightarrow \mathbb{P}^n$, una definición más intrínseca es que $X$ tiene buena reducción si existe un modelo donde $X$ tiene buena reducción en el sentido ingenuo. Con ambas definiciones el hecho más importante es el siguiente

**Proposición 13.2.** *Sea $X$ una variedad proyectiva lisa definida sobre un cuerpo de números $K$. Existe un conjunto finito $S$ de ideales primos de $K$, tal que para todo $\mathfrak{p}_v \notin S$, la variedad $X$ tiene buena reducción en $\mathfrak{p}_v$.*

*Demostración.* Utilizando la caracterización de la propiedad de ser liso por el criterio de Jacobi (un punto es liso si un menor de tamaño adecuado de la matriz de la diferencial de las ecuaciones es no nulo), esta propiedad sigue siendo verdad módulo $\mathfrak{p}_v$ para cada $\mathfrak{p}_v$ que no divide este determinante. □

Volvemos a las variedades abelianas. Si $A$ es una variedad abeliana definida sobre $K$ y tiene buena reducción en $v$, entonces $\tilde{A}_v$ es también una variedad abeliana (definida sobre $\mathbb{F}_v$) y el morfismo

$$\mathrm{red}_v : A(K) \to \tilde{A}_v(\mathbb{F}_v) \tag{13.2}$$

es un homomorfismo de grupos. Obviamente este homomorfismo no es en general inyectivo (ejercicio) pero una propiedad importante tiene que ver con inyectividad.

**Lema 13.3.** *Sea $m \geqslant 2$ un entero y $A$ una variedad abeliana definida sobre un cuerpo de números $K$. Sea $\mathfrak{p}_v$ un ideal primo de $K$ que no divide $m$ y donde $A$ tiene buena reducción. El homomorfismo de reducción*

$$\mathrm{red}_v : A[m](K) \to A(\mathbb{F}_v)$$

*es inyectivo.*

*Demostración.* Ver [7, Teorema C.1.4, p. 263]. □

Observamos que el análogo para el grupo $\mathbb{G}_m$ puede ser demostrado de manera elemental.

**Lema 13.4** (Análogo del lema 13.3 para $\mathbb{G}_m$)**.** *Sean $p$ y $m$ coprimos. Sea $\mathbb{G}_m[m] = \mu_m$ el grupo de las raíces $m$-ésima de la unidad, $K = \mathbb{Q}(\mu_m)$ y $\mathfrak{p}$ un ideal de $K$ con característica residual $p$. Entonces la reducción módulo $\mathfrak{p}$ es inyectiva sobre $\mu_m$.*

*Demostración.* Sean $\zeta \neq \zeta'$ dos raíces $m$-ésimas de la unidad. Si $\zeta \equiv \zeta' \mod \mathfrak{p}$ tenemos también $1 - \zeta^{-1}\zeta' \equiv 0 \mod \mathfrak{p}$. Pero es elemental ver que si $\zeta'' \neq 1$ es una raíz de la unidad entonces $1 - \zeta''$ es una unidad o una $p$-unidad cuando el orden de $\zeta''$ es una potencia de $p$. □

Volviendo a la noción intrínseca de buena reducción, tenemos la caracterización siguiente en términos de la representación sobre el módulo de Tate.

**Teorema 13.5** (Criterio de Néron-Ogg-Shafarevich)**.** *Una variedad abeliana $A$ tiene buena reducción en $v$ si y sólo si el subgrupo de inercia de $v$ en $G_K$ actúa trivialmente sobre $T_\ell(A)$.*

*Demostración.* Para el caso de las curvas elípticas, ver [15, Teorema VII.7.1]. La prueba se adapta al caso general, utilizando los modelos de Néron de una variedad abeliana [36]. □

**13.2. Alturas de Weil.** Cada cuerpo de números $K$ es dotado de un conjunto de lugares: un lugar para cada ideal primo de $K$ y un lugar para cada inmersión $\sigma : K \hookrightarrow \mathbb{R}$ o par de inmersión conjugada $\tau, \bar{\tau} : K \hookrightarrow \mathbb{C}$. Denotamos $n_v = [K_v : \mathbb{Q}_v]$ por un ideal primo y $n_v = 1$ (resp. $n_v = 2$) si $v$ es real (resp. compleja). Asociamos a cada lugar un valor absoluto $|\cdot|_v : K \to \mathbb{R}$, normalizando de manera que se cumple la fórmula siguiente

**Teorema 13.6** (fórmula del producto)**.** *Para $\alpha \in K^\times$,*

$$\prod_{v \in M_K} |\alpha|_v^{n_v} = 1. \tag{13.3}$$

**Definición 13.7.** Sea $P = (x_0, \ldots, x_n) \in \mathbb{P}^n(K)$, la *altura* (logarítmica) de Weil está dada por la fórmula:

$$h(P) := \frac{1}{[K : \mathbb{Q}]} \sum_{v \in M_K} n_v \log \max_{i=0}^{n} |x_i|_v. \tag{13.4}$$

Observamos que la definición no depende de las coordenadas proyectivas de $P$, gracias a la fórmula del producto (13.6). Además, $h(P)$ no depende del cuerpo de racionalidad de $P$, así podemos ver a $h$ como una función $h : \mathbb{P}^n(\bar{\mathbb{Q}}) \to \mathbb{R}$. La propiedad importante más básica es la siguiente.

**Teorema 13.8** (Northcott)**.** *El conjunto siguiente es finito para todo $D \geqslant 1$, $T \geqslant 1$:*

$$\{P \in \mathbb{P}^n(\bar{\mathbb{Q}}) \mid h(P) \leqslant T \text{ y } [\mathbb{Q}(P) : \mathbb{Q}] \leqslant D\}.$$

En particular para un cuerpo de números $K$, vemos que el conjunto $\{P \in \mathbb{P}^n(K) \mid h(P) \leqslant T\}$ es finito.

Podemos extender la noción de alturas a variedades proyectivas considerando inmersiones $\phi : V \hookrightarrow \mathbb{P}^n$ y definiendo $h_\phi(P) := h_{\mathbb{P}^n}(\phi(P))$. Cuando $\mathcal{L}$ es un fibrado amplio sobre $V$, se puede asociar una inmersión $\phi_{\mathcal{L}} : V \hookrightarrow \mathbb{P}^n$ que es única sólo módulo una transformación lineal $\alpha \in \mathrm{PGL}_{n+1}$. El lema elemental siguiente muestra que eso no altera mucho las alturas (ejercicio).

**Lema 13.9.** *Sea $\alpha \in \mathrm{PGL}_{n+1}$ un automorfismo $\alpha : \mathbb{P}^n \to \mathbb{P}^n$, existe una constante $C = C_\alpha$ tal que*

$$|h(\alpha(P)) - h(P)| \leqslant C.$$

Sea $\mathcal{L}$ un fibrado sobre una variedad proyectiva $V$, se puede escribir como la diferencia de dos fibrados muy amplios: $\mathcal{L} = \mathcal{L}_1 \otimes \mathcal{L}_2^{-1}$ y definir la altura asociada a $\mathcal{L}$ por

$$h_{\mathcal{L}}(P) = h_{\mathcal{L}_1}(P) - h_{\mathcal{L}_2}(P) = h(\phi_{\mathcal{L}_1}(P)) - h(\phi_{\mathcal{L}_2}(P)) \tag{13.5}$$

Observamos que $h_{\mathcal{L}}$ es única salvo una función acotada; se denota esto tradicionalmente con $h_{\mathcal{L}} = h'_{\mathcal{L}} + O(1)$. [19]

**Teorema 13.10** (Máquina de las alturas de Weil)**.** *A cada fibrado $\mathcal{L}$ sobre $V$ variedad proyectiva, es asociada una altura $h_{\mathcal{L}} : V(\bar{Q}) \to \mathbb{R}$, única módulo funciones acotadas. Estas alturas verifican las propiedades siguientes:*

1. *(normalización) Sea $\mathcal{O}(1)$ el fibrado de Serre sobre $\mathbb{P}^n$, entonces*
$$h_{\mathcal{O}(1)} = h_{\mathbb{P}^n} + O(1)$$
2. *(aditividad) Si $\mathcal{L}$ y $\mathcal{M}$ son dos fibrados sobre $V$, entonces*
$$h_{\mathcal{L}\otimes\mathcal{M}} = h_{\mathcal{L}} + h_{\mathcal{M}} + O(1)$$
3. *(functorialidad) Sean $\phi : V \to W$ un morfismo de variedades proyectivas y $\mathcal{L}$ un fibrado sobre $W$, entonces*
$$h_{\mathcal{L}} \circ \phi = h_{\phi^*\mathcal{L}} + O(1)$$
4. *(positividad) Sea $\mathcal{L}$ un fibrado sobre $V$ con secciones no nulas; denotamos $Z$ el conjunto de los ceros comunes de todas las secciones, entonces*
$$\forall P \in V(\bar{\mathbb{Q}})\backslash Z, \quad h_{\mathcal{L}}(P) \geqslant -c.$$

*Demostración.* Ver [7, Teoremas B.3.2 y B.3.6]. □

[19] El contexto y la tipografía permite distinguir entre el fibrado $\mathcal{O}(1)$ y la función acotada $O(1)$.

**13.3. Alturas sobre variedades abelianas.** Utilizando la máquina de las alturas de Weil y las relaciones entre fibrados sobre variedades abelianas obtenemos la fórmulas siguientes (ejercicio).

**Proposición 13.11.** *Sean $A$ una variedad abeliana y $\mathcal{L}$ un fibrado sobre ella.*

1. *Si $\mathcal{L}$ es simétrico, entonces*
$$h_{\mathcal{L}}([n](P)) = n^2 h_{\mathcal{L}}(P) + O(1)$$
*y también*
$$h_{\mathcal{L}}(P+Q) + h_{\mathcal{L}}(P-Q) = 2h_{\mathcal{L}}(P) + 2h_{\mathcal{L}}(Q) + O(1)$$
2. *Si $\mathcal{L}$ es antisimétrico, tenemos:*
$$h_{\mathcal{L}}([n](P)) = nh_{\mathcal{L}}(P) + O(1)$$
*y también*
$$h_{\mathcal{L}}(P+Q) = h_{\mathcal{L}}(P) + h_{\mathcal{L}}(Q) + O(1)$$

*Demostración.* Damos la prueba de la primera relación, mientras que las otras se demuestran de manera similar. Sigue de la functorialidad $h_{\mathcal{L}}([n](P)) = h_{[n]*\mathcal{L}}(P) + O(1)$, y de las relaciones de fibrados y la aditividad $h_{[n]*\mathcal{L}}(P) = h_{n^2\mathcal{L}}(P) = n^2 h_{\mathcal{L}}(P) + O(1)$. □

Estas relaciones nos dicen que la altura asociada a un fibrado simétrico (resp. antisimétrico) es casi-cuadrática (resp. casi-lineal). Gracias al siguiente lema de Tate podemos definir las alturas canónicas de Néron-Tate.

**Lema 13.12** (Tate)**.** *Sean $S$ un conjunto, $\alpha > 1$ y dos aplicaciones $h : S \to \mathbb{R}$ y $\phi : S \to S$ tales que $|h(\phi(x)) - \alpha h(x)| \leqslant c_1$ entonces la sucesión $\alpha^{-n} h(\phi^n(x))$ es convergente y la función*
$$\hat{h}(x) := \lim_{n\to\infty} \frac{h(\phi^n(x))}{\alpha^n},$$
*cumple las dos propiedades*

1. $|\hat{h}(x) - h(x)| \leqslant c_1/(\alpha - 1)$;
2. $\hat{h}(\phi(x)) = \alpha \hat{h}(x)$.

*Demostración.* Empezamos por verificar que $u_n := \alpha^{-n} h(\phi^n(x))$ es una sucesión de Cauchy. De hecho, como $-c_1 \leqslant h(\phi^n(x)) - \alpha h(\phi^{n-1}(x)) \leqslant c_1$, multiplicando por $\alpha^{-n}$ y sumando las desigualdades, obtenemos
$$-c_1 \left( \frac{1}{\alpha^n} + \dots \frac{1}{\alpha^{m+1}} \right) \leqslant u_n - u_m \leqslant c_1 \left( \frac{1}{\alpha^n} + \dots \frac{1}{\alpha^{m+1}} \right)$$
Esto comprueba que $u_n$ es una sucesión de Cauchy. Tomando $n$ infinito obtenemos
$$-\frac{c_1}{\alpha^m(\alpha-1)} \leqslant \hat{h}(x) - \alpha^{-m} h(\phi^m(x)) \leqslant \frac{c_1}{\alpha^m(\alpha-1)}$$
y en particular que $|\hat{h}(x) - h(x)|$ es acotada por $c_1/(\alpha - 1)$. Finalmente
$$\hat{h}(\phi(x)) = \lim_{n\to\infty} \frac{h(\phi^n(\phi(x)))}{\alpha^n} = \alpha \lim_{n\to\infty} \frac{h(\phi^{n+1}(x))}{\alpha^{n+1}} = \alpha \hat{h}(x).$$
□

Este lema 13.12 juntado con la proposición 13.11 nos permite definir las *alturas canónicas*, también llamadas *alturas de Néron-Tate*.

**Definición 13.13.** Sea $A$ una variedad abeliana y $\mathcal{L}$ un fibrado sobre ella.

- Si $\mathcal{L}$ es simétrico, ponemos:
$$\hat{h}_{\mathcal{L}}(P) = \lim_{n\to\infty} 4^{-n} h_{\mathcal{L}}([2^n](P))$$
- Si $\mathcal{L}$ es antisimétrico, ponemos:
$$\hat{h}_{\mathcal{L}}(P) = \lim_{n\to\infty} 2^{-n} h_{\mathcal{L}}([2^n](P))$$

**Teorema 13.14.** *Sea $\mathcal{L}$ un fibrado amplio simétrico sobre una variedad abeliana $A$ definida sobre un cuerpo de números $K$. La altura canónica $\hat{h}_{\mathcal{L}}$ satisface las propiedades:*

1. *Es una forma cuadrática, en particular verifica la ley del paralelogramo:*
$$\hat{h}_{\mathcal{L}}(P+Q) + \hat{h}_{\mathcal{L}}(P-Q) = 2\hat{h}_{\mathcal{L}}(P) + \hat{h}_{\mathcal{L}}(Q)$$
2. *Es definida positiva, es decir que, después de tensorizar por $\mathbb{R}$, la forma cuadrática real $\hat{h}_{\mathcal{L},\mathbb{R}} : A(K)\otimes_{\mathbb{Z}}\mathbb{R} \to \mathbb{R}$ es definida positiva en el sentido usual. En particular: $\hat{h}_{\mathcal{L}}(P) = 0$ si y sólo si $P$ es torsión.*

*Demostración.* Sea $h_{\mathcal{L}}$ una altura de Weil asociada a $\mathcal{L}$, utilizando la relación (9.4) y la máquina de alturas, deducimos que $h_{\mathcal{L}}(P+Q)+h_{\mathcal{L}}(P-Q) = 2h_{\mathcal{L}}(P)+2h_{\mathcal{L}}(Q)+O(1)$. Remplazando $P$ y $Q$ por $2^nP$ y $2^nQ$ y dividiendo por $4^n$ y haciendo $n\to\infty$ nos da la ley del paralelogramo, que caracteriza formas cuadráticas. Utilizando el teorema de Northcott, se puede verificar la segunda parte. □

**13.4. Teorema de Mordell-Weil.** La finalidad de esta sección es de dar un esbozo de demostración del teorema siguiente.

**Teorema 13.15** (Mordell-Weil)**.** *Sea $A$ una variedad abeliana definida sobre un cuerpo de números $K$, el grupo $A(K)$ es un grupo de tipo finito, o sea, existe $r \geqslant 0$ y puntos $P_1,\dots, P_r$ en $A(K)$ tales que:*
$$A(K) = A(K)_{\mathrm{tor}} \oplus \mathbb{Z}P_1 \oplus \cdots \oplus \mathbb{Z}P_r$$
*donde el grupo de torsión $A(K)_{\mathrm{tor}}$ es finito.*

*Demostración.* Como para las curvas elípticas, la prueba combina una versión "débil" del teorema con la teoría de alturas. Damos debajo un esbozo de la prueba del Teorema débil de Mordell-Weil y el lema que junta los dos argumentos. □

**Lema 13.16** (lema del descenso)**.** *Sea $G$ un grupo abeliano tal que $G/2G$ es finito y el grupo es dotado de una forma cuadrática $q : G \to \mathbb{R}$ tal que para todo real $X$ el conjunto $\{x \in G \mid q(x) \leqslant X\}$ es finito. Entonces el grupo $G$ es un grupo de tipo finito.*

Observamos que se podría remplazar 2 en este lema por cualquier $m \geqslant 2$.

*Demostración.* Empezamos por notar que $q$ es positiva (si existe $x \in G$ tal que $q(x) < 0$, entonces tenemos $q(nx) = n^2q(x)$ y el conjunto $\{x \in G \mid q(x) \leqslant 0\}$ seria infinito) y así podemos definir una semi-norma $|x| = \sqrt{q(x)}$. Sean $y_1, \dots y_m$ representantes de $G/2G$, denotamos $C = \max_i |y_i|$ y $S := \{x \in G \mid q(x) \leqslant C^2\}$; podemos demostrar que $S$ genera el grupo $G$. Sea $x$ un punto de $G$, su clase módulo $2G$ es igual a la clase de $y_{i_1}$, es decir existe $x_1 \in G$ tal que $x = 2x_1 + y_{i_1}$. Observamos que
$$2|x_1| = |2x_1| = |x - y_{i_1}| \leqslant |x| + |y_{i_1}| \leqslant |x| + C$$

entonces o $x \in S$ o tenemos $|x| > C$ y entonces $|x_1| \leqslant \frac{|x|+C}{2} < |x|$. Iterando el proceso encontramos una sucesión de $x_k \in G$ tales que $x_k = 2x_{k+1} + y_{i_k}$ con la propiedad que

$$|x_k| < |x_{k-1} < \dots |x_1| < |x|.$$

El conjunto de los puntos $x_k$ tales que $|x_k| \leqslant |x|$ es finito y entonces existe un $k$ tal que $|x_k| \leqslant C$. Por tanto se puede expresar el punto $x$ como combinación lineal de $x_k \in S$ y dos $y_i$ que también pertenecen a $S$. □

**Teorema 13.17** (Teorema débil de Mordell-Weil)**.** *Sea $A$ una variedad abeliana definida sobre un cuerpo de números $K$, el grupo $A(K)/2A(K)$ es un grupo finito.*

El primer paso de la demostración es de agrandar el cuerpo hasta que contenga los coordenadas de los puntos de 2-torsión:

**Paso 1.** El siguiente lema permite de agrandar $K$ hasta que $A[2] \subset A(K)$.

**Lema 13.18.** *Si $L/K$ es galoisiana finita y si $A(L)/2A(L)$ es finito, entonces $A(K)/2A(K)$ es finito.*

*Demostración.* Podemos construir una inyección del núcleo de $A(K)/2A(K) \to A(L)/2A(L)$ en el conjunto de las funciones de $G = \mathrm{Gal}(L/K)$ hacia $A[2]$. □

**Paso 2.** Supongamos ahora que $K$ es tal que $A[2] \subset A(K)$. Se define un emparejamiento llamado *emparejamiento de Kummer* $\lambda : A(K) \times G_K \longrightarrow A[2]$ de la manera siguiente: sea $(P, \sigma) \in A(K) \times G_K$, escogemos $Q \in A(\bar{K})$ tal que $2Q = P$, entonces se define $\lambda(P, \sigma) = \sigma(Q) - Q$, observando que $2\lambda(P, \sigma) = [2]\sigma(Q) - [2]Q = \sigma([2](Q)) - [2]Q = \sigma(P) - P = 0$ y por consiguiente $\lambda(P) \in A[2]$.

**Lema 13.19.** *Sea $L = K([2]^{-1}A(K))$ el compositum de los cuerpos donde son definidos los puntos $Q$ tal que $2Q \in A(K)$. El emparejamiento de Kummer induce un emparejamiento perfecto (es decir el núcleo a la derecha y el núcleo a la izquierda son triviales)*

$$\lambda : A(K)/2A(K) \times \mathrm{Gal}(L/K) \longrightarrow A[2].$$

De este lema, deducimos que $A(K)/2A(K)$ es finito si y sólo si $L/K$ es una extensión finita.

**Paso 3.** Demostramos que los cuerpos $K(Q)$ con $Q \in [2]^{-1}A(K)$ son no ramificados afuera de un conjunto finito de lugares de $K$, más precisamente si $S$ es el conjunto de los ideales primos de $K$, quienes dividen 2 o son primos de mala reducción, entonces $K(Q)/K$ no es ramificada fuera de $S$. Un teorema de Minkowski muestra entonces que hay un número finito de tales extensiones $K(Q)$ de $K$ y deducimos de esto que $L/K$ es finita. Este paso es basado sobre el lema 13.3. Se utiliza este lema para demostrar el hecho fundamental:

**Lema 13.20.** *Sea $A$ una variedad abeliana definida sobre un cuerpo de números $K$ y $m \geqslant 2$, suponemos que $[m](Q) \in A(K)$; sea $S$ es el conjunto de los ideales primos de $K$, quienes dividen $m$ o son primos de mala reducción, entonces $K(Q)/K$ no es ramificada fuera de $S$.*

*Demostración.* Denotamos $F := K(Q)$; la extensión $F/K$ es no ramificada en $v$ si y sólo si el grupo de inercia $I_v$ actúa trivialmente sobre $F$. Por definición, si $\sigma \in I_v$, la reducción módulo $v$ de $\sigma$ actúa trivialmente. Así tenemos $\sigma(Q) = Q + \lambda(P, \sigma)$ y $\tilde{Q} = \tilde{\sigma}(\tilde{Q}) = \tilde{Q} + \widetilde{\lambda(P, \sigma)}$. Entonces $\widetilde{\lambda(P, \sigma)} = 0$ y, gracias al Lema 13.3 concluimos que $\lambda(P, \sigma) = 0$ y, por consiguiente, $\sigma$ actúa trivialmente sobre $F$. □

## 14. Ejercicios

1. Sea $C$ una curva hiperelíptica y $\iota$ su involución canónica. Consideramos $L = \{(P, \iota(P)) \mid P \in C\}$. Aplicando la fórmula de adjunción a $L \subset C \times C$ (si $C \subset X$ es una curva en una superficie y $K_S$ es el divisor canónico, $C^2 + K_S \cdot C = 2g(C) - 2$) mostrar que $L^2 = -2g + 2$. Sea $\pi : C \times C \to X$ el cociente por $\sigma(P, Q) = (Q, P)$ y $L_0 = \pi(L)$, mostrar que $L_0 \cdot L_0 = -g + 1$ y $L_0 \cong \mathbb{P}^1$. En el caso $g = 2$ concluir que $L_0$ es una curva excepcional (i.e. auto-intersección $-1$ e isomorfa a $\mathbb{P}^1$).
2. Sea $E$ una variedad abeliana de dimensión 1. Definimos $\mathrm{End}^0(E) = \mathrm{End}(E) \otimes \mathbb{Q}$. Mostrar que $\mathrm{End}^0(E)$ puede ser $\mathbb{Q}$, una extensión cuadrática imaginaria o una álgebra de cuaterniones $[\mathrm{End}^0(E) : \mathbb{Q}] = 4$. Si la característica es cero, mostrar que sólo los dos primeros casos existen. (Indicación: ver [15, Teorema III.9.3]. El ejercicio siguiente muestra que el tercer caso puede ocurrir en característica positiva.)
3. Consideramos la curva elíptica $E$ sobre $\mathbb{F}_2$ definida por $y^2 + y = x^3$. Mostrar que $\mathrm{Frob}^2(x, y) = (x^4, y^4)$ coincide con $[+2]$ o $[-2]$ y deducir que $T_2(E) = 0$. Sea $a^3 = 1$ y $e^2 + e = 1$, mostrar que $\phi_{a,e}(x, y) = (a(x + 1), y + x + e)$ es un automorfismo de $E$. Verificar que en general $\phi_{a,e}$ no conmuta con $\phi_{a',e'}$ y concluir que $\mathrm{End}(E)$ no es conmutativo y debe ser un orden en una álgebra de cuaterniones.
4. Sea $A/\mathbb{F}_q$, mostrar que $|A(\mathbb{F}_q)| = \deg(Id_A - \mathrm{Frob}_A)$. Suponemos que $\phi : A \to B$ es una isogenía definida sobre $\mathbb{F}_q$, mostrar que $|A(\mathbb{F}_q)| = |B(\mathbb{F}_q)|$ (N.B. en general los grupos $A(\mathbb{F}_q)$, $B(\mathbb{F}_q)$ no son isomorfos). [Indicación: utilizar $\phi \circ (Id_A - \mathrm{Frob}_A) = (Id_B - \mathrm{Frob}_B) \circ \phi$, observar que un punto $x \in A$ pertenece a $A(\mathbb{F}_q)$ si y sólo si $\mathrm{Frob}_A(x) = x$ y probar que $Id_A - \mathrm{Frob}_A$ es una isogenía separable.]
5. Sea $K$ cuerpo de característica $\neq 2$ y $f(x) = (x - a_1) \dots (x - a_{2g+1})$ un polinomio separable (es decir que los $a_i$ son distintos). Consideramos la curva proyectiva $C$ con ecuación afín $y^2 = f(x)$ y los puntos $P_i = (a_i, 0)$ y el punto al infinito que denotamos $\infty$. Denotamos $J$ la jacobiana de $C$ y $j : C \to J$ la inmersión $j(P) := Cl\,((P) - (\infty))$.

   a) Mostrar que $\mathrm{div}(x - a_i) = 2(P_i) - 2(\infty)$ y $\mathrm{div}(y) = \sum_i (P_i) - (2g + 1)(\infty)$.

   b) Mostrar que las únicas relaciones entre los puntos $j(P_i)$ son dadas por $[2]j(P_i) = 0$ y $\sum_i j(P_i) = 0$.

   c) Mostrar que los puntos $j(P_i) \in J$ tienen orden 2 y generan el grupo $J[2] \cong (\mathbb{Z}/2\mathbb{Z})^{2g}$.
6. Utilizar el Teorema 12.4 para mostrar que si $C$ es una curva de género 2 y $N_i = |C(\mathbb{F}_{q^i})|$, entonces

$$|J_C(\mathbb{F}_q)| = \frac{N_1^2 + N_2}{2} - q.$$

   Aplicar eso a la curva $C : y^2 = x^5 + 1$, mostrando que por $p \neq 2, 5$, si $p \equiv 2, 3$ mód 5 tenemos $|J_C(\mathbb{F}_p)| = p^2 + 1$. Utilizar el Lema 13.3 y deducir que $|J_C(\mathbb{Q})_{\mathrm{tor}}|$ divide 10. Sea $\infty$ el punto "en el infinito", $P_0 = (-1, 0)$ y $Q = (0, 1)$, verificar que $\mathrm{div}(y - 1) = 5(Q) - 5(\infty)$ y también $\mathrm{div}(x + 1) = 2(P_0) - 2(\infty)$ y concluir que:

$$J_C(\mathbb{Q})_{\mathrm{tor}} =< j(P_0), j(Q) > \cong \mathbb{Z}/10\mathbb{Z}.$$

7. Sean $A$ y $B$ dos variedades abelianas definidas sobre un cuerpo de números $K$ y $v$ un lugar donde ambas tienen buena reducción; denotamos $\tilde{A}_v$ y $\tilde{B}_v$ las reducciones. Mostrar que la aplicación natural:

$$\mathrm{Hom}(A,B) \longrightarrow \mathrm{Hom}(\tilde{A}_v, \tilde{B}_v)$$

es inyectiva. [Indicación: utilizar el Lema 13.3 para demostrar que si $\Phi \neq 0$, la reducción $\tilde{\Phi}$ no puede anularse sobre todos los puntos de torsión.] Construir un ejemplo donde la aplicación no es sobreyectiva [Indicación: examinar el ejemplo del ejercicio 3.]

## Parte 3. Variedades abelianas: Aritmética

Esta tercera parte se enfoca en el difícil problema de *la clasificación de variedades abelianas principalmente polarizadas* de una dimensión $g$ dada sobre un cuerpo $K$ dado por medio de *invariantes aritméticos explícitos*.

Sección 15 establece un marco general. Dentro del marco, presenta el caso relativamente simple de $K = \mathbb{F}_p$, donde los fundamentos teóricos de una clasificación completa para todo $g$ son conocidos. Sección 16 describe conjeturas profundas dando la naturaleza de la clasificación para $K = \mathbb{Q}$ para todo $g$. Describe cómo las conjeturas son conocidas para $g = 1$ y más aún los cálculos han dado extensas tablas. Sección 17 aún asume $K = \mathbb{Q}$, aunque entra en el caso mucho más complicado de $g \geqslant 2$. Explica que las conjeturas son conocidas sólo para casos especiales, pero que hay muchos cálculos que apoyan las conjeturas.

En ambos casos, $K = \mathbb{F}_p$ y $K = \mathbb{Q}$, la clasificación se centra en funciones $L$. En el caso de $\mathbb{F}_p$, la función $L$ $L_p(A,s)$ asociada a una variedad abeliana $A$ de dimensión $g$ proviene de un solo polinomio

$$F_p(A,T) = \sum_{j=0}^{2g} a_{p,j} T^j$$

como uno tiene la fórmula $L_p(A,s) = F_p(A,p^{-s})$. En el caso de $\mathbb{Q}$, la función $L$ $L(A,s)$ es un objeto de enorme riqueza, siendo de la forma $\prod_p L_p(A,s)^{-1}$, con la definición de $L_p(A,s)$ requiriendo modificaciones importantes en los primos malos de $A$. La cuestión central en el caso del cuerpo base $\mathbb{Q}$ es cómo se comporta $F_p(A,T)$ cuando uno varía $p$.

La dificultad de consideraciones explícitas aumenta muy rápidamente con la dimensión $g$. Asimismo una "curva abeliana principalmente polarizada" es solo una curva de género uno con un punto distinguido, i.e. una curva elíptica. Una superficie abeliana principalmente polarizada es o la Jacobiana de una curva de género dos, el producto de dos curvas elípticas, o la restricción de Weil de una curva elíptica sobre una extensión cuadrática. Por lo tanto, las dos últimas secciones se enfocarán sobre todo en curvas.

La clasificación explícita de variedades abelianas principalmente polarizadas no es una cuestión puramente matemática. De hecho, es posible obtener tablas completas de tamaño modesto sólo con el uso sistemático de computadoras. Esta parte apunta a reflejar un equilibrio teórico/computacional apropiado, presentando cálculos explícitos que ilustran diferentes aspectos de la situación general. Incluimos algunos fragmentos del código en *Magma* para que incluso los principiantes sin copias de *Magma* puedan hacer algunos cómputos en la versión en línea de *Magma*. La clasificación explícita de objetos por medio de funciones $L$ es el objetivo principal

de la base de datos *L-functions and modular forms database*. Esta parte del curso también sirve como una introducción a la LMFDB, ya que cada clase corresponde directamente a una gran parte particular de la base de datos.

## 15. Invariantes geométricos y de isogenía

Esta primera sección se centra en definir invariantes y discutir la clasificación para un cuerpo base $K$. Dado $K$, fijamos una clausura algebraica $\overline{K}$. Si $K$ tiene característica finita $p$, denotamos por $\mathbb{F}_{p^e}$ el subcuerpo de $\overline{K}$ con $p^e$ elementos. Esta sección provee ejemplos para el caso relativamente fácil de cuerpos bases $K = \mathbb{F}_q$ a manera de calentamiento para el caso principal de $K = \mathbb{Q}$.

**15.1. Cuatro conjuntos interrelacionados.** Para un cuerpo arbitrario $K$ y un entero positivo $g$, existen cuatro conjuntos interrelacionados dignos de atención:

$$\begin{array}{ccc} \mathrm{PPAb}_g(K) & \to & \mathrm{Ab}_g(K) \\ m \downarrow & & \downarrow \\ A_g(K) & & \mathrm{IsAb}_g(K) \end{array} . \tag{15.1}$$

El conjunto $\mathrm{Ab}_g(K)$ es el conjunto de variedades abelianas sobre $K$ de dimensión $g$ salvo isomorfismo. Es el conjunto en el que uno podría pensar que es mejor estudiarlo primero, pero de hecho los otros tres se comportan mejor.

**15.2. Variedades abelianas principalmente polarizadas.** Nuestro objetivo principal es la descripción explícita del conjunto $\mathrm{PPAb}_g(K)$ de variedades abelianas principalmente polarizadas sobre $K$ de dimensión $g$.

*El caso* $g = 1$. El caso unidimensional puede hacerse de manera muy concreta. Para $\mathrm{car}(K) > 3$, cualquier curva elíptica $E/K$ puede ser dada por una ecuación afín

$$y^2 = x^3 + bx + c \tag{15.2}$$

con $\Delta := -4b^3 - 27c^2 \neq 0$. Sustituyendo $(x, y) \to (x/u^2, y/u^3)$ y luego multiplicando por $u^6$ obtenemos que

$$y^2 = x^3 + bu^4x + cu^6 \tag{15.3}$$

también define $E$.

En efecto, esta construcción identifica $\mathrm{PPAb}_1(K)$ con el conjunto cociente

$$\{(b, c) \in K^2 | -4b^3 - 27c^2 \neq 0\}/K^\times,$$

donde la acción está dada por $(b, c)u = (bu^4, cu^6)$. Denotemos $\mu_m(K)$ el conjunto de raíces $m$-ésimas de la unidad en $K$. Entonces el estabilizador de $(0, c)$ es $\mu_6(K)$ mientras que el de $(b, 0)$ es $\mu_4(K)$. En el caso de que ambas coordenadas sean no nulas, el estabilizador de $(b, c)$ es $\mu_2(K) = \{\pm 1\}$.

Ahora supongamos que $K$ es un cuerpo finito $\mathbb{F}_q$. Entonces,

$$|\mu_6(\mathbb{F}_q)| = \begin{cases} 6 & \text{si } q \equiv 1\ (6) \\ 2 & \text{si } q \equiv 5\ (6) \end{cases}, \qquad |\mu_4(\mathbb{F}_q)| = \begin{cases} 4 & \text{si } q \equiv 1\ (4) \\ 2 & \text{si } q \equiv 3\ (4) \end{cases}.$$

El conjunto $\{(b, c)| - 4b^3 - 27c^2 \neq 0\}$ tiene $q^2 - q$ elementos. Contando el número de órbitas, concluimos que

$$|\mathrm{PPAb}_1(\mathbb{F}_q)| = \begin{cases} 2q + 6 & \text{si } q \equiv 1\ (12) \\ 2q + 2 & \text{si } q \equiv 5\ (12) \\ 2q + 4 & \text{si } q \equiv 7\ (12) \\ 2q & \text{si } q \equiv 11\ (12) \end{cases}.$$

¡Uno querría un conteo explícito similar a este para género arbitrario $g$!

**15.3. El espacio de móduli $A_g$.** La teoría profunda de los esquemas de móduli dice que existe un esquema de módulos gruesos $A_g$ sobre $\mathbb{Z}$ para variedades abelianas. El mapa $m$ en (15.1) envía $A \in \mathrm{PPAb}_g(K)$ a su punto móduli $m(A) \in A_g(K)$. Para $K$ algebraicamente cerrado, $m$ is biyectiva. Una función sobre $\mathrm{PPAb}_g(K)$ es llamada un *invariante geométrico* si proviene de una función sobre $A_g(\overline{K})$.

*El caso $g = 1$.* Claramente $b^3/c^2$ es un invariante geométrico de una curva elíptica $E$ dada por (15.2). Por uniformidad en las características 2 y 3 que estamos excluyendo aquí, uno se concentra en

$$j = \frac{6912b^3}{4b^3 + 27c^2} = \frac{-2^8 3^3 b^3}{\Delta}.$$

El invariante $j$ identifica $A_1$ con la línea afín con coordenada $j$. En otras palabras, con una definición distinta de $j$ en los casos excluidos por la característica, $A_1 = \mathrm{Spec}(\mathbb{Z}[j])$. De esta manera uno tiene la simple fórmula

$$|A_1(\mathbb{F}_q)| = q.$$

¡Uno querría generalizar esta fórmula para $g$ arbitrario!

*Enfoque a través de curvas.* La Jacobiana de una curva es una variedad abeliana principalmente polarizada. Nuestros ejemplos provienen de curvas hiperelípticas de la forma afín

$$y^2 = f(x).$$

Aquí $\mathrm{car}(K) \neq 2$ y $f(x) \in K[x]$ es separable de grado $2g + 1$ o $2g + 2$. Para un género dado $g$, considere los espacios de móduli ásperos de curvas hiperelípticas, de todas las curvas, y de las variedades abelianas principalmente polarizadas. Por medio de la inyectividad del mapeo jacobiano, uno tiene

$$H_g \subseteq M_g \subseteq A_g. \tag{15.4}$$

Para $g = 1$, todas las inclusiones son igualdades. También $H_2 = M_2$, pero todas las demás inclusiones son estrictas.

*Dimensiones en general.* Para $g > 1$, las dimensiones relativas sobre $\mathbb{Z}$ de los tres esquemas móduli son $(2g - 1, 3g - 3, g(g + 1)/2)$. Luego para $g = 2$, las dimensiones son $(3, 3, 3)$. Como explicaremos en la tercera sección de esta parte, $|H_2(\mathbb{F}_q)| = |M_2(\mathbb{F}_q)| = q^3$ mientras que $|A_2(\mathbb{F}_q)| = q^3 + q^2$. Para $g = 3$, uno necesita ir más allá de las curvas hiperelípticas, pero aún puede usar la estrategia de las Jacobianas, pues las dimensiones son $(5, 6, 6)$. En general, $A_g$ es geométricamente conexo, lo que implica que

$$|A_g(\mathbb{F}_q)| \approx q^{g(g+1)/2}.$$

Aquí el radio de los dos lados tiene límite 1 para $g$ fijo y $q \to \infty$.

*La suryectividad de m falla.* Para $j \neq 0, 1728$, la curva elíptica

$$y^2 = x^3 - \frac{3jx}{j-1728} + \frac{2j}{j-1728} \tag{15.5}$$

tiene invariante $j$ igual a $j$. Además, $y^2 = x^3 - 1$ y $y^2 = x^3 - x$ tienen invariantes $j$ igual a 0 y 1728 respectivamente. Por lo tanto, en el caso $g = 1$, el mapeo $m$ es suryectivo. Para $g \geqslant 2$, $m$ no es suryectiva. La obstrucción a la suryectividad en el caso de género 2 es descripta en términos muy concretos en [34]. Para cuerpos finitos, $m$ es siempre suryectiva pues cuerpos finitos tienen dimensión cohomológica uno y las obstrucciones viven en el segundo grupo de cohomología.

*La inyectividad de m falla.* Sea $x \in A_g(K)$ representado por $A \in \mathrm{PPAb}_g(K)$. Para $K \subseteq K' \subseteq \overline{K}$, denotamos por $A_{K'}$ el cambio de base de $A$ a una variedad abeliana sobre $K'$. Entonces uno tiene no solo $\mathrm{Aut}_K$, sino también los grupos $\mathrm{Aut}(A_{K'})$ que pueden ser más grandes. Para $K^s$ la clausura separable de $K$ en $\overline{K}$, el grupo $\mathrm{Gal}(K^s/K)$ actúa en $\mathrm{Aut}(A_{K^s})$ con $\mathrm{Aut}(A)$ como el conjunto de puntos fijos. La fibra arriba de $m$ es entonces un conjunto de un punto indexado por un grupo de cohomología de Galois:

$$m^{-1}(x) = H^1(\mathrm{Gal}(K^s/K), \mathrm{Aut}(A_{K^s})).$$

Cuando $K$ es finito, uno tiene que $\mathrm{Gal}(K^s/K) = \hat{\mathbb{Z}}$ y la cohomología puede ser expresada en términos elementales. En particular, uno tiene

$$\sum_{t \in m^{-1}(x)} \frac{1}{\mathrm{Aut}(A_t)} = 1.$$

En el caso de las curvas elípticas, el lado izquierdo es $m$ veces $1/m$ para $m \in \{2, 4, 6\}$. Como las variedades abelianas sobre cuerpos arbitrarios siempre tienen al menos el automorfismo negación $-1$, los grupos $\mathrm{Aut}(A_t)$ son siempre no triviales, mostrando que la falla de la inyectividad es más seria que en el caso paralelo donde $A_g$ es reemplazado por $M_g$.

**15.4. Clases de isogenía.** Por definición, $\mathrm{IsAb}_g(K)$ es el conjunto de clases de isogenías de variedades abelianas de dimensión $g$ sobre $K$. Una función sobre $\mathrm{Ab}_g(K)$ es llamada *invariante de isogenía* si proviene de una función sobre $\mathrm{IsAb}_g(K)$. Para $K = \mathbb{F}_q$, uno tiene una función obvia sobre $\mathrm{Ab}_g(K)$ para cada entero positivo $e$. Ésta es $A \mapsto |A(\mathbb{F}_{q^e})|$. Notablemente, la cantidad $|A(\mathbb{F}_{q^e})|$ es un invariante de isogenía. Más aún, como describiremos, estos números pueden ser usados para indexar $\mathrm{IsAb}_p(\mathbb{F}_p)$ con un conjunto $\mathcal{L}_g(\mathbb{F}_p)$ fácil de describir.

*Conteo de puntos.* La famosa hipótesis de Riemann de Weil para variedades abelianas sobre $\mathbb{F}_q$ es la siguiente.

> *Sea $A$ una variedad abeliana $g$-dimensional sobre $\mathbb{F}_q$. Entonces existen números complejos $\alpha_1, \ldots, \alpha_{2g}$ tales que*
>
> $$|A(\mathbb{F}_{q^e})| = \prod_{j=1}^{2g} (1 - \alpha_j^e)$$
>
> *para todos los enteros positivos $e$. Más aún, estos números complejos tiene valor absoluto $\sqrt{q}$.*

La lista desordenada de los $2g$ números $\alpha_j$ está determinada por $|A(\mathbb{F}_{q^e})|$ para $e \leqslant g$, como parte del formalismo presentado a continuación.

Si $A$ es la Jacobiana de una curva $C$ de género $g$, entonces los mismo números determinan el número de puntos en $C$:

$$|C(\mathbb{F}_{q^e})| = q^e + 1 - \sum_{j=1}^{2g} \alpha_j^e.$$

Por ejemplo, si $C$ está dada por $y^2 = f(x)$ con $f(x) \in \mathbb{F}_q[x]$ de grado $2g+1$, entonces una cuenta ingenua puede ser conceptualmente formulada como

$$|C(\mathbb{F}_{q^e})| = q^e + 1 - \sum_{x \in \mathbb{F}_{q^e}} \left( \frac{f(x)}{q^e} \right). \tag{15.6}$$

Aquí estamos usando el símbolo del residuo cuadrático,

$$\left( \frac{z}{q^e} \right) = (\text{número de raíces cuadradas de } z \text{ en } \mathbb{F}_{q^e}) - 1.$$

Existen maneras mucho más rápidas de calcular $|C(\mathbb{F}_{q^e})|$ que evaluando directamente el lado derecho de (15.6).

*Funciones L desde distintos puntos de vista.* Se puede pensar a los $2g$ números $\alpha_j$ de varias formas, y diferentes términos estrechamente relacionados están involucrados. Primero que todo, un número algebraico que tiene todos sus conjugados con valor absoluto $\sqrt{q}$ es llamado un *q-número de Weil*. Luego, $\alpha_j$ es un $q$-número de Weil para todo $j$.

Como segundo punto de vista, se puede eliminar la ambigüedad del orden formando el *polinomio de Frobenius*

$$F_q(A,T) = \prod_{j=1}^{2g} (1 - \alpha_j T) =: \sum_{j=0}^{2g} a_j T^j.$$

Aquí los coeficientes pertenecen a $\mathbb{Z}$ y el polinomio es conformalmente palíndromo en el sentido que

$$a_{2g-j} = q^j a_j.$$

Así $a_0 = 1$, $a_{2g} = q^g$ y el polinomio es determinado por los coeficientes $a_1, \ldots, a_g$.

Como tercera opción, se puede escalar las raíces para obtener el *polinomio de Frobenius unitarizado*

$$f_q(A,t) = \prod_{j=1}^{2g} (1 - \frac{\alpha_j}{\sqrt{q}} t) =: \sum_{j=0}^{2g} u_j t^j.$$

Este escalamiento tiene la ventaja que el polinomio es realmente un palíndromo, aunque la desventaja que los coeficientes $u_i$ tienen en general denominadores que involucran $\sqrt{q}$.

Como una cuarta manera, se puede ver los coeficientes de $f_q(A,t)$ como el vector

$$\mathrm{fr}_p = (u_1, \ldots, u_g).$$

La hipótesis de Riemann se traduce en desigualdades entre las coordenadas, que no hacen mención de $q$ debido a la normalización. El simplex curvilineal en el cual los vectores viven se puede ver de manera natural como el conjunto $Sp_{2g}^\natural$ de clases de conjugación en el grupo simpléctico compacto $Sp_{2g}$. Notar que $Sp_{2g}$ es un subgrupo maximal del grupo complejo $Sp_{2g}(\mathbb{C})$ y una forma interior del grupo real $Sp_{2g}(\mathbb{R})$.

El caso $g = 2$ está dibujado en Figura 1. Las curvas fronterizas de arriba, a la izquierda, y a la derecha, corresponden a $f_p(t)$ con raíces de la forma $(\alpha, \bar{\alpha}, \alpha, \bar{\alpha})$, $(\alpha, \bar{\alpha}, 1, 1)$, y $(-1, -1, \alpha, \bar{\alpha})$ respectivamente. Así, los vértices de la izquierda, de la derecha, y de abajo, corresponden a las raíces $(1, 1, 1, 1)$, $(-1, -1, -1, -1)$, y $(-1, -1, 1, 1)$.

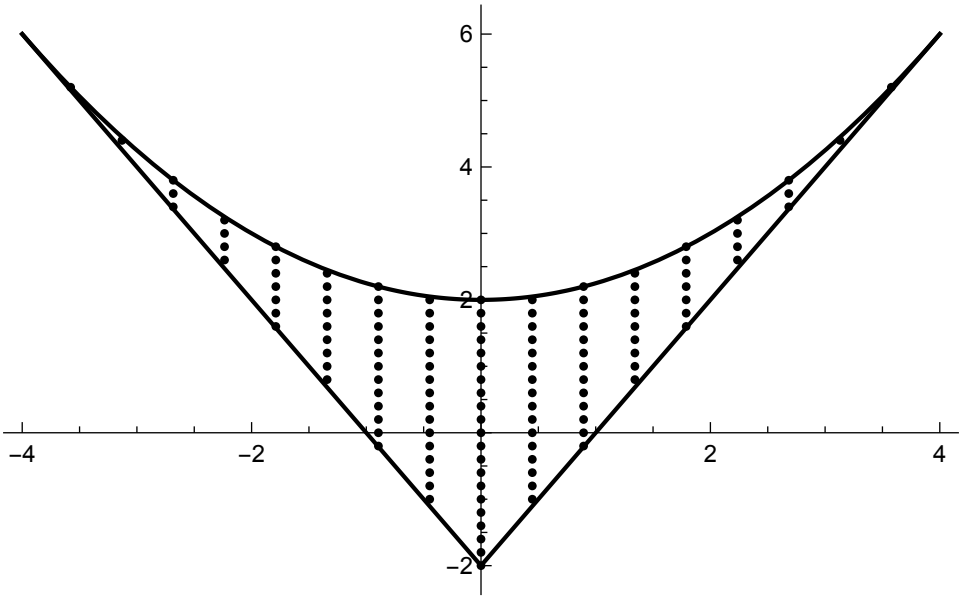



FIGURA 1. El simplex curvilineal $Sp_4^\natural$. Los 129 puntos $\mathrm{fr}_5 = (u_1, u_2) = (a_1/\sqrt{5}, a_2/5)$ corresponden a las 129 funciones $L$ $1 + a_1 5^{-s} + a_2 5^{-2s} + a_1 5^{1-3s} + 5^{2-4s}$ en $\mathcal{L}_2(\mathbb{F}_5)$.

Como una quinta opción, se puede trasladar al lenguaje de *funciones L*, escribiendo

$$L_q(A, s) = F_q(A, q^{-s}).$$

Éste es el punto de vista que destacaremos, escribiendo $\mathcal{L}_g(\mathbb{F}_q)$ para el conjunto de todas las funciones $L$ que surjan de cualquier colección de $2g$ $q$-números de Weil definidos sobre $\mathbb{Q}$ y estables bajo $\alpha \mapsto q/\alpha$.

*Teorema de Honda-Tate.* Este teorema describe completamente el conjunto *a priori* muy complicado $\mathrm{IsAb}(\mathbb{F}_q)$ en términos del conjunto elemental $\mathcal{L}(\mathbb{F}_q)$.

**Teorema 15.1** (Parte del Teorema de Honda-Tate)**.** *El mapeo que envía una clase de isogenía de una variedad abeliana $A \in \mathrm{IsAb}(\mathbb{F}_q)$ a su función $L$ $L(A, s) \in \mathcal{L}(\mathbb{F}_q)$ es inyectivo. Para $q = p$ primo, es también suryectivo.*

Cuando $q$ no es primo, una función $L$ está en la imagen si ciertas obstrucciones se anulan. Estas obstrucciones son extrañas. No entraremos en esta hermosa teoría de la obstrucción porque nuestro objetivo principal es describir un marco simple que sirva de guía para las próximas dos secciones. Tate probó en [37] la parte de la inyectividad del teorema y describió las obstrucciones. Honda probó en [30] para $q$ general que la imagen es en efecto igual a todas las funciones $L$ no obstruidas.

*Volúmenes de espacio de clases.* El Teorema de Honda-Tate hace que sea importante entender bien el conjunto más simple $\mathcal{L}_g(\mathbb{F}_q)$. Como la Figura 1 sugiere, conteos exactos son posibles, y de hecho muchos conteos exactos de $\mathrm{IsAb}(\mathbb{F}_q)$ vía el Teorema de Honda-Tate están en la LMFDB. En un nivel aproximado, los conteos provienen de volúmenes como sigue. El intervalo $Sp_2^\natural = [-2, 2]$, que sirve como espacio ambiental de todos los $\mathcal{L}_1(\mathbb{F}_q)$, obviamente tiene longitud 4. Las curvas fronteras en Figura 1 tienen ecuaciones que se pueden ver en (17.9) abajo, y una integración muestra que el "escudo" $Sp_4^\natural$ conteniendo todo $\mathcal{L}_2(\mathbb{F}_q)$ tiene área 16/3. Un cómputo más sofisticado ([24]) del volumen Euclidiano $V_g$ de $Sp_{2g}^\natural$ para $g$ arbitrario da

$$V_g = \prod_{j=1}^{g} \frac{2^{j+1}(j-1)!}{(2j-1)!!}.$$

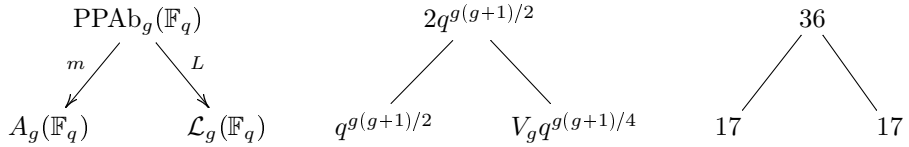
El $j$-ésimo factor es asintótico a $\sqrt{\pi/j}$, por lo que en particular $V_g$ tiende a 0. Reescalando la $j$-ésima coordenada por $q^{j/2}$, obtenemos

$$|\mathcal{L}_g(\mathbb{F}_q)| \approx V_g q^{g(g+1)/4}. \tag{15.7}$$

Una interpretación rigurosa de esta aproximación es que el radio de los dos lados tiene a 1 cuando $q$ va al infinito.

**15.5. El diagrama principal revisado.** Los tres conjuntos en los que nos hemos concentrado en el caso $K = \mathbb{F}_q$ están a la izquierda:

(15.8)




Fórmulas aproximadas para sus tamaños, provenientes de nuestras consideraciones previas, están en el medio. Estas fórmulas muestran que $\mathcal{L}_g(\mathbb{F}_q)$ es mucho más pequeño que los otros dos conjuntos.

El mapeo $L$ es complicado: algunas fibras puedan ser vacías pero la mayoría de las fibras son grandes. Describimos aquí el caso $g = 1$ y $q = p$ primo, siguiendo [38]. Aquí el mapeo $L$ es trivialmente suryectivo pues todas las curvas elípticas vienen con una polarización principal canónica. La descripción está en términos de discriminantes cuadráticos negativos, es decir, discriminantes de ordenes cuadráticos imaginarios. Estos son enteros negativos congruentes a 0 o 1 módulo 4. Tienen una factorización canónica como $D = dc^2$, donde $d$ es el discriminante del cuerpo $\mathbb{Q}(\sqrt{d})$. Los números $d$ son llamados discriminantes fundamentales y son reconocidos entre todos los discriminantes como los únicos libre de cuadrados si $d \equiv 1 \pmod 4$ y 4 veces un entero libre de cuadrados si $d \equiv 0 \pmod 4$.

El número de clase $h(D)$ de un discriminante general se puede expresar en términos del discriminante fundamental asociado $d$:

$$h(dc^2) = h(d)\frac{w(dc^2)}{w(d)} c \prod_{p|c} \left(1 - \left(\frac{d}{p}\right)\frac{1}{p}\right).$$

Aquí, $w(D)$ cuenta raíces de la unidad, así que $w(-3) = 6$, $w(-4) = 4$, y $w(D) = 2$ en caso contrario. Definimos $H(dc^2) = \sum_{j|c} h(dj^2)$. Entonces la fórmula simple es que el tamaño de la fibra arriba $1 + ap^{-s} + p^{1-2s}$ es $H(a^2 - 4p)$.

| $j$ | **0** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | **11** | 12 | 13 | 14 | 15 | 16 | $a$ | $D$ | $H(D)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | 1 | | | | | | 8 | $-4$ | 1 |
| | | | | | | | | | | | | | 1 | | | | | 7 | $-19$ | 1 |
| | | 1 | | | | | 1 | | | | 1 | | | | | | | 6 | $-8 \cdot 2^2$ | $1+2$ |
| | | | | | | | | | | | | | | | | 1 | | 5 | $-43$ | 1 |
| | | | | | 1 | 1 | | | | | | | | | | | | 4 | $-52$ | 2 |
| | | 1 | | | | | | 1 | | | | | | 1 | | | | 3 | $-59$ | 3 |
| | 1 | | | | | | | | 1 | | | | | | 1 | | 1 | 2 | $-4 \cdot 4^2$ | $1+1+2$ |
| | | | 1 | | | | | | | | | | | | | | | 1 | $-67$ | 1 |
| | | | | | | | | 2 | | | 2 | | | | | | | 0 | $-68$ | 4 |
| | | | 1 | | | | | | | | | | | | | | | $-1$ | $-67$ | 1 |
| | 1 | | | | | | | | 1 | | | | | | 1 | | 1 | $-2$ | $-4 \cdot 4^2$ | $1+1+2$ |
| | | 1 | | | | | | 1 | | | | | | 1 | | | | $-3$ | $-59$ | 3 |
| | | | | | 1 | 1 | | | | | | | | | | | | $-4$ | $-52$ | 2 |
| | | | | | | | | | | | | | | | | 1 | | $-5$ | $-43$ | 1 |
| | | 1 | | | | | 1 | | | | 1 | | | | | | | $-6$ | $-8 \cdot 2^2$ | $1+2$ |
| | | | | | | | | | | | | | 1 | | | | | $-7$ | $-19$ | 1 |
| | | | | | | | | | | | | 1 | | | | | | $-8$ | $-4$ | 1 |

CUADRO 1. El número de curvas elípticas sobre $\mathbb{F}_{17}$ con invariante $j$ igual a $j$ y función $L$ $1 + ap^{-s} + p^{1-2s}$. El correspondiente discriminante $D = a^2 - 4p$ y el número de clases $h(D)$ están dados a la derecha.

La parte derecha de (15.8) es el caso del cuerpo base $\mathbb{F}_{17}$. Con mucho más detalle, Cuadro 1 muestra explícitamente cómo las fibras de $L$ son gobernadas por los números de clases, mientras que las fibras de $m$ tiene tamaño 2, excepto quizás sobre los invariantes $j$ excepcionales 0 y 1728. En este caso, 0 aún tiene una fibra de tamaño 2 pues $|\mu_6(\mathbb{F}_{17})| = 2$, pero 1728, visto como 11 en $\mathbb{F}_{17}$, tiene una fibra de tamaño 4, pues $|\mu_4(\mathbb{F}_{17})| = 4$.

**15.6. Grupos de Galois motívicos y grupos de Sato-Tate.** La teoría de Galois juega un papel más grande en esta situación que la que hemos indicado. Concluimos la primera sección de esta parte definiendo grupos de Galois motívicos y grupos de Sato-Tate de variedades abelianas sobre cuerpos finitos. Tal como $\mathrm{Gal}(\overline{\mathbb{F}}_q/\mathbb{F}_q)$, ellos son cíclicos en un sentido apropiado, siendo generados por un elemento de Frobenius. En contraste, los grupos de Galois motívicos y los grupos de Sato-Tate de las secciones siguientes, como $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, estarán lejos de ser abelianos. Nuestro propósito principal con esta subsección es dar alguna idea de lo que son estos grupos, antes de ingresar al más sofisticado marco de la próxima sección.

*Grupos de números de Weil.* Dada $A \in \mathrm{IsAb}(\mathbb{F}_q)$, sea $\Pi$ el subgrupo de $\mathbb{C}^\times$ generado por sus $q$-números de Weil $\alpha$. De manera similar, sea $\Theta$ el subgrupo del círculo unitario generado por los números normalizados de Weil $\alpha/\sqrt{q}$. El grupo $\Pi$ es el *grupo de números de Weil* de $A$ y el grupo $\Theta$ es el *grupo de ángulos* de $A$. Claramente, $\Pi$ y $\Theta$ son versiones similares una de otra.

Entre las diferentes cosas contabilizadas por la LMFDB para una clase de isogenía es su *rango angular* $r$, que significa el rango del grupo abeliano finitamente generado

$\Theta$. Como los números $\alpha/\sqrt{q}$ vienen en pares uno inverso del otro, este rango está en $\{0, \ldots, g\}$. El rango de $\Pi$ es $r+1$. Sea $t$ el tamaño del subgrupo de torsión de $\Pi$. El tamaño del subgrupo de torsión de $\Theta$ es generalmente $t$, pero excepcionalmente puede ser $2t$. Este último caso se da cuando por ejemplo $q = p^2$ por $p \in \{2, 3\}$ y $F_q(T) = 1 + pT + q$, donde $t = 2p$.

*Rango angular en dimensión* 2. Como un ejemplo que será de ayuda después, tomamos $g = 2$ y $q = p \geqslant 7$. Entonces, existen siempre exactamente cinco polinomios de Frobenius de rango cero. Sus versiones normalizadas $f_p(t)$ son $1 + bt^2 + t^4$ con $b \in \{-2, -1, 0, 1, 2\}$. Los polinomios $f_p(t)$ son productos de polinomios ciclotómicos. Por ejemplo, cuando $b = -1$, $1 - t^2 + t^4 = (1 + t + t^2)(1 - t + t^2) = \Phi_3(t)\Phi_6(t)$.

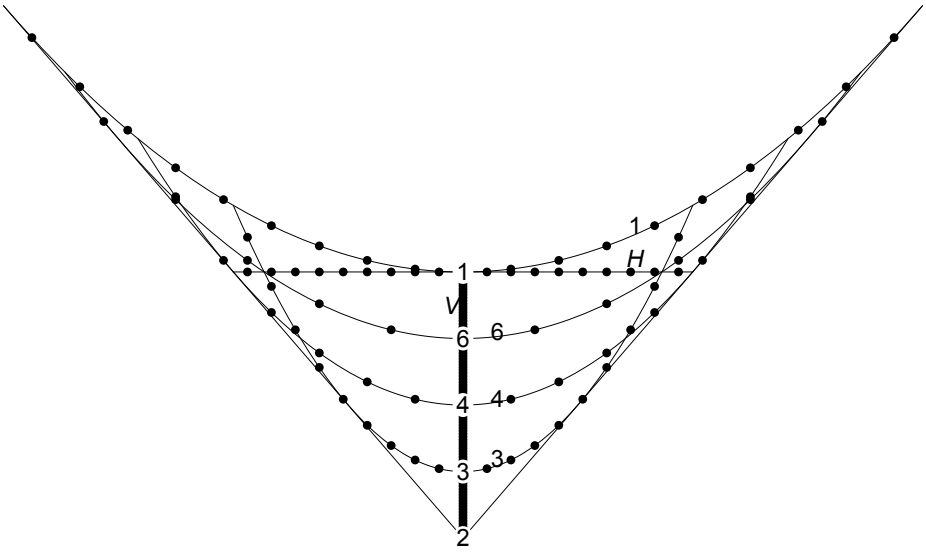



FIGURA 2. El espacio de clases $Sp_4^\natural$ de Figura 1, ahora con cinco puntos etiquetados $p_i$ correspondientes al rango angular cero y seis curvas etiquetadas correspondientes a rango angular uno. Los 164 puntos de $\mathcal{L}_2(\mathbb{F}_{23})$ que tienen rango angular uno están también dibujados, con 88 de ellos en la línea vertical $V$.

El comportamiento es nuevamente uniforme con respecto al rango angular uno: los puntos $(u, v) \in Sp_4^\natural$ indexando los polinomios unitarizados $f_p(t) = 1 + ut + vt^2 + ut^3 + t^4$ viven todos en seis curvas. Estas curvas son cuatro parábolas $C_k$, una línea vertical $V_2$ y una línea horizontal $H_4$. Los subíndices dan el número de torsión. Figura 2 dibuja y etiqueta cada una de las seis curvas. Coloca un índice $k$ para indicar un punto $p_k$ donde el rango de ángulo es cero. La notación de los puntos se hereda de la de las curvas, y también tiene la propiedad que $p_k = (0, 4\cos^2(\pi/k) - 2)$.

Más detalles son dados en Cuadro 2. Sobre la derecha $f_p(t)$ es dado si se factoriza en factores de menor grado. Esta factorización muestra cómo puntos genéricos en $C_1$ y $H_4$ tienen de hecho rango uno.

| Punto | $f_p(t)$ |
|---|---|
| $p_2 = (0,-2)$ | $(1-t)^2\,(1+t)^2$ |
| $p_3 = (0,-1)$ | $1-t^2+t^4$ |
| $p_4 = (0,0)$ | $1+t^4$ |
| $p_6 = (0,1)$ | $\Phi_3(t)\,\Phi_6(t)$ |
| $p_1 = (0,2)$ | $(1+t^2)^2$ |

| Curva | $f_p(t)$ | $g_p(t)$ |
|---|---|---|
| $V_2:\ u=0$ | | $(t+1)^2\,(t^2-vt+1)$ |
| $C_3:\ v=u^2-1$ | | $(t^2+t+1)\ \ (-tu^2+t^2+2t+1)$ |
| $C_4:\ v=\frac{u^2}{2}$ | | $(t^2+1)\ \ (-\frac{tu^2}{2}+t^2+2t+1)$ |
| $C_6:\ v=\frac{u^2}{3}+1$ | | $(t^2-t+1)\ \ (-\frac{tu^2}{3}+t^2+2t+1)$ |
| $C_1:\ v=\frac{u^2}{4}+2$ | $(t^2+\frac{ut}{2}+1)^2$ | $(1-t)^2\,(-\frac{tu^2}{4}+t^2+2t+1)$ |
| $H_4:\ v=2$ | $(t^2+1)(t^2+ut+1)$ | |

CUADRO 2. Información de los cinco puntos correspondientes a rango angular cero y las seis curvas correspondientes a rango angular uno.

Para ver lo especial de las otras curvas, sean $\alpha$, $\beta$, $\bar{\alpha}$, y $\bar{\beta}$ las raíces de $f_p(t)$. Entonces

$$\begin{aligned} g_p(t) &:= (1-\alpha\beta t)(1-\bar{\alpha}\beta t)(1-\alpha\bar{\beta}t)(1-\bar{\alpha}\bar{\beta}t) \\ &= 1+(2-v)t+(2+u^2-2v)t^2+(2-v)t^3+t^4. \end{aligned}$$

Cuando usamos la ecuación de la curva para remover la variable, entonces $g_p(t)$ se factoriza en los casos listados en Cuadro 2. Nuevamente estas factorizaciones muestran que los puntos genéricos sobre las curvas restantes tienen rango uno. Las factorizaciones también muestran que los números de torsión dados como subíndices son correctos.

*Definiciones vía dualidad.* Sea $A$ una variedad abeliana con grupo de Weil $\Pi$ y grupo angular $\Theta$. Sea $r$ el rango angular de $\Theta$ y sea $\delta t$ el número de torsión de $\Theta$ como arriba, por lo que $\Pi$ tiene rango $r+1$ y número de torsión $t$.

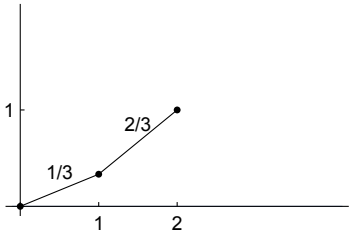
Sea $ST$ el grupo dual de $\Theta$. Aquí vemos $\Theta$ como un grupo discreto tal que $ST$ es compacto. Su componente de la identidad $ST^0$ es el producto de $r$ círculos. El grupo $ST/ST^0$ es isomorfo a $\mathbb{Z}/(\delta t)$. El grupo $ST$ es el *grupo de Sato-Tate* de $A$.

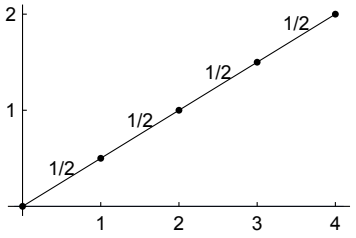
Para $\Pi$ procedemos de manera similar. Sin embargo, esta vez, prestamos atención a la acción natural de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Sea $G$ el grupo dual de $\Pi$, ahora considerado en el marco de grupos algebraicos conmutativos sobre $\mathbb{Q}$. La componente identidad $G^0$ satisface $G^0(\mathbb{C}) \cong (\mathbb{C}^\times)^{r+1}$. Su grupo de componentes $Q = G/G^0$ satisface $Q(\mathbb{C}) = \mathbb{Z}/t$. El grupo $G$ es el *grupo de Galois motívico* de $A$. El hecho que ambos grupos, $\Pi$ y $\Theta$, están dentro de $\mathbb{C}^\times$ provee a los grupos recién definidos generadores canónicos, los cuales son denotados $\mathrm{Fr}_q \in G(\mathbb{Q})$ y $\mathrm{fr}_q \in ST$.

**15.7. Ejercicios.**

1. Verificar que la curva elíptica (15.5) realmente tiene invariante $j$ igual a $j$.
2. Construir el cuadro análogo a Cuadro 1 para $p = 5$.
3. Explorar la sección de la LMFDB sobre variedades abelianas sobre $\mathbb{F}_q$. Algunos posibles temas son:
   - ¿Cuán común es para $A$ no ser una Jacobiana porque la "curva correspondiente" tendría un número negativo de puntos?
   - ¿Cuántos polígonos de Newton pueden ocurrir para un $g$ dado y cuáles son sus frecuencias relativas aproximadas?
   - ¿Qué porcentaje de las clases de isogenía en la página $(g, q)$ es primitivo?
4. Los puntos $(u_1, \ldots, u_g)$ en $Sp^\natural_{2g}$ correspondientes al rango angular 0 son aquellos con coordenadas enteras. Usar la factorización de polinomios de Frobenius unitarizados $f_p(t) = 1 + u_1 t + \cdots + u_1 t^{2g-1} + t^{2g}$ en polinomios ciclotómicos para contar el número $N_g$ de tales puntos vía funciones generatrices. (Se necesita considerar $\Phi_1(t) = t - 1$ y $\Phi_2(t) = t + 1$ de manera diferente que los otros $\Phi_k(t)$, y tu respuesta debería dar $N_{10} = 20399$ como un caso especial.)
5. En la página de la LMFDB para superficies abelianas simples sobre cuerpos primos $\mathbb{F}_p$, encontrarás exactamente una clase de isogenía tal que el polinomio de Frobenius es reducible. ¿Qué es esto? Para una variedad abeliana de otras dimensiones sobre cuerpos primos $\mathbb{F}_p$, no encontrarás ningún polinomio irreducible. Explica cómo esto sigue del Teorema 15.1.
6. Lee en la literatura sobre el Teorema de Honda-Tate sobre $\mathbb{F}_q$ general. Da una segunda explicación para el fenómeno del ejercicio anterior en términos de obstrucciones reales. Considera también los siguientes fenómenos que son visibles en la LMFDB.
   - Los polinomios $1 \pm 2T + 8T^2$ no están en la página para curvas elípticas sobre $\mathbb{F}_8$. Sin embargo, $(1 \pm 2T + 8T^2)^3$ aparecen en la página de 3-variedades abelianas sobre $\mathbb{F}_8$ como los polinomios de Frobenius de variedades simples. Todas los demás 6458 polinomios de grado seis para variedades abelianas simples son irreducibles.
   - Los polinomios $1 + pT + p^2T^2 + p^3T^3 + p^4T^4$ se ven en las páginas para superficies abelianas sobre $\mathbb{F}_{p^2}$ para $p \leqslant 7$, pero no en la página para superficies abelianas sobre $\mathbb{F}_{11^2}$.

   Explica estos dos fenómenos en términos de obstrucciones $p$-ádicas. Tu explicación debe hacer referencia a los siguientes polígonos de Newton, donde los números son las subidas verticales de los segmentos.




7. El grupo de Galois $G$ de un polinomio conformalmente palíndromo $F_q(T) = 1 + a_1 T + \cdots + a_1 q^{g-1} T^{2g-1} + q^g T^{2g}$ está en $2^g.S_g$, el grupo de orden $2^g g!$ que consiste en permutaciones de las raíces los cuales conmutan con la involución

$\alpha \mapsto q/\alpha$ sobre las raíces. Probar que si $g \geqslant 2$ y $G = 2^g.S_g$, entonces el rango angular de $F_q(T)$ es $g$.

## 16. Variedades abelianas sobre $\mathbb{Q}$: generalidades ilustradas por curvas elípticas

Esta segunda sección de la tercera parte discute invariantes y la clasificación de variedades abelianas sobre $\mathbb{Q}$. Mostraremos el marco teórico para $g$ arbitrario, pero centrándonos en el escenario relativamente familiar de $g = 1$. En particular, recalcaremos tres conjeturas de la década de 1960 para $g$ general, las cuales están completamente resueltas únicamente para $g = 1$. Estas conjeturas y algunas otras abordan la cuestión de por qué uno querría tabular minuciosamente las variedades abelianas principalmente polarizadas: su aritmética es extremadamente rica.

Como ejemplos explícitos, tomamos

$$E_1 : y^2 = x^3 - x, \qquad \Delta_1 = 4 \qquad = 2^2, \qquad j_1 = 1728 \ = 2^6 3^3,$$

$$E_2 : y^2 = x^3 + 6x - 7, \qquad \Delta_2 = -2187 = 3^7, \qquad j_2 = \frac{2048}{3} = \frac{2^{11}}{3}.$$

La curva $E_1$ tiene un automorfismo extra, $(x, y) \mapsto (-x, iy)$, definido sobre $\mathbb{Q}(i)$. En otras palabras, tiene multiplicación compleja potencial. Veremos de distintas formas que $E_2$ no tiene multiplicación compleja potencial; en otras palabras, es genérica.

**16.1. Reducción buena versus reducción mala.** Sea $A/\mathbb{Q}$ una variedad abeliana. Para $p$ un primo, sea $\mathbb{Z}_{(p)}$ el anillo de números racionales con denominador coprimo a $p$. Entonces, se dice que $A$ tiene *reducción buena* en $p$ si existe un esquema abeliano $\underline{A}$ sobre $\mathbb{Z}_{(p)}$ con fibra genérica $A$. En caso contrario, se dice que $A$ tiene *reducción mala* en $p$. El conjunto $S$ de los primos malos es un invariante de isogenía.

Puede ser difícil identificar el conjunto $S$ de primos malos de una $A/\mathbb{Q}$ dada, pero es usualmente fácil dar una cota superior razonable $S'$ para él. Por ejemplo supongamos que $A$ es la Jacobiana de $y^2 = f(x)$ con $f(x)$ un polinomio mónico en $\mathbb{Z}[x]$. Entonces uno puede tomar $S'$ como 2 y los primos en los cuales $f(x)$ tiene un factor irreducible repetido cuando es reducido a $\mathbb{F}_p[x]$. Estos últimos primos son exactamente aquellos que dividen al discriminante $\Delta$ de $f(x)$. Luego, en nuestros ejemplos, $S'_1 = \{2\}$ y $S'_2 = \{2, 3\}$.

Un invariante fundamental de $A \in \mathrm{IsAb}_g(\mathbb{Q})$, más refinado que $S$, es el entero positivo $N = \prod_p p^{c_p}$ llamado *conductor* de $A$. Los factores primos $p$ de $N$ son los primos de reducción mala de $A$. El exponente $c_p$ en un primo $p$ mide la naturaleza de la reducción mala: a mayor $c_p$, peor es la reducción. La magnitud de $N$ es una medida importante de la complejidad aritmética de $A$.

En nuestros casos,

$$N_1 = 32 = 2^5, \qquad N_2 = 72 = 2^3 3^2.$$

El factor $3^2$ será explicado vía representaciones módulo 2 al final de §16.10. De manera similar, los factores $2^c$ serán explicados allí vía representaciones módulo 3.

**16.2. Estrategia de clasificación.** El punto de vista más usado para la aritmética es el de concentrarse primeramente en las clases de isogenía, y en las variedades abelianas dentro de una clase de isogenía en segundo lugar. Uno ordena las clases de isogenía con respecto al valor del conductor. Los invariantes geométricos juegan un papel secundario.

Para $g = 1$, las tablas clásicas ([19]) de Swinnerton-Dyer et al., publicadas en 1975, consideraron todos los casos hasta la cota 200 para el conductor. En el libro de Cremona de 1992 ([22]) se incrementó esta cota a 1000. La base de datos de Cremona actualmente llega hasta 400,000, y está en la LMFDB. Existen $1,741,002$ clases de isogenía y $2,483,649$ curvas, alrededor de 1.43 curvas por clase de isogenía.

La lista comienza a la izquierda de la tabla de abajo.

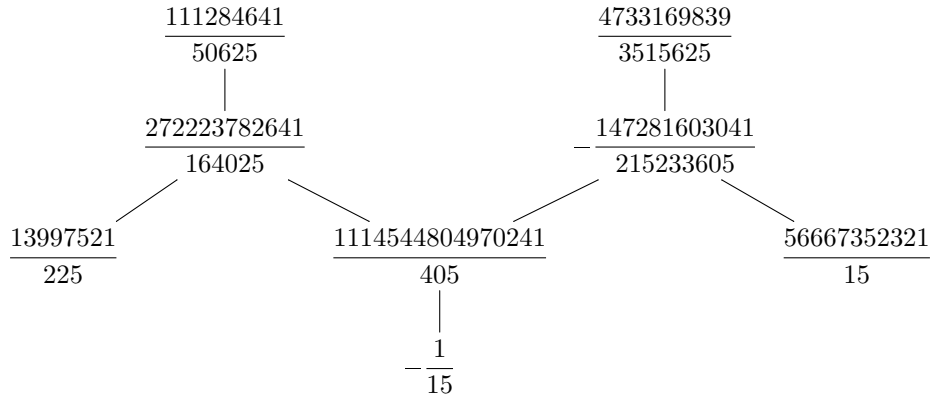
| Conductor $N$ | Número de curva elípticas | Número encontrado por una búsqueda muy rápida |
|---|---|---|
| 11 | 3 | 1 |
| 14 | 6 | 1 |
| 15 | 8 | 1 |
| 17 | 4 | 1 |
| 19 | 3 | 1 |
| 20 | 4 | 1 |
| 21 | 6 | 1 |
| 24 | 6 | 1 |

En este rango, una clase de isogenía es determinada por su conductor, aunque ya para $N = 26$ existen dos clases de isogenía. También en este rango curvas diferentes dentro de una misma clase de isogenía tiene diferentes invariantes $j$, aunque para $N = 27$ hay dos curvas isógenas con el mismo invariante $j$.

Por diversión, los resultados de una muy corta búsqueda son presentados en la última columna. La búsqueda consideró curvas elípticas en la "forma larga" estándar

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

con $a_k \in \{-1, 0, 1\}$ para todo $k$. Encontramos exactamente una curva por cada uno de los primeras ocho clases de isogenía. El caso de $N = 15$, con los ocho invariantes $j$ conectados por 2-isogenías, es




La búsqueda encontró solo la curva con invariante $j$ igual a $-1/15$. El diagrama ilustra que el tamaño de $N$ y la altura de $j$ están muy débilmente relacionadas, así que es difícil encontrar todas las curvas elípticas de conductor pequeño buscando por las ecuaciones. La lista de Cremona fue calculada por el método modular de Teorema 16.8.

**16.3. Funciones $L$ como series de Dirichlet definidas por productos de Euler.** Dada $A/\mathbb{Q}$ con reducción mala dentro de $S'$, uno tiene inmediatamente

infinitos invariantes, los factores locales $L_p(A,s) = F_p(A,p^{-s})$ de la sección 15 para cualquier $p$ que no está en $S'$. Estos son los correspondientes polinomios de Frobenius para nuestras dos curvas:

| $p$ | $F_p(E_1,T)$ | $F_p(E_2,T)$ |
|---|---|---|
| 2 | **1** | **1** |
| 3 | $1+3T^2$ | **1** |
| 5 | $1+2T+5T^2$ | $1-2T+5T^2$ |
| 7 | $1+7T^2$ | $1+7T^2$ |
| 11 | $1+11T^2$ | $1+4T+11T^2$ |
| 13 | $1-6T+13T^2$ | $1+2T+13T^2$ |
| 17 | $1-2T+17T^2$ | $1+2T+17T^2$ |
| 19 | $1+19T^2$ | $1+4T+19T^2$ |
| 23 | $1+23T^2$ | $1-8T+23T^2$ |
| 29 | $1+10T+29T^2$ | $1+6T+29T^2$ |

Como lo indican los primeros tres símbolos **1** en la tabla, para $p \in S'$ existen también polinomios $F_p(A,T)$ bien definidos, que dan un factor local $L_p(A,s)$ por la misma substitución $T = p^{-s}$. Tiene grado $\leqslant 2g$ con desigualdad estricta exactamente cuando $p \in S$.

La función $L$ asociada a $A$ es

$$L(A,s) = \prod_p \frac{1}{L_p(A,s)} = \sum_{n=1}^{\infty} \frac{a_n}{n^s}. \tag{16.1}$$

(Notar que la inversión es forzada por el requerimiento de que $L(A,s)$ es la función $L$ estándar sobre $\mathbb{Q}$ y $L_p(A,s)$ es la función $L$ estándar sobre $\mathbb{F}_p$. En la literatura enfocada únicamente en funciones $L$ sobre $\mathbb{Q}$, uno usualmente encuentra que $L_p(A,s)$ significa lo que nosotros llamamos $1/L_p(A,s)$.)

El producto y la suma en (16.1) convergen absolutamente en el semi-plano derecho $\Re(s) > 3/2$. Por el momento, asuntos analíticos no jugarán ningún papel, y uno puede considerar $L(A,s)$ como un paquete formal de cantidades $L_p(A,s)$, permitiendo las modificaciones en un número finito de primos malos que describiremos en §16.9. Sea $\mathcal{L}_g(\mathbb{Q})$ el conjunto de todos estos productos formales.

Faltings [27] generalizó la parte de la inyectividad en el Teorema de Honda-Tate, de tal manera que $\mathrm{IsAb}_g(\mathbb{Q}) \to \mathcal{L}_g(\mathbb{Q}) : A \mapsto L(A,s)$ es inyectiva. Luego, el problema fundamental es caracterizar el conjunto numerable de la imagen dentro del dominio no numerable. En otras palabras, cuáles relaciones debe satisfacer una sucesión de polinomios para que aparezcan como una sucesión $F_p(A,T)$. Las tres conjeturas de abajo dan condiciones que se esperan que sean necesarias. Tal como lo veremos, se espera que la última condición esté cerca de ser suficiente.

**16.4. Anillos de endomorfismos.** ¡No todas las clases de isogenía de variedades abelianas son creadas iguales! Uno de los propósitos de los grupos de Galois motívicos $G$, y de sus variantes fáciles, los grupos de Sato-Tate $ST$, es hacer distinciones cualitativas entre clases de isogenía. Un principio simple es, *mientras más grande sea el grupo, más difícil será la aritmética.* Como un preludio de $G$ y $ST$, discutiremos anillos de endomorfismos.

Una variedad abeliana $A$ sobre un cuerpo $K$ tiene un anillo de endomorfismos $\mathrm{End}(A)$ y un anillo de endomorfismos geométricos $\mathrm{End}(A_{\overline{K}}) \supseteq \mathrm{End}(A)$. Para todo

anillo de endomorfismos geométricos posible $R$, existe una correspondiente subvariedad $X_R$ de $A_g$. Sus puntos complejos $X_R(\mathbb{C})$ por definición son la clausura del conjunto de los elementos $x$ de $\mathrm{End}(A_x)$ isomorfos a $R$.

Para las curvas elípticas $E$ sobre $\mathbb{Q}$, el anillo de endomorfismos $\mathrm{End}(E)$ es siempre $\mathbb{Z}$. Mientras que $\mathrm{End}(E_{\overline{\mathbb{Q}}})$ es genéricamente $\mathbb{Z}$, también puede ser un anillo cuadrático $R$ con discriminante negativo $D$. En este caso, se dice que $E$ tiene multiplicación compleja potencial por $\mathbb{Q}(\sqrt{D})$. La subvariedad $X_R$ es irreducible de grado $h(D)$, lo que significa que $X_R(\mathbb{C})$ contiene $h(D)$ puntos, todos conjugados por $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.

Los siguientes casos donde $h(D) = 1$ son famosos:

| $D$ | $j_D$ |
|---|---|
| $-3$ | $0$ |
| $-4$ | $1728$ |
| $-7$ | $-3375$ |
| $-8$ | $8000$ |
| $-11$ | $-32768$ |
| $-12 = -3 \cdot 2^2$ | $54000$ |
| $-16 = -4 \cdot 2^2$ | $287496$ |
| $-19$ | $-884736$ |
| $-27 = -3 \cdot 3^2$ | $-12288000$ |
| $-28 = -7 \cdot 2^2$ | $16581375$ |
| $-43$ | $-884736000$ |
| $-67$ | $-147197952000$ |
| $-163$ | $-262537412640768000$ |

Ya que los invariantes $j$ son $j_1 = 1728$ y $j_2 = 2048/3$, la curva $E_1$ tiene multiplicación compleja potencial por $\mathbb{Q}(i)$ mientras que $E_2$ es genérica.

**16.5. Grupos de Galois motívicos.** Asociado a una variedad abeliana $A$ sobre un subcuerpo $K$ de $\mathbb{C}$ está su grupo de Galois motívico $G$. Esto es un subgrupo del grupo simpléctico conforme $\mathrm{GSp}_{2g}$. Existen un número de definiciones competentes para $G$, las cuales no se sabe si son equivalentes en general. Nosotros tomamos la de Parte I de [23], donde requiere que $G$ fije los "ciclos de Hodge absolutos" en su acción natural $H^1(A(\mathbb{C}), \mathbb{Q})^{\otimes 2j} \otimes \mathbb{Q}(j)$ donde $\mathbb{Q}(j)$ indica un "giro de Tate".

Omitiremos la definición completa de estos $G$, ya que tres propiedades de ellos son un substituto adecuado para estas notas. Primeramente, $G$ siempre conmuta con $\mathrm{End}(A)$. En segundo lugar, la componente de la identidad $G^0$, también conocida como el grupo de Mumford-Tate, siempre conmuta con $\mathrm{End}(A_{\mathbb{C}})$. Finalmente, para $g \leqslant 3$, $G^0$ es siempre igual al conmutador completo en $\mathrm{GSp}_{2g}$ de $\mathrm{End}(A_{\mathbb{C}})$.

En el caso $g = 1$, el grupo simpléctico conforme $\mathrm{GSp}_2$ no es más que otro nombre para $GL_2$, el cual es bien conocido por jugar un papel central en la teoría de curvas elípticas. El siguiente gráfico determina $G$:

| | $\mathrm{End}(E)_{\mathbb{Q}}$ | $\mathrm{End}(E_{\mathbb{C}})_{\mathbb{Q}}$ | $G(\mathbb{Q})$ |
|---|---|---|---|
| Genérico: | $\mathbb{Q}$ | $\mathbb{Q}$ | $GL_2(\mathbb{Q})$ |
| CM potencial : | $\mathbb{Q}$ | $F$ | $N(F^\times)$ |
| CM: | $F$ | $F$ | $F^\times$ |

Luego en el caso CM, $G$ es un toro de dimensión dos. En el caso CM potencial, es el normalizador de este toro y por lo tanto tiene dos componentes.

**16.6. Restricciones en los polinomios de Frobenius.** Sea $A$ una variedad abeliana sobre $\mathbb{Q}$. Su grupo de Galois motívico $G$ actúa sobre sí mismo por conjugación y el espectro del anillo de funciones invariantes es su variedad de clases $G^\natural$. Para el mismo $\mathrm{GSp}_{2g}$, la variedad de clases puede ser identificada con el conjunto de polinomios conformalmente palíndromos de grado $2g$, donde el factor conforme está dado. Los polinomios de Frobenius necesariamente viven en la imagen de $G^\natural(\mathbb{Q})$ en $\mathrm{GSp}_{2g}^\natural(\mathbb{Q})$. Cuando $G$ es estrictamente más pequeño que $\mathrm{GSp}_{2g}$, se reduce drásticamente el conjunto de posibles polinomios de Frobenius para cualquier primo dado.

En el caso de curvas elípticas sobre $\mathbb{Q}$, las restricciones para los polinomios de Frobenius de curvas elípticas con CM potencial por $D$ son como siguen. Primeramente, si $(D/p) = -1$ entonces $F_p(T) = 1 + pT^2$, tal como está ilustrado cinco veces por $E_1$ arriba. En segundo lugar, para $(D/p) = 1$, el discriminante de $F_p(T)$ debe ser $D$ multiplicado por un cuadrado.

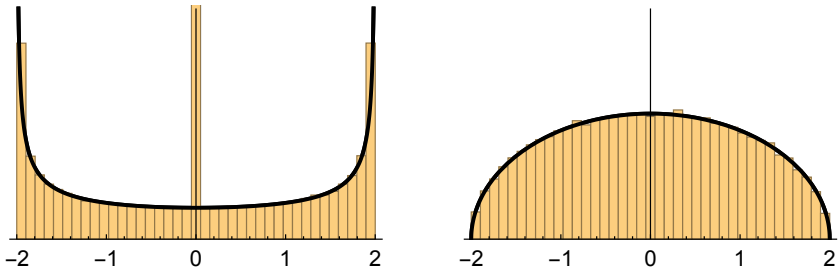
Para probar que una curva elíptica $E$ sobre $\mathbb{Q}$ no tiene CM potencial, no se debe usar el invariante $j$. Sólo se necesita mostrar que las condiciones mencionadas no son satisfechas. Por ejemplo, 5 y 13 son los primos menores $p$ tales que $F_p(E_2, T)$ tiene un término lineal no nulo. Sus discriminantes módulo cuadrados son $d_5 = -1$ y $d_{11} = -7$. El hecho que $-1 \neq -7$ implica que $G = GL_2$. La Proposición 17.1 de abajo explica cómo este simple cálculo tiene su análogo para $g \geqslant 2$.

**16.7. Equidistribución arquimediana.** La intersección de $G$ con $Sp_{2g}$ tiene una forma real compacta $ST$ llamada el grupo de Sato-Tate de $A$. Como en §15.6, se puede pensar a $ST$ como una versión no aritmética del grupo de Galois motívico: los giros de Tate han sido eliminados y el marco refinado de grupos reductivos ha sido reemplazado por los grupos compactos que son más familiares. Para curvas elípticas sobre $\mathbb{Q}$, existen sólo dos posibilidades para $ST$. El grupo $ST$ es $Sp_2$ si $E$ no tiene potencial CM. Es el normalizador $U_1.2$ de un toro $U_1$ en caso que sí lo tenga.

El grupo de Sato-Tate $ST$ tiene una medida de probabilidad de Haar, la cual induce una medida de probabilidad $\mu_{ST}$ sobre el espacio de polinomios palíndromos $Sp_{2g}^\natural$. Para curvas elípticas $E$ sobre $\mathbb{Q}$ las medidas en el $u$-intervalo $Sp_2^\natural = [-2, 2]$ para las dos posibilidades son las siguientes:

$$\mu_{U_1.2} = \frac{1}{2}\delta_0 + \frac{1}{2\pi\sqrt{4-u^2}}du, \qquad \mu_{Sp_2} = \frac{\sqrt{4-u^2}}{2\pi}du. \tag{16.2}$$

Los siguientes gráficos consideran nuestros dos ejemplos, ubicando las primeras 100,000 trazas buenas de Frobenius en 39 compartimientos del mismo ancho. La barra del medio en el dibujo de la izquierda ha sido cortada, ya que debería ser nueve veces más alta. El acuerdo con las medidas de (16.2) es visualmente evidente.



En los primeros años de la década del 1960, Sato y Tate conjeturaron lo siguiente para el caso de curvas elípticas, con Sato inspirado por los datos que recién presentamos. Poco después, la siguiente conjetura general era de esperar, módulo el hecho que una definición rigurosa del grupo $ST$ aún no había sido realizada.

**Conjetura 16.1** (Conjetura de Sato-Tate)**.** *Los polinomios buenos de Frobenius $F_p(A,T)$, considerados como puntos en $Sp_{2g}^\natural$, están equidistribuidos con respecto a $\mu_{ST}$.*

Una razón inicial para creer en esta conjetura fue que Deligne había probado un análogo con el cuerpo base $\mathbb{Q}$ reemplazado por $\mathbb{F}_p(t)$. También muchas personas habían encontrado evidencias numéricas para muchos $ST$ sobre $\mathbb{Q}$, el cual es uno de los tópicos de la siguiente sección. El hecho de que la Conjetura 16.1 parezca verdadera es quizás la forma más rápida de ver la importancia de grupos de Galois motívicos.

La conjetura ya era conocida en los años sesenta para curvas elípticas con CM. El caso $g = 1$ fue probado completamente en una sucesión de artículos hace diez años, comenzando con [21].

**Teorema 16.2** (Taylor et al)**.** *La conjetura de Sato-Tate es verdadera para $g = 1$.*

La extensa demostración de este teorema usa propiedades analíticas de no solo $L(E,s)$, sino también de las funciones $L(\mathrm{Sym}^k E, s)$ relacionadas a potencias simétricas.

**16.8. Representaciones de Galois y equidistribución $\ell$-ádica.** Sea $\ell$ un número primo. Entonces $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ actúa en $H^1(A(\mathbb{C}), \mathbb{Q}_\ell)$ vía la teoría de cohomología de étale. Como los ciclos de Hodge se comportan como ciclos algebraicos para variedades abelianas, lo cual fue probado por Deligne en la primera parte de [23], la imagen vive en $G(\mathbb{Q}_\ell)$.

Llevando la medida de probabilidad de Haar de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ hacia $\mathrm{GSp}_{2g}^\natural(\mathbb{Z}_\ell)$ da una medida $\mu_\ell$. El Teorema de Densidad de Chebotarev nos dice que los polinomios característicos están definitivamente equidistribuidos con respecto a esta medida. Este hecho es obviamente un modelo para la conjetura general de Sato-Tate. Aunque en un sentido diferente, la situación $\ell$-ádica es más complicada que la situación arquimediana, pues existen muchas posibilidades para la imagen $K_\ell$ de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ en $G(\mathbb{Q}_\ell)$, y entonces muchas posibilidades para $\mu_\ell$.

Otra conjetura que resalta la importancia fundamental esperada de los grupos de Galois motívicos es la conjetura de la imagen abierta.

**Conjetura 16.3** (Conjetura de la imagen abierta)**.** *Sea $A$ una variedad abeliana sobre $\mathbb{Q}$ con grupo de Galois motívico $G$. Entonces, para todo número primo $\ell$, la imagen $K_\ell$ de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ es un subgrupo abierto de $G(\mathbb{Q}_\ell)$.*

Con la idea de ser menos abstractos en el caso $G = \mathrm{GSp}_{2g}$, la conjetura dice que la imagen tiene índice finito en el grupo $\mathrm{GSp}_{2g}(\mathbb{Z}_\ell)$ de puntos enteros.

De las tres conjeturas que estamos destacando, la actual es la que está establecida con mayor generalidad.

**Teorema 16.4** (Serre et al.)**.** *La Conjetura de la imagen abierta es verdadera para $g = 1$. Es también cierta si $\mathrm{End}(A_\mathbb{C}) = \mathbb{Z}$ y $g$ es impar.*

El primer enunciado fue probado en 1972 por el artículo más citado de Serre [35]. Para el segundo, la hipótesis implica que $G = \mathrm{GSp}_{2g}$.

En el resto de esta subsección, damos una idea de cómo se ve en términos computacionales. Consideramos únicamente representaciones mód $\ell$. Este es el primer y más importante paso para el caso $\ell$-ádico completo. Estas representaciones mód $\ell$ provienen de las acciones de Galois en $H^1(A(\mathbb{C}), \mathbb{F}_\ell)$. Si $A$ varía en una clase de isogenía, estas representaciones pueden cambiar. Sin embargo, sus semisimplificaciones son todas iguales.

Para curvas elípticas, ésto puede hacerse explícitamente para cualquier $\ell$ de manera uniforme. Nosotros tratamos aquí solo el caso $\ell = 2$ y 3, con Figura 3 dándonos una guía.

*Mód 2.* Las representaciones mód 2 de una curva elíptica $y^2 = x^3 + bx + c$ depende del polinomio cúbico $x^3 + bx + c$ vía $GL_2(\mathbb{F}_2) = S_3$. Particiones de factorización $\lambda_p$ y trazas $a_p$ son coordenadas como en las dos columnas de la izquierda.

(16.3)

| $\lambda_p$ | $a_p$ | masas genéricas | # para $E_1$ | # para $E_2$ |
|---|---|---|---|---|
| 3 | 1 | 1/3 | | |
| 2 1 | 0 | 1/2 | | 50038 |
| $1^3$ | 0 | 1/6 | 100000 | 49962 |

Cuando $x^3 + bx + c$ es irreducible con grupo de Galois $S_3$, la distribución del par $(\lambda, a_p)$ entre las tres posibilidades depende de las masas en las columnas del medio. Ninguno de nuestros ejemplos se ajusta a este patrón, porque los polinomios $x^3 + bx + c$ son reducibles:

$$x^3 - x = (x + 1)x(x - 1), \qquad x^3 + 6x - 7 = (x - 1)(x^2 + x + 7).$$

Las masas que gobiernan estos dos casos no son $(1/3, 1/2, 1/6)$ sino $(0, 0, 1)$ y $(0, 1/2, 1/2)$. En el ejemplo, las representaciones mód 2 son diferentes, aunque sus semisimplificaciones son la misma, lo que significa que $a_p$ es siempre par. La curva de dos componentes $E_1(\mathbb{R})$ está graficada en Figura 3 y los tres puntos de 2-torsión están sobre el eje real, con $x = -1$, 0, y 1. Para una curva con una componente, como $E_2(\mathbb{R})$, existe exactamente un punto real de 2-torsión, el cual en el caso de $E_2$ es racional.

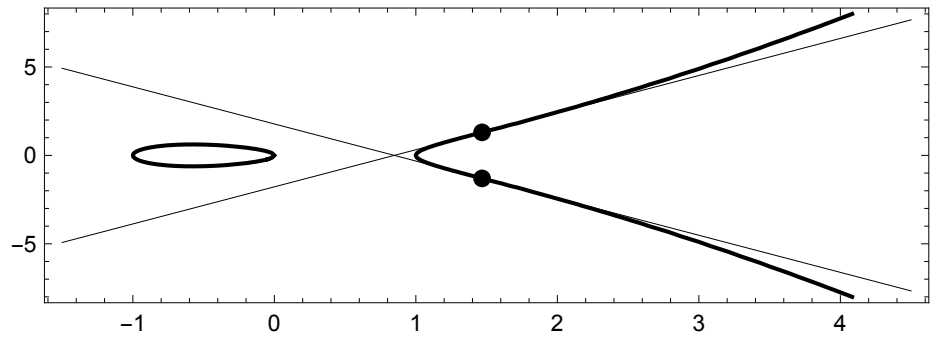



FIGURA 3. La curva $E_1(\mathbb{R})$. Se destaca los dos puntos 3-torsión reales y sus tangentes inflexivas.

*Mód* 3. Para $\ell$ un primo impar, hay relaciones de recursión clásicas que dan "polinomios de división" $f_\ell(y)$ de grado $\ell^2-1$ con raíces las $y$-coordenadas de los puntos de torsión en $E$ de orden $\ell$. Para abreviar, tratamos solo el caso $\ell=3$, donde la geometría es particularmente atractiva.

Puntos distintos $P$, $Q$ y $R$ sobre una curva elíptica $E: y^2=x^3+bx+c$ suman cero si y sólo si $P$, $Q$ y $R$ viven sobre una línea. Por supuesto que, un punto $P$ es un punto de 3-torsión si y sólo si $P+P+P=0$. La descripción geométrica de adición dice que $P$ es un punto de 3-torsión si y sólo si es un punto de inflexión de la curva. Calculando puntos de inflexión de la forma que en un curso de cálculo de primer año, obtenemos que

$$f_3(y)=27y^8+216cy^6-18\Delta y^4-\Delta^2$$

es el polinomio de división buscado. Aquí no importa si $E(\mathbb{R})$ tiene una o dos componentes; siempre exactamente dos de los ocho puntos de 3-torsión son reales.

En los dos casos, las álgebras $\mathbb{Q}[y]/f(b_j,c_j,y)$ también son presentadas como $\mathbb{Q}[z]/g_j(z)$ para polinomios con coeficientes mucho más pequeños:

$$g_1(z)=z^8+6z^4-3, \qquad |\operatorname{Gal}_1|=16, \quad D_1=-2^{16}3^7, \quad d_1=-2^6 3^3,$$
$$g_2(z)=z^8+4z^6-12z^2-12, \quad |\operatorname{Gal}_2|=48, \quad D_2=-2^{10}3^{11}, \quad d_2=-2^4 3^5.$$

El tamaño del grupo de Galois $|\operatorname{Gal}_i|$, el discriminante $D_j$ de la álgebra $\mathbb{Q}[z]/g_j(z)$, y el discriminante $d_j$ de $\mathbb{Q}[z]/g_j(\sqrt{z})$ son también indicados.

La siguiente tabla es análoga a (16.3), pero ahora para $\ell=3$.

(16.4)

| $\lambda_p$ | $F_p(T)$ | masas genéricas | # para $E_1$ | # para $E_2$ |
|---|---|---|---|---|
| $1^8$ | $1+T+T^2$ | 1/48 | 6253 | 2042 |
| $2^4$ | $1-T+T^2$ | 1/48 | 6246 | 2094 |
| $3^2\,1^2$ | $1+T+T^2$ | 1/6 | | 16584 |
| $6\ \ 2$ | $1-T+T^2$ | 1/6 | | 16686 |
| $4^2$ | $1\quad\ +T^2$ | 1/8 | 37463 | 12556 |
| $2^3\,1^2$ | $1\quad\ -T^2$ | 1/4 | 25027 | 24952 |
| 8 | $1-T-T^2$ | 1/8 | 12520 | 12545 |
| 8 | $1+T-T^2$ | 1/8 | 12491 | 12541 |

Aquí, las última columnas corresponde al número de primos entre $5, 7, \ldots, p_{100002}$ que tienen invariantes $(\lambda_p, a_p)$. Como $\operatorname{Gal}_2 = GL_2(\mathbb{F}_3)$, la columna para $E_2$ es gobernada por la columna de masa impresa. Como $\operatorname{Gal}_1$ es solo el subgrupo de 2-Sylow de $GL_2(\mathbb{F}_3)$, se rige por estadísticas diferentes. Uno puede correctamente suponer de la columna de $E_1$ que las frecuencias límites son $(1/16, 1/16, 0, 0, 3/8, 1/4, 1/8, 1/8)$.

Para $\ell$ general, las clases de Frobenius pertenecen a $GL_2(\mathbb{F}_\ell)^\sharp$, el conjunto de clases de conjugación del grupo $GL_2(\mathbb{F}_\ell)$. Notamos que el par de invariantes $(\lambda_p, F_p(T))$ determina estas clases completamente, pero ningún invariante por sí mismo es suficiente. En el caso $\ell=3$, el invariante $\lambda_p$ determina un conjunto cociente de seis elementos mientras que $F_p(T)$ determina un conjunto cociente de siete elementos. Para $\lambda_p$, el problema es la repetición de 8 en su columna, mientras que para $F_p(T)$ los dos problemas son la repetición de $1+T+T^2$ y $1-T+T^2$.

Recordemos que un problema fundamental es caracterizar la imagen de $\operatorname{IsAb}_g(\mathbb{Q})$ en $\mathcal{L}_g(\mathbb{Q})$. El hecho que para cualquier $\ell^e$, los coeficientes de $L(A,s)$ son completamente determinados en $\mathbb{Z}/\ell^e$ por un cuerpo de números es una restricción muy fuerte.

**16.9. Reducción mala en casos fáciles.** Hicimos hincapié en §16.1 y §16.2 que la manera natural para clasificar variedades abelianas principalmente polarizadas es aumentando el conductor. ¡Pero desde entonces no hemos dicho nada sobre reducción mala! En las dos subsecciones siguientes discutiremos brevemente este aspecto fundamental.

El estudio de la reducción mala de variedades abelianas es extremadamente complicado. En general, dados $A$ sobre $\mathbb{Q}$ y un primo $p$, se tiene una descomposición de dimensión

$$g = g_{\text{good}} + g_{\text{mult}} + g_{\text{add}}.$$

El polinomio de Frobenius $F_p(A,T)$ tiene grado $2g_{\text{good}} + g_{\text{mult}}$. Raíces inversas correspondientes a la parte buena tienen el valor absoluto usual $\sqrt{p}$. De todas maneras, aquellos que corresponden a $g_{\text{mult}}$ son raíces de la unidad. Abstractamente, los tres términos son respectivamente la dimensión de la parte buena, la parte toroidal, y la parte unipotente de la fibra especial del modelo de Néron para $A$.

En casos fáciles, las cantidades son calculables. Por ejemplo, en el marco hiperelíptico supongamos que el polinomio $f(x)$ tiene discriminante divisible exactamente por $p^k$ con $k \leqslant g$ y el polinomio reducido a $\mathbb{F}_p[x]$ tiene la forma $a(x)b(x)^2$ con $b(x)$ de grado $k$. Entonces $(g_{\text{good}}, g_{\text{mult}}, g_{\text{add}}) = (g-k, k, 0)$. La parte buena de $F_p(T)$ viene desde la curva $y^2 = a(x)$ de género $g-k$, y la parte multiplicativa de $F_p(T)$ puede ser calculada desde las raíces de $b(x)$ y sus tangentes.

Escribiendo al conductor como $N = \prod p^{c_p}$, uno generalmente calcula los individuos $c_p$ por separado. Se obtiene

$$c_p \geqslant g_{\text{mult}} + 2g_{\text{add}}.$$

La igualdad vale si y sólo si la ramificación es moderada. Una condición suficiente para que la ramificación sea moderada es que $p > 2g+1$. En el marco de la teoría de la ramificación, esta condición proviene del hecho que un grupo cíclico de orden $p$ no puede actuar de manera no trivial sobre el espacio vectorial racional $H^1(A(\mathbb{C}), \mathbb{Q})$ de dimensión $2g$.

**16.10. Reducción mala en casos difíciles.** El famoso algoritmo de Tate determina las deseadas cantidades $F_p(A,T)$ y $c_p$ directamente desde la ecuación de la curva elíptica. Un uso de polinomios de división es que, para un número primo $p$ diferente de $\ell$, la ramificación $p$-ádica en $\mathbb{Q}[y]/f_\ell(y)$ también da información sobre el exponente $c_p = \text{ord}_p(N)$. Para esta aplicación, algunas veces es suficiente usar solo los dos primeros $\ell$. En efecto, uno usa $\ell = 2$ para obtener información sobre los primos impares $p$, y luego uno usa $\ell = 3$ para resolver ambigüedades para $p \geqslant 5$ y obtener información sobre el caso más difícil $p = 2$.

**Ejemplo 16.5.** *Ejemplo de ramificación* 3*-ádica vía representaciones mód* 2. Sea $y^2 = x^3 + bx + c$ con exponente conductor $c_3 = \text{ord}_3(N)$ y sea $x^3 + bx + c$ con exponente discriminante $\delta_3 = \text{ord}_3(D)$. Entonces, $c_3 \geqslant 3$ si y sólo si $\delta_3 \geqslant 3$; en este caso $c_3 = \delta_3$.

**Ejemplo 16.6.** *Ejemplo de ramificación* 2*-ádica vía representaciones mód* 3. Sea $y^2 = x^3 + bx + c$ con exponente conductor $c_2 = \text{ord}_2(N)$ y sea $f_3(y)$ con exponente discriminante relativo $\delta_2 = \text{ord}_2(D/d)$. Supongamos que $\delta_2/4$ es la pendiente más grande en el cuerpo $\mathbb{Q}_2[y]/f_3(y)$, tal como ocurre en nuestros ejemplos. Entonces $c_2 = \delta_2/2$. En nuestro primer ejemplo se obtiene $c_2 = \text{ord}_2(D_1/d_1)/2 = 5$, mientras que en el segundo ejemplo se obtiene $c_2 = \text{ord}(D_2/d_2)/2 = 3$.

**16.11. Funciones $L$ como funciones analíticas de $s$.** Definamos la siguiente modificación en la función Gamma estándar: $\Gamma_{\mathbb{C}}(s) = 2(2\pi)^{-s}\Gamma(s)$. La función $L$ completa de una variedad abeliana $g$-dimensional es

$$\Lambda(A, s) := N^{s/2}\Gamma_{\mathbb{C}}(s)^g L(A, s) \tag{16.5}$$

Como mencionamos previamente, el producto que define a $L(A, s)$ converge sólo para $\mathrm{Re}(s) > 3/2$. Asimismo, nuevamente desde la década de 1960, se espera mucho más:

**Conjetura 16.7** (Conjetura de la función $L$)**.** *Para cualquier variedad abeliana $A$ sobre $\mathbb{Q}$, $\Lambda(A, s)$ es una función entera, acotada en bandas verticales, y satisfaciendo*

$$\Lambda(A, s) = \pm\Lambda(A, 2 - s). \tag{16.6}$$

La conjetura fue primeramente conocida para curvas elípticas con CM potencial. En los años noventa, la demostración para curvas elípticas fue muy famosa.

**Teorema 16.8** (Wiles et al.)**.** *La Conjetura de la función $L$ es verdadera para $g = 1$.*

La larga demostración comienza en [39]. Conecta curvas elípticas vía representaciones de Galois con formas modulares, y para las funciones $L$ de formas modulares ya se sabía que tenían las propiedades analíticas deseadas.

Sea $A$ una variedad abeliana sobre $\mathbb{Q}$ con álgebra de endomorfismos $D$ con centro $F$. Sea $\dim_F(D) = d^2$. Entonces la función $L$ $L(A, s)$ es la $d$-ésima potencia de una función que denotamos formalmente $L(A^{1/d}, s)$. Definimos $\Lambda(A^{1/d}, s) = N^{s/(2d)}\Gamma_{\mathbb{C}}(s)^{g/d}L(A^{1/d}, s)$. Entonces nuevamente se espera que Conjetura 16.7 sea cierta con $A$ reemplazado por $A^{1/d}$. Más aún, podríamos ser más optimista y esperar que cualquier función que satisfaga Conjetura 16.7 provenga de una variedad abeliana de esta manera. Esto sería una descripción de $\mathrm{IsAb}_g(\mathbb{Q})$ paralela a la descripción de Honda-Tate sobre $\mathrm{IsAb}_g(\mathbb{F}_q)$.

**16.12. Ejercicios.**

1. Realizar una búsqueda más extensa de curvas elípticas con $|a_1|, |a_2|, |a_3| \leqslant 1$ como antes, pero ahora con $|a_4|, |a_6| \leqslant 10$. ¿Cuántas de las 93 clases de isogenía con conductor $\leqslant 100$ se encontraron? ¿Cuántas de las 306 curvas se encontraron?
2. Explorar la sección de la LMFDB de curvas elípticas sobre $\mathbb{Q}$. Algunos posibles temas son:
   - ¿Cuáles conductores tienen una gran cantidad de clases de isogenía?
   - ¿Cuál es el significado de los enormes picos en la función $Z$ para la única curva en la base de datos con rango 4?
   - ¿Cuál es la cota de conductor mínima para la cual los trece invariantes $j$ aparecen?
   - Confirmar en unos pocos casos que toda curva con conductor divisible exactamente por $2^4$ o $2^6$ es un giro cuadrático de una curva de conductor menor.
   - La curva $X_0(1200)$ tiene género 205. Es de notar que su Jacobiana es isógena al producto de 205 curvas elípticas. ¿Cuántas clases de isogenía están involucradas? ¿Con cuáles multiplicidades?

3. Gross y Zagier probaron que todas las diferencias $j_D - j_{D'}$ con los $j_D$ como en §16.4 se factoriza en primos pequeños solamente. Por ejemplo, el código en *Magma* `Factorization(-3375+32768);` revela que la diferencia $j_{-7} - j_{-11}$ es $7 \cdot 13 \cdot 17 \cdot 19$. Intenta adivinar rasgos de la fórmula general sin leer la referencia [29].

4. El código de *Magma*
```
E2 := EllipticCurve([6,7]);
L2 := LSeries(E2);
&+[(Coefficient(L2,NthPrime(j))/Sqrt(NthPrime(j)))^4 :
   j in [1..100000]]/100000;
```
devuelven $1.995\ldots$, mientras que el cuarto momento de la correspondiente medida es $m_4 = \int_{-2}^{2} u^4 \mu_{Sp_2} = 2$. Este es un ejemplo cuantitativo de cuán bien las sucesiones $u_2$, $u_3$, $u_5$, ... encajan con la medida $\mu_{Sp_2}$. Los $m_k$ correctos son dados en la página de $\mu_{Sp_2}$ en la sección de Sato-Tate en la LMFDB. ¿Para cuáles $k$ aseguran los primeros 100000 $u_p$ el correcto $m_k$ luego del redondeo?

5. El código
```
F5T<T>:=PolynomialRing(FiniteField(5));
E2 := EllipticCurve([6,7]);
{* F5T!EulerFactor(E2,NthPrime(j)): j in [1..100000] *};
```
obtiene los primeros 100000 $F_p(E_2, T)$ como elementos de $\mathbb{F}_5[T]$. ¿Cuál subgrupo de $GL_2(\mathbb{F}_5)$ es la imagen de la representación mód 5? Repetirlo para $E_1$. ¿En qué lugar en la LMFDB está la respuesta?

6. *Magma* implementa para $\ell$ impares un polinomio de división de grado $(\ell^2 - 1)/2$ dando las coordenadas $x$ de los puntos $\ell$-división. El código
```
Qx<x>:=PolynomialRing(Rationals());
E1 := EllipticCurve([-1,0]);
DivisionPolynomial(E1,3);
```
devuelve este polinomio para la curva $E_1$ y $\ell = 3$. Repetirlo para $\ell = 5$, y utiliza el "identifier" a [32] para obtener información sobre el comportamiento 2-ádico de los dos polinomios. ¿Cómo comparan las pendientes 2-ádicas (dadas en la columna "Galois slope content")? Repetirlo para $E_2$.

7. Ve en la LMFDB de la página de $E_1$ hasta la página de su forma modular $f_1$, para aprender que $f_1 = q\prod_{n=1}(1-q^{4n})^2(1-q^{8n})^2 = \sum_{n=1}^{\infty} a_n q^n$, con los $a_n$ exactamente los coeficientes de Dirichlet de $L(E_1, s) = \sum_{n=1}^{\infty} a_n n^{-s}$. Escoge un número primo grande $p$ y compute $a_p$ de $E_1$. Independientemente, computa $a_p$ de $f_1$. ¿Cómo comparan el tiempo de ejecución de los computaciones?

## 17. Variedades abelianas sobre $\mathbb{Q}$: ejemplos de superficies

En esta última sección continuamos la discusión de invariantes y clasificación, aunque ahora considerando ejemplos del caso menos familiar de Jacobianas de curvas de género dos. Haremos contacto con cada una de las tres conjeturas mostradas en la sección anterior. Sin embargo, el punto principal es ilustrar cómo se ven las cosas desde un punto de vista computacional.

Como ejemplos explícitos de curvas, sean

$$C_1 : y^2 = (x-2)(x-1)(x+1)(x+2)\left(x^2-5\right), \qquad \begin{array}{l}\hat{\Delta}_1 = 2^{24}3^4 5,\\ \Delta_1 = 2^4 3^4 5,\\ N_1 = 360 = 2^3 3^2 5,\end{array} \tag{17.1}$$

$$C_2 : y^2 = x\left(x^2+1\right)\left(x^3-3x-4\right), \qquad \begin{array}{l}\hat{\Delta}_2 = 2^{26}3^4,\\ \Delta_2 = 2^6 3^4,\\ N_2 = 2592 = 2^5 3^4.\end{array} \tag{17.2}$$

La curva $C_1$ es especial porque no tiene solo la involución hiperelíptica $(x, y) \mapsto (x, -y)$, sino que además tiene la involución independiente $(x, y) \mapsto (-x, y)$. En contraste, veremos que $C_2$ tiene un comportamiento genérico.

**17.1. Tablas de curvas con conductor pequeño.** En general, consideremos una curva $C$ de género dos presentada de la forma $y^2+h(x)y = f(x)$, con $f(x) \in \mathbb{Z}[x]$ de grado seis y $h(x) \in \mathbb{Z}[x]$ de grado $\leqslant 3$. Su discriminante es $\hat{\Delta} = 2^{10}\,|\operatorname{disc}(f + \frac{h^2}{4})|$. El discriminante $\Delta$ es el mínimo de todos estos $\hat{\Delta}$ y otros que provienen de presentaciones donde $f(x)$ tiene grado cinco (estos usualmente no son necesarios).

En general el conductor divide al discriminante: $N|\Delta$. La desigualdad $\operatorname{ord}_p(N) \leqslant \operatorname{ord}_p(\Delta)$ usualmente está cerca de ser una igualdad, tal como lo ilustran nuestros ejemplos. En particular, todos los primos que dividen exactamente a $\Delta$, también dividen exactamente a $N$.

La LMFDB contiene los resultados de una extensa búsqueda [20] usando métodos muy eficientes para toda curva con $\Delta \leqslant 10^6$. La lista obtenida contiene 66158 curvas. La lista comienza con

| Conductor $N$ | Número de curvas de género dos en la LMFDB |
|---|---|
| 169 | 1 |
| 196 | 1 |
| 249 | 2 |
| 256 | 1 |
| 277 | 2 |
| 294 | 2 |
| 295 | 2 |
| 324 | 1 |

Nuevamente la lista comienza con conductores determinando clases de isogenía. La primera repetición se da cuando $N = 576 = 2^6 3^2$.

Pero ahora la situación con respecto a la completitud está lejos de ser la configuración óptima de curvas de género 1. Primero, es probable que haya algunas curvas con $\Delta \leqslant 10^6$ perdidas por la búsqueda. Mucho más en serio para las aplicaciones, hay muchas curvas con conductor $N$ pequeño, que no aparecen porque su discriminante $\Delta$ es más que $10^6$. El artículo [20] da evidencia de que la lista puede estar completa con respecto a las clases de isogenía para $N \leqslant 1000$. El final de §17.3 nos muestra que a la clase de isogenía de $C_1$ con $N = 360$ le faltan al menos cinco curvas. Los ejercicios da una clase de isogenía con $N = 1024$ también faltante.

**17.2. Análogos de invariantes $j$.** La fórmulas clásicas de esta subsección se dan con más información en la LMFDB. Sea

$$C : y^2 = f(x)$$

una curva de género dos con el polinomio $f(x) = cx^6 + \cdots$ teniendo raíces $\alpha_1, \ldots, \alpha_6$. Abreviamos $(\alpha_i - \alpha_j)^2$ por $[i,j]$. Entonces los invariantes de Igusa-Clebsch de la curva $C$ explícitamente presentada son

$$\begin{aligned}
I_2 &= c^2\,([1,2][3,4],[5,6] + \textit{14 términos parecidos}),\\
I_4 &= c^4\,([1,2][2,3][3,1][4,5][5,6][6,4] + \textit{9 términos parecidos}),\\
I_6 &= c^6\,([1,2][2,3][3,1][4,5][5,6][6,4][1,4][2,5][3,6] + \textit{59 términos parecidos}),\\
I_{10} &= c^{10}\prod_{i<j}[i,j].
\end{aligned}$$

Cada $I_k$ puede ser escrito como un polinomio en los coeficientes de $f(x)$, homogéneo de grado $k$.

Para estar a gusto con el resto de la literatura, uno debe conocer varias ligeras variantes. Por ejemplo, los invariantes de Igusa son

$$\begin{aligned}
J_2 &= I_2/8,\\
J_4 &= (4J_2^2 - I_4)/96,\\
J_6 &= (8J_2^3 - 160J_2J_4 - I_6)/576,\\
J_{10} &= I_{10}/4096.
\end{aligned}$$

La variedad de móduli compactada $\overline{M}_2$ para las curvas de género dos es el espectro proyectivo $\operatorname{Proj} R$ del anillo graduado

$$R = \mathbb{Q}[I_2, I_4, I_6, I_{10}] = \mathbb{Q}[J_2, J_4, J_6, J_{10}].$$

La variedad $M_2$ en sí misma es el complemento de la hipersuperficie discriminante $I_{10} = 0$, o equivalentemente $J_{10} = 0$. Los invariantes de Igusa fueron introducidos ya que ellos se comportaban mejor cuando eran reducidos módulo 3 y 5. Cuando se complementa con un invariante similar $J_8$ se comportan bien cuando se reduce módulo 2.

El espacio $M_2$ es de dimensión 3 y singular. Sin embargo, se puede diseccionar inteligentemente en tres partes y volver a montar para crear el espacio afín ordinario de la siguiente manera. Definamos el invariante $g$ por

$$(g_1, g_2, g_3) = \begin{cases} (J_2^5/J_{10}, & J_2^3/J_{10}, & J_2^2J_6/J_{10}) & \text{si } J_2 \neq 0,\\ (0, & J_4^5/J_{10}^2, & J_4J_6/J_{10}) & \text{si } J_2 = 0 \text{ y } J_4 \neq 0,\\ (0, & 0, & J_6^5/J_{10}^3) & \text{si } J_2 = J_4 = 0. \end{cases} \tag{17.3}$$

Entonces vía $(g_1, g_2, g_3)$, tenemos $M_2(K) = K^3$.

Tal como el invariante $j$, los invariantes $g$ son números racionales típicamente de altura grande. Por ejemplo,

| para $C_1$, | para $C_2$, |
|---|---|
| $g_1 = 28596971960000/81$, | $g_1 = 0$, |
| $g_2 = 1150492082200/81$, | $g_2 = 3125/3456$, |
| $g_3 = 6677950400/9$, | $g_3 = -110/27$. |

Nuevamente los denominadores son significativos, ya que reflejan que $J_{10}$ es divisible por $p$ si y sólo si $f(x) \in \mathbb{Z}[x]$ continúa teniendo seis raíces distintas en la línea proyectiva en característica $p$.

Es fácil calcular todos estos invariantes con *Magma*. Por ejemplo,

```
Qx<x>:=PolynomialRing(Rationals());
C2 := HyperellipticCurve([x*(x^2+1)*(x^3-3*x-4),0]);
G2Invariants(C2);
```

da el vector $(g_1, g_2, g_3)$ de $C_2$ muy rápidamente. Se tiene $A_2 = M_2 \coprod S$ donde $S$ es el producto simétrico de dos copias de la línea $j$. Entonces para entender $A_2$, entender $M_2$ es el paso principal.

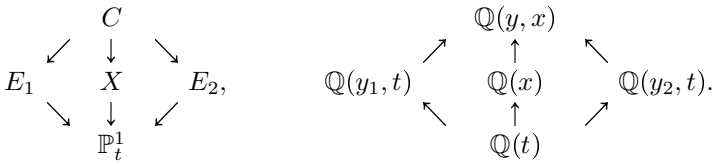
**17.3. Una subvariedad clásica de $A_2$.** Muchos anillos $R$ surgen como anillos de endomorfismos de superficies abelianas que uno prácticamente puede enumerar. En consecuencia, hay muchas subvariedades naturales $X_R$ de $A_2$, y la situación es mucho más complicada que la colección de subvariedades de dimensión cero $X_D$ de $A_1$ tratada en §16.4. Discutimos solo una de las subvariedades más simples y clásicas, el de $R = \{(x, y) \in \mathbb{Z}^2 : x \equiv y\ (2)\}$. La ecuación para este $X_R$ es ya muy complicada.

Sean $E_1$ y $E_2$ curvas elípticas sobre $\mathbb{Q}$ con todos los puntos de 2-torsión racionales. Se pueden escribir en la forma de Legendre como

$$y_1^2 = t(t-1)(t-\lambda), \qquad y_2^2 = t(t-1)(t-\mu). \tag{17.4}$$

Entonces, Legendre mostró que se puede "pegar" $E_1 = E_\lambda$ y $E_2 = E_\mu$ en una curva $C = C_{\lambda,\mu}$ de género dos como sigue.

A la izquierda tenemos un diagrama de curvas:




Aquí $C$ es el producto fibrado de $E_1$ y $E_2$. Su cuerpo de funciones $\mathbb{Q}(C) = \mathbb{Q}(t, y_1, y_2)$ es una extensión de grado cuatro del cuerpo base $\mathbb{Q}(t)$ con grupo de Galois que tiene cuatro elementos $(1,1)$, $(1,-1)$, $(-1,1)$, y $(-1,-1)$. El elemento $(\epsilon_1, \epsilon_2)$ actúa por $y_1 \mapsto \epsilon_1 y_1$ y $y_2 \mapsto \epsilon_2 y_2$. La conducta de la ramificación nos dice que $C$ tiene género dos y el cociente $X := C/(-1,-1)$, con cuerpo de funciones $\mathbb{Q}(t, y_1 y_2)$, tiene género cero.

Hay una coordenada $x$ en $X$ que lo identifica con la línea proyectiva $\mathbb{P}^1_x$ de tal manera que el mapeo a $\mathbb{P}^1_t$ toma la forma

$$t = \frac{\mu x^2 - \lambda}{x^2 - 1}. \tag{17.5}$$

Define

$$y = (-1+x)^2(1+x)(y_1 + y_2). \tag{17.6}$$

Eliminando las variables $y_1$, $y_2$ y $t$ del sistema (17.4),(17.5),(17.6), obtenemos una ecuación estándar para $C$,

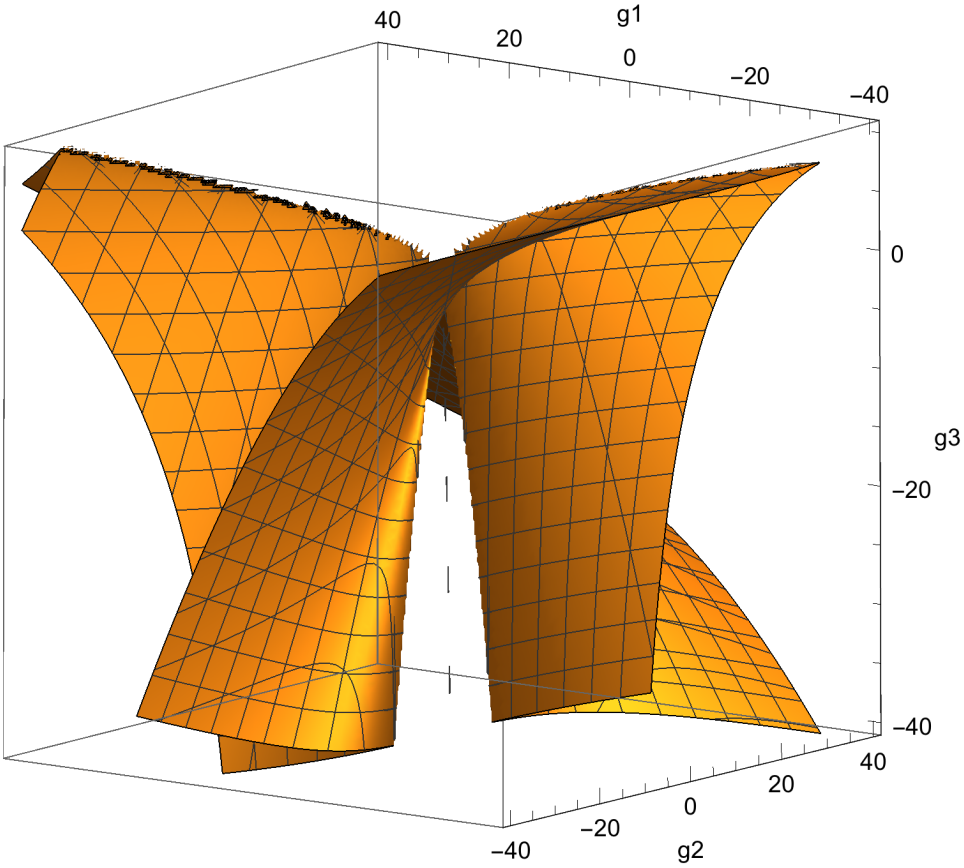
$$y^2 = (\mu - \lambda)(x-1)(x+1)\left(\lambda - \mu x^2\right)\left(\lambda - \mu x^2 + x^2 - 1\right). \tag{17.7}$$

Para obtener una ecuación para la subvariedad de $M_2$ correspondiente a esta construcción, calculamos los invariantes de Igusa $(J_2, J_4, J_6, J_{10})$ y buscamos una relación lineal entre los monomios en $J_i$ de un grado dado. Estos monomios son polinomios gigantes en $\lambda$ y $\mu$. La primera relación lineal ocurre en grado 30, donde

existen 47 monomios. Solo 29 de ellos están involucrados en la relación. Traduciendo a invariantes absolutos para el primer régimen de (17.3), la relación es

$$
\begin{aligned}
&-51200000g_1^4+432g_1^5-28800g_1^4g_2+512000g_1^3g_2^2-8g_1^3g_2^3+512g_1^2g_2^4\\
&-8192g_1g_2^5+96000g_1^4g_3-11520000g_1^3g_2g_3+72g_1^4g_2g_3-4816g_1^3g_2^2g_3\\
&+84480g_1^2g_2^3g_3-g_1^2g_2^4g_3+64g_1g_2^5g_3-1024g_2^6g_3+48g_1^4g_3^2+12960g_1^3g_2g_3^2\\
&-691200g_1^2g_2^2g_3^2+2g_1^3g_2^2g_3^2-136g_1^2g_2^3g_3^2+2304g_1g_2^4g_3^2+129600g_1^3g_3^3\\
&-g_1^4g_3^3+72g_1^3g_2g_3^3-1080g_1^2g_2^2g_3^3-6912g_1g_2^3g_3^3-216g_1^3g_3^4+7776g_1^2g_2g_3^4\\
&-11664g_1^2g_3^5=0.
\end{aligned}
\tag{17.8}
$$

Por diversión, aquí tenemos un vistazo de la superficie $X_R(\mathbb{R})$ en el espacio real con coordenadas $(g_1, g_2, g_3)$.




La curva $C_1$ fue construida por el método de esta subsección, con $(\lambda, \mu) = (-15, -3)$. En efecto, dividiendo ambos lados de la ecuación (17.7) de $C_{-15,-3}$ por $12^2$, obtenemos la ecuación (17.1) de $C_1$. El primer factor $E_{-15}$ es uno de las ocho curvas elípticas con conductor 15. Tres de estas curvas tiene 2-torsión partida, estas son $E_\lambda$ con

$$\lambda \in \{-15, -9/16, 81\}.$$

El segundo factor $E_{-3}$ es una de las seis curvas elípticas con conductor 24. Dos de estas curvas tienen 2-torsión partida, a saber $E_\mu$ con

$$\mu \in \{-3, 9\}.$$

Cuando se pegan las curvas elípticas de esta manera, las funciones $L$ y los conductores se multiplican, así que $C_{-15,-3}$ tiene conductor $15 \cdot 24 = 360$. Cuando los factores de curvas elípticas no son isógenas entre sí, la clase de isogenía se comporta multiplicativamente. Así la clase de isogenía de la Jacobiana de $C_1 = C_{-15,-3}$ contiene exactamente $8 \times 6 = 48$ elementos. Al menos seis de estas 48 superficies abelianas tiene polarización principal, a saber las seis $C_{\lambda,\mu}$. Asimismo, la LMFDB actualmente tiene sólo $C_1$.

**17.4. Polinomios de Frobenius y grupos de Galois motívicos.** Evaluando (15.6) se obtienen polinomios de Frobenius buenos que ahora veremos. Los polinomios malos correctos son mostrados nuevamente en negrita. En el caso de $C_1$, todos los polinomios, buenos y malos, son conocidos por la construcción de pegar, como $F_p(C_1, T) = F_p(E_{-15}, T)F_p(E_{-3}, T)$. La columna $F_p(C_1, T)$ da $F_p(E_{-15}, T)$ seguido de $F_p(E_{-3}, T)$.

| $p$ | $F_p(C_1, T)$ | | $F_p(C_2, T)$ |
|---|---|---|---|
| 2 | $(\mathbf{1 + T + 2T^2})$ | $\mathbf{1}$ | $\mathbf{1 + T + 2T^2}$ |
| 3 | $(\mathbf{1 + T})$ | $(\mathbf{1 + T})$ | $\mathbf{1 + 2T + 3T^2}$ |
| 5 | $(\mathbf{1 - T})$ | $(\mathbf{1 + 2T + 5T^2})$ | $1 + T + 5T^3 + 25T^4$ |
| 7 | $(1 + 7T^2)$ | $(1 + 7T^2)$ | $1 + 6T + 18T^2 + 42T^3 + 49T^4$ |
| 11 | $(1 + 4T + 11T^2)$ | $(1 - 4T + 11T^2)$ | $1 - 2T + 6T^2 - 22T^3 + 121T^4$ |
| 13 | $(1 + 2T + 13T^2)$ | $(1 + 2T + 13T^2)$ | $1 + 5T + 24T^2 + 65T^3 + 169T^4$ |
| 17 | $(1 - 2T + 17T^2)$ | $(1 - 2T + 17T^2)$ | $1 - T - 4T^2 - 17T^3 + 289T^4$ |
| 19 | $(1 - 4T + 19T^2)$ | $(1 + 4T + 19T^2)$ | $1 + 30T^2 + 361T^4$ |
| 23 | $(1 + 23T^2)$ | $(1 + 8T + 23T^2)$ | $1 + 4T - 2T^2 + 92T^3 + 529T^4$ |
| 29 | $(1 + 2T + 29T^2)$ | $(1 - 6T + 29T^2)$ | $1 - 3T + 32T^2 - 87T^3 + 841T^4$ |

Recordemos de §16.6 el formalismo de polinomios de Frobenius buenos $F_p(A, T)$ para una variedad abeliana $A$ con grupo de Galois motívico $G$. Ellos viven en la imagen de $G^\natural(\mathbb{Q})$ en $\mathrm{GSp}_{2g}^\natural(\mathbb{Q})$. Así que calculando secuencialmente $F_p(A, T)$ para más y más $p$, se obtiene una mejor cota inferior para $G$. Rápidamente se tiene un "buen pálpito" para $G$, el cual en la práctica es generalmente correcto. Por ejemplo, la columna $F_p(C_1, T)$ dice que la Jacobiana $J_1$ se parece al producto de dos curvas elípticas.

En este contexto, hay una proposición general que es muy útil:

**Proposición 17.1.** *Sea $A$ una variedad abeliana $g$-dimensional sobre $\mathbb{Q}$. Sean $F_p(A, T)$ y $F_q(A, T)$ dos polinomios de Frobenius con* $\mathrm{Gal}(F_p(A, T)F_q(A, T))$ *tan largo como sea posible, es decir, de orden $(2^g g!)^2$. Entonces, el grupo de Galois motívico $G$ de $A$ es tan grande como es posible, a saber,* $\mathrm{GSp}_{2g}$.

De hecho, si para un primo $p$ se tiene que $|\mathrm{Gal}(F_p(A, T))| = 2^g g!$, entonces restan muy pocas posibilidades para $G$, por la clasificación de subgrupos de grupos reductivos que contienen un toro maximal. Si para un segundo primo $q$, el subgrupo contiene un toro maximal completamente diferente, la única posibilidad es $G = \mathrm{GSp}_{2g}$.

Es fácil de aplicar Proposición 17.1. Por ejemplo, $F_p(C_2, T)$ tiene grupo de Galois de orden 8 cuando $p \in \{5, 11, 13, 17, 23\}$. Ya los primeros dos de estos primos son suficientes, pues

```
Order(GaloisGroup(EulerFactor(C2,5)*EulerFactor(C2,11)));
```

devuelve 64.

**17.5. Equidistribución arquimediana.** Preparamos el escenario describiendo la equidistribución arquimediana en nuestros ejemplos. Las medidas de Sato-Tate en nuestros dos casos pueden ser escritos como densidades $f_{ST}dudv$. Las densidades son

$$f_{Sp_2\times Sp_2} = \frac{1}{2\pi^2}\sqrt{\frac{(-2u+v+2)(2u+v+2)}{u^2-4v+8}}, \tag{17.9}$$
$$f_{Sp_4} = \frac{\sqrt{(u^2-4v+8)\,(-2u+v+2)(2u+v+2)}}{4\pi^2}.$$

La equidistribución de clases de Frobenius $\mathrm{fr}_p = (u_p, v_p)$ es parcialmente sabida en el primer caso, ya que por los dos factores se puede aplicar Teorema 16.1. Casi nada se conoce en el segundo caso. Los primeros cien $\mathrm{fr}_p$ en nuestros dos casos coinciden en la densidad de Sato-Tate, todas ilustradas en Figura 4.

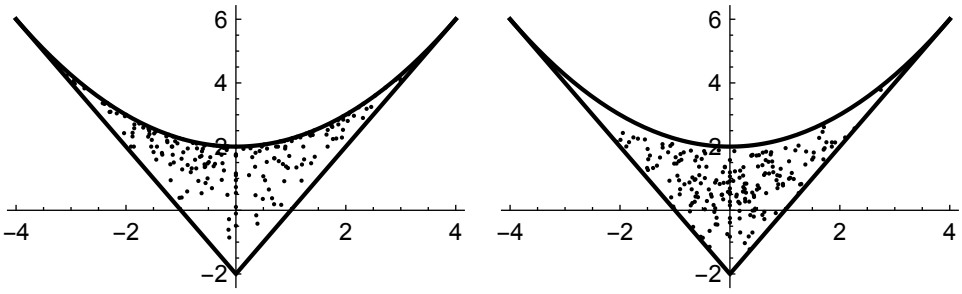



FIGURA 4. Puntos en el escudo $Sp_4^\natural$ representando el $\mathrm{fr}_p$ para los primeros cien primos buenos para $C_1$ (izquierda) y $C_2$ (derecha).

Algunas veces, como veremos pronto, uno se interesa solo en la distribución de $a_1$. La conjetura de Sato-Tate dice que estos números $a_1$ son controlados por la medida de probabilidad inducida por $\mu_{ST}$ en $[-2g, 2g]$. Si la medida tiene una densidad, la cual es garantizada si $ST$ es conexo, escribimos la densidad como $\phi_{ST}$.

Para calcular la medida inducida en el eje $u_1$, hay que integrar las variables restantes. En nuestros casos integramos sobre $v = u_2$ para obtener funciones en $u = u_1$:

$$\frac{24\pi^2\phi_{Sp_2\times Sp_2}}{u+4} = \left(u^2+16\right)E\left(\frac{(u-4)^2}{(u+4)^2}\right) - 8uK\left(\frac{(u-4)^2}{(u+4)^2}\right), \tag{17.10}$$
$$\frac{240\pi^2\phi_{Sp_4}}{u+4} = \left(u^4+224u^2+256\right)E\left(\frac{(u-4)^2}{(u+4)^2}\right)$$
$$- 8u\left(u^2+24u+16\right)K\left(\frac{(u-4)^2}{(u+4)^2}\right).$$

Aquí, $E$ y $K$ son integrales elípticas completas clásicas. A pesar de la formas funcionales complicadas de las $\phi_{ST}$, los gráficos tienen una apariencia simple, como se muestra en Figura 5.

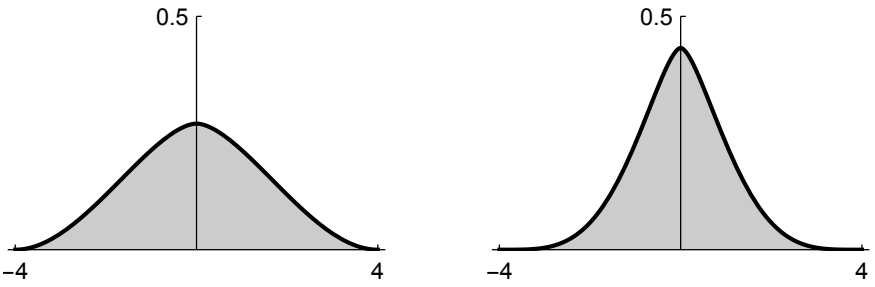



FIGURA 5. Densidades $\phi_{Sp_2\times Sp_2}$ y $\phi_{Sp_4}$ con varianza 2 y 1.

La fórmula del carácter de Weyl da expresiones explícitas para $f_{Sp_2^g}$ y $f_{Sp_{2g}}$ para $g$ general, similar en apariencia al caso $g = 2$ (17.9). Sin embargo, las funciones $\phi_{Sp_2^g}$ y $\phi_{Sp_{2g}}$ se vuelven más complicadas cuando $g$ crece. La ecuación diferencial lineal natural que ellos satisfacen tiene grado $g$ y puntos singulares en $-2g, -2g + 4, \ldots, 2g - 4, 2g$ y $\infty$.

*Rangos de anillos de endomorfismos vía el segundo momento.* La dificultad de calcular $u_j = a_j/p^{j/2}$ en una clase de Frobenius $\mathrm{fr}_p = (u_1, \ldots, u_g)$ se aumenta rápidamente con $j$. Una situación típica cuando $g$ es grande es que se puede calcular una gran cantidad de $u_1$ pero ningún $u_g$. En esta situación, Proposición 17.1 no está disponible para ayudar a determinar el grupo de Galois motívico $G$.

En este contexto se puede a veces hacer buenos pálpitos sobre $G$ si uno asume la Conjetura de Sato-Tate. Escribiendo ahora $u_p$ para la primera coordenada de $\mathrm{fr}_p$, la conjetura asegura en particular que

$$\lim_{x\to\infty} \frac{1}{\pi(x)} \sum_{p\leqslant x} u_p^k = \int_{-2g}^{2g} \phi_{ST}(u) u^k du. \tag{17.11}$$

Calculando los radios finitos análogos para una $x$ grande, uno puede intentar usar esta información para determinar algunos momentos $m_k$ y luego $ST$ mismo.

Los momentos para $k$ impar son todos cero, y la atención se concentra en los momentos pares. El primer momento par no trivial $m_2$ es particularmente interesante, ya que es el rango de $\mathrm{End}(A)$. Luego, si $m_2 = 1$, existe solo la posibilidad de $\mathrm{End}(A) = \mathbb{Z}$. En general, escribiendo $E_1$, $E_2$ y $E_4$ para $\mathbb{R}$, $\mathbb{C}$ y $\mathbb{H}$ respectivamente, las posibilidades para $\mathrm{End}(A)_{\mathbb{R}} := \mathrm{End}(A) \otimes_{\mathbb{Z}} \mathbb{R}$ son

$$\bigoplus_i M_{j_i}(E_{d_i}) \tag{17.12}$$

con $\sum d_i j_i^2 = m_2$.

*Clasicidad vía el cuarto momento.* Se hace más difícil utilizar las clases de Frobenius para determinar con precisión $m_k$ cuando $k$ crece, ya que un análisis probabilístico suponiendo válida la Conjetura de Sato-Tate dice que la convergencia se vuelve más lenta. Sin embargo, uno aún puede identificar $m_4$ con confianza.

El cuarto momento es particularmente importante. Escribimos $G_{1,g}$, $G_{2,g}$, y $G_{4,g}$ para los grupos compactos $Sp_{2g}$, $U_g$, y $O_{g/2}$ en sus representaciones simplécticas naturales de dimensión $2g$. Candidatos fuertes para $ST$ correspondiente al álgebra de endomorfismos (17.12) son

$$\prod_i G_{d_i,g_i}. \tag{17.13}$$

con $\sum_i g_i = g$. Tengamos en cuenta que si $m_2$ es grande admite una gran lista de posibilidades. Por ejemplo, $m_2 = 2$ permite muchos $Sp_{2g_1} \times Sp_{2g_2}$, y también $U_g$ si $g$ es par.

Un análisis relativamente fácil de los momentos, extendido en la discusión de los límites de Gauss que haremos a continuación, dice que

$$m_4 \geqslant 3m_2. \tag{17.14}$$

Supongamos, para hacer un enunciado limpio, que todos los $g_i$ son al menos 2. Entonces, la alternativa de Larsen [33] nos dice que la igualdad vale si y sólo si $G$ tiene el mismo grupo derivado que (17.13). "El mismo grupo derivado" es realmente necesario, ya que $m_2$ y $m_4$ no pueden distinguir entre $U_g$ y $SU_g$, ni tampoco entre $O_{g/2}$ y $SO_{g/2}$. El caso más simple es el que se encuentra con mayor frecuencia en la práctica: *si* $(m_2, m_4) = (1, 3)$ *entonces A tiene grupo de Sato-Tate* $Sp_{2g}$.

*Límites gaussianos.* La medida gaussiana con promedio cero y varianza $v$ en la línea $u$ es $\mu_v = e^{-u^2/(2v)}du/\sqrt{2\pi v}$. Sus momentos pares son $m_k = v^{k/2}(k-1)!!$. Aquí, el doble factorial es como un factorial regular, excepto que uno baja por dos como en $7!! = 7 \cdot 5 \cdot 3 \cdot 1 = 105$. El grupo $Sp_{2g}$ en su representación estándar de dimensión $2g$ tiene los mismos momentos para $k \leqslant g$ que $\mu_1$, y entonces momentos más pequeños. Por ejemplo, los primeros momentos pares para $Sp_4$ son $(m_2, m_4, m_6) = (1, 3, 14)$, lo cual es apenas inferior a los valores asintóticos $(1, 3, 15)$ alcanzados ya en $g = 3$. Del mismo modo, $SU_g$ y $SO_{g/2}$ en sus representaciones estándares de dimensión $2g$ tienen momentos coincidiendo con $\mu_2$ y $\mu_4$ para $k \leqslant g - 1$.

En general la medida $\mu$ en $\mathbb{R}$ asociada con la representación de $G_1 \times G_2$ en $V_1 \oplus V_2$ es la convolución de las medidas $\mu_i$ asociadas con $(G_1, V_1)$ y $(G_2, V_2)$: $\mu = \mu_1 * \mu_2$. Las varianzas siempre se suman cuando convolucionamos y la convolución de dos gaussianas es gaussiana. La medida en $\mathbb{R}$ asociada a $(G, V^m)$ es el reescalamiento por $m$ de la medida asociada con $(G, V)$, por lo que las varianzas aumentan por el factor $m^2$. Este dibujo muestra la densidad $\phi_G$ perteneciendo al grupo $G$ de la forma (17.13) con mín$(g_i)$ grande es muy cercana a una Gaussiana con promedio cero y varianza $m_2$.

*Un ejemplo exótico en género dos.* Describimos tres posibles grupos de Sato-Tate para curvas elípticas: $Sp_2$ y $U_1.2$ ocurren sobre $\mathbb{Q}$ y $U_1$ no lo hace. Para género dos, fue probado en [28] que son 34 las posibilidades que ocurren sobre $\mathbb{Q}$ y entonces 18 más posibilidades que solo ocurren sobre cuerpos de números más grandes. Cada uno de los 52 grupos tiene su propia página web en la sección de Sato-Tate de la LMFDB. Por supuesto, el número de posibilidades aumenta rápidamente con $g$.

Presentamos ahora un ejemplo de [28], el grupo llamado $J(O)$ allí. Como muchos de los 52 grupos, es construido a partir de un grupo finito $G_1$ y un grupo infinito $G_2$, cada uno en su propia representación 2-dimensional $V_1$ y $V_2$, y cada uno conteniendo la matriz escalar $-1$. El grupo de Sato-Tate es entonces $(G_1 \times G_2)/\{\pm(1,1)\}$, actuando en el espacio 4-dimensional $V_1 \otimes V_2$.

En el caso que $ST = J(O)$, el grupo finito $G_1$ es $\tilde{S}_4 \subset Sp_2$, un cubrimiento doble de $S_4 \subset SO_3$, mejor pensado como rotaciones de un cubo en el 3-espacio. El grupo infinito es $G_2 = O_2$. Las medidas de probabilidad inducidas en el intervalo $[-2, 2]$ son

$$\mu_1 = \frac{1}{48}\delta_{-2} + \frac{1}{8}\delta_{-\sqrt{2}} + \frac{1}{6}\delta_1 + \frac{3}{16}\delta_0 + \frac{1}{6}\delta_1 + \frac{1}{8}\delta_{\sqrt{2}} + \frac{1}{48}\delta_2\,,$$
$$\mu_2 = \frac{1}{2}\mu_0 + \frac{dy}{2\pi\sqrt{4-y^2}}\,.$$

Mientras que el producto semidirecto $O_2 = SO_2.2$ es un grupo diferente de la extensión no partida $U_1.2$, induce la misma medida en $[-2, 2]$.

Para esta construcción de producto tensorial en general, el mapa natural envía un punto $(x, y) \in [-2, 2] \times [-2, 2]$ al punto $(u, v) = (xy, x^2 + y^2 - 2)$ en el escudo $Sp_4^\natural$. La medida $\mu_1 \times \mu_2$ avanza hacia la medida deseada $\mu_{ST}$. En el caso $ST = J(O)$ uno obtiene

$$\mu_{J(O)} = \frac{3}{16}\delta_{p_2} + \frac{1}{6}\delta_{p_3} + \frac{1}{8}\delta_{p_4} + \frac{1}{48}\delta_{p_1} + \frac{3}{16}\nu_{V_2} + \frac{1}{6}\nu_{C_3} + \frac{1}{8}\nu_{C_4} + \frac{1}{48}\nu_{C_1}. \tag{17.15}$$

Así, la medida $\mu_{J(O)}$ tiene la mitad de su soporte sobre cuatro puntos especiales en la Figura 2, y la otra mitad en cuatro curvas especiales. Las $\nu_C$ son medidas de probabilidad. Son todas trasladadas de la medida con densidad $f(y) = 1/(\pi\sqrt{4-y^2})$ sobre $[-2, 2]$ y cero afuera. Por consecuencia, le medida unidimensional sobre $[-4, 4]$ es

$$\nu_{J(O)} = \frac{11}{16}\delta_0 + \left(\frac{f(u)}{6} + \frac{f\left(u/\sqrt{2}\right)}{8\sqrt{2}} + \frac{f\left(u/2\right)}{96}\right) du. \tag{17.16}$$

La imagen de esta medida en la página web de $J(O)$ en la LMFDB es redibujada en Figura 6. Desde un punto de vista inocente, es sorprendente que uno podría mirar

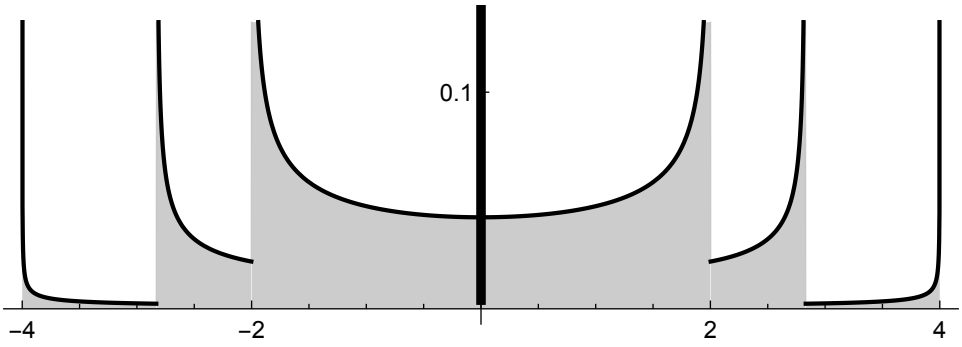



FIGURA 6. La medida Sato-Tate $\nu_{J(O)}$ de (17.16), con masa 11/16 a 0 y la masa restante dada por una densidad discontinua.

cientos de curvas y ver sólo distribuciones del tipo gaussiano de Figura 5, y luego de repente encontrarse con $\nu_{J(O)}$ desde la inofensiva curva $y^2 = x^6 - 5x^4 + 10x^3 - 5x^2 + 2x - 1$.

La receta para momentos es aún más fácil. Supongamos $\mu_{G_1}$ y $\mu_{G_2}$ en el intervalo $[-2, 2]$ tienen momentos $m'_k$ y $m''_k$ respectivamente. Entonces los momentos de $\nu_{ST}$

en $[-4,4]$ son $m_k = m'_k m''_k$. Siempre, todos los momentos impares se anulan. En nuestro ejemplo, los momentos pares son

(17.17)

| | $m_2$ | $m_4$ | $m_6$ | $m_8$ |
|---|---|---|---|---|
| $G_1 = \tilde{S}_4$ | 1 | 2 | 5 | 15 |
| $G_2 = O_2$ | 1 | 3 | 10 | 35 |
| $ST = J(O)$ | 1 | 6 | 50 | 525 |

Los momentos para $G_1$ y $G_2$ son calculados por métodos diagramados simples en el último ejercicio. Los momentos para $ST$ están dados en la página web de $J(O)$ en la LMFDB.

**17.6. Representaciones de Galois mód $\ell$.** Para curvas hiperelípticas $y^2 = f(x)$ la representación mód 2 está dada con la curva. A saber, el grupo de Galois $\mathrm{Gal}(f(x))$ está en $S_{2g+2}$ y se tiene una inclusión

$$S_{2g+2} \to \mathrm{GSp}_{2g}(\mathbb{F}_2).$$

Para $g = 1$ y $g = 2$, esta inclusión es suryectiva, reflejando el hecho que curvas elípticas y curvas de género dos son siempre hiperelípticas. Para $g = 3$, las 28 bitangentes sobre una curva cuártica le permiten a uno obtener la representación mód 2 nuevamente, aunque para $g \geqslant 4$ fórmulas explícitas parecen fuera del alcance para curvas generales.

Para $g \geqslant 2$ y $\ell \geqslant 3$, solo hay un caso para el cual uno tiene polinomios universales. Este caso único es el primer caso, $g = 2$ y $\ell = 3$. El método clásico para producir el polinomio de división se describe con detalles en [26]. El polinomio para $y^2 = x^5 + bx^3 + cx^2 + dx + e$ es par y comienza

$$f_3(b,c,d,e;x) = x^{80} + 15120\,b\,x^{76} + +2620800\,c\,x^{74} + (419237280\,d - 35394408\,b^2)\,x^{72} + \cdots \tag{17.18}$$

Expandido como un elemento de $\mathbb{Z}[b,c,d,e,x]$, tiene 1673 términos.

Tal como enfatizamos en el caso de género uno, representaciones mód $\ell$ tienen diferentes propósitos. Uno de ellos es dar acceso independiente a polinomios de Frobenius reducidos a $\mathbb{F}_\ell[T]$. Nuestros dos casos son muy degenerados para $\ell = 2$. Las distribuciones de $(\lambda_p, F_p(T))$ para los primeros $10^5$ primos buenos están en las dos últimas columnas:

| $\lambda_p$ | $F_p(T) \in \mathbb{F}_2[T]$ | masas genéricas | # para $C_1$ | #para $C_2$ |
|---|---|---|---|---|
| $1^6$ | $(1+T)^4$ | 1/720 | 49977 | 16569 |
| $2\ 1^4$ | ” | 1/48 | 50023 | 50051 |
| $2^3$ | ” | 1/48 | | |
| $2^2\,1^2$ | ” | 1/16 | | |
| $4\ 2$ | ” | 1/8 | | |
| $4\ 1^2$ | ” | 1/8 | | |
| $3\ 1^3$ | $(1+T)^2\left(1+T+T^2\right)$ | 1/18 | | 33380 |
| $3\ 2\ 1$ | ” | 1/6 | | |
| $3^2$ | $\left(1+T+T^2\right)^2$ | 1/18 | | |
| $6$ | ” | 1/6 | | |
| $5\ 1$ | $1+T+T^2+T^3+T^4$ | 1/5 | | |

La división de la tabla en cuatro bloques muestra muy claramente cómo un polinomio de Frobenius determina solo la parte semisimple de una clase de conjugación. Aunque la masa de $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ se convierte en equidistribuida en el espacio de polinomios característicos $\mathrm{GSp}^\natural_{2g}(\mathbb{F}_\ell)$ en el límite $\ell \to \infty$, existen notables discrepancias para $\ell$ pequeño. Los cuatro bloques en orden contienen respectivamente $35.\overline{5}\,\%$, $22.\overline{2}\,\%$, $22.\overline{2}\,\%$ y $20\,\%$ de la masa.

Otro propósito de representaciones mód $\ell$, descriptas ya en §16.10, es el de analizar la mala reducción. Como un ejemplo de esto, aplicamos (17.18) a la curva $C_2$ buscando información sobre la reducción mala de $C_2$ en 2. Cambiando coordenadas en (17.2) vía $(x, y) \mapsto (-100/(15 + x), -20y/(15 + x)^3)$ para expresar $C_2$ vía un polinomio de quinto grado, evaluamos (17.18) en $(b, c, d, e) = (7750, -117500, -9009375, 2418212500)$. La factorización en irreducibles en $\mathbb{Q}_2[x]$ tiene la forma $f_{8r}(x)f_{8u}(x)f_{64}(x)$. El factor mayor no tiene que ser estudiado y el sitio web de [32] dice que $f_{8u}(x)$ es no ramificado. Aplicando este sitio web en una manera menos trivial muestra que $f_{8r}(x)$ tiene grupo de Galois $D_2$ de orden 16. Muestra también que el grupo de inercia es el grupo de cuaterniones de orden 8. El "Galois slope content" allí, $[2, 2, 5/2]^2$, indica la filtración de ramificación de $D_2$. En particular, la única posibilidad para la valuación 2-ádica de la conductor de $C_2$ es dos veces la mayor pendiente, a saber $2 \cdot (5/2) = 5$.

**17.7. Cálculos numéricos con funciones $L$.** Para curvas de género dos con grupo de Sato-Tate genérico, Conjetura 16.7 es desconocida. Notablemente, uno puede todavía calcular con una precisión muy alta. Hay un paquete muy útil en *Magma*, que viene de [25]. Nuestra curva $C_2$ da un ejemplo representativo. Dados los polinomios de Frobenius en (17.4), las posibilidades localmente permitidas por el conductor son $2^a3^b$ con $0 \leqslant a \leqslant 8$ y $0 \leqslant b \leqslant 5$, como en el caso de curvas elípticas. Usando `CFENew` de *Magma* para examinar todas las posibilidades da los siguientes números.

| $a\backslash b$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| 0 | 0.65071 | 0.53189 | 0.41151 | 0.29208 | 0.16978 | 0.02654 |
| 1 | 0.57620 | 0.45586 | 0.33611 | 0.21589 | 0.08433 | 0.10492 |
| 2 | 0.50034 | 0.38017 | 0.26069 | 0.13567 | 0.02104 | 0.37675 |
| 3 | 0.42438 | 0.30489 | 0.18337 | 0.04423 | 0.18654 | 3.82310 |
| 4 | 0.34890 | 0.22900 | 0.09975 | 0.07776 | 0.66956 | 0.62849 |
| 5 | 0.27357 | 0.14983 | 0.00069 | 0.30666 | 0.00000 | 0.23470 |
| 6 | 0.19678 | 0.06112 | 0.14992 | 1.69170 | 0.40104 | 0.07216 |
| 7 | 0.11473 | 0.05313 | 0.51913 | 0.79266 | 0.15720 | 0.04627 |
| 8 | 0.01843 | 0.25073 | 3.19710 | 0.27365 | 0.02061 | 0.15698 |

Estos números sugieren enfáticamente que el conductor correcto es $2^53^4$. Los escépticos podrían probablemente considerar todavía a $2^53^2$ como una posibilidad. Se puede calcular con más precisión así:

```
ZT<T>:=PolynomialRing(Integers());
CFENew(LSeries(C2:
    LocalData:=[<2,5,1+T+2*T^2>,<3,2,1+2*T+3*T^2>],Precision:=30));
```

Este código nos da

$$0.00069191183229691135350870962 1 \text{ para } 2^53^2 \text{ en } 0.39 \text{ segundos.}$$

Pero un cambio de $c_3 = 2$ a $c_3 = 4$ da

0.00000000000000000000000000000000 para $2^5 3^4$ en 1.26 segundos.

Una prueba de Conjetura 16.7 espera progreso en las conexiones con formas automorfas. Pero cálculos significativos ya son posibles, incluso en mayores dimensiones $g$.

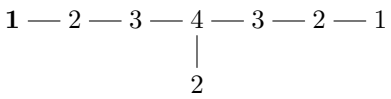
### 17.8. Ejercicios.

1. La curva

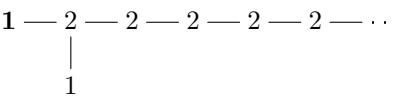
$$C_3 : y^2 + x^3 y = x^5 - 5x^3 - 10x^2 - 8x - 2 \tag{17.19}$$

tiene conductor $40000 = 2^6 5^4$ y está en la LMFDB. ¿Su punto móduli $(g_1, g_2, g_3)$ vive en la subvariedad $X_{1+2\mathbb{Z}^2}$ de (17.8)?.

2. La LMFDB reporta que el grupo de Sato-Tate $ST$ de (17.19) es $J(C_4)$. Tiene dimensión uno y grupo componente $C_2 \times C_4$. Más aún, la componente que contiene a $\mathrm{Fr}_p$ depende solo del símbolo del residuo cuadrático $(-8/p)$ y la clase de $p$ en $\mathbb{F}_5^\times$. Identificar la medida $\mu_{J(C_4)}$ en el escudo $Sp^\natural_{2g}$. (El soporte consiste en tres puntos especiales con medida $1/2$, y entonces tres curvas, también con medida $1/2$. Mucha orientación adicional se da en la página de $J(C_4)$.)
3. Pegue $E_1 : y^2 = x(x-1)(x+1)$ a una segunda copia de la misma curva dada por $E_2 : y^2 = x(x-1)(x-2)$, usando (17.7) para obtener una curva $C_4$. El discriminante de $C_4$ debería ser el número grande $\Delta = 2056589122535424 = 2^{16} 3^{22}$. ¿Por qué es `LocalData:=[<2,10,1>,<3,0,(1+T^2)^2>]` la información local correcta? Verifica que $C_4$ pasa `CFENew` a treinta decimales. (Fíjate que esta curva tiene grupo de Sato-Tate $C_{2,1}$ y ningún ejemplo de este grupo existe actualmente en la LMFDB. Pon atención también que hay un salto grande del gran discriminante $\Delta$ al pequeño conductor $N = 1024$.)
4. Consideremos la curva $C_5$ dada por $y^2 = x^6 - 1$. ¿Cuál es su discriminante? ¿Es isomorfo a $C_4$? ¿Parece su Jacobiana ser isógena a la Jacobiana de $C_4$?
5. Los momentos de $\tilde{S}_4$ se calculan por el diagrama extendido de Dynkin de tipo $\tilde{E}_7$:




A saber, $m_k$ es el número de paseos de longitud $k$ que empiezan y terminan al punto **1**. Verifica que los valores dados en (17.17) sean correctos. Repítelo para $G_2 = O_2$ y el diagrama extendido de Dynkin de tipo $\tilde{D}_\infty$:




¿Cuál es el significado de los números en los vértices?

## Referencias

[1] C. Birkenhake y H. Lange, *Complex Abelian varieties*. Springer-Verlag, second edition, 2004.
[2] G. Cornell y J. H. Silverman (ed.), *Arithmetic Geometry*. Springer-Verlag, 1986.
[3] O. Debarre, *Tores et variétés abéliennes complexes*. EDP Sciences, 1999.
[4] G. van de Geer y B. Moonen, *Notes on Abelian varieties*. preliminary notes accessible on: `http://www.mi.fu-berlin.de/users/elenalavanda/BMoonen.pdf`
[5] R. Hartshorne, *Algebraic geometry*. Springer-Verlag, 1977.

[6] M. Hindry y M. Rebolledo, *Introducción a la teoría de las curvas elípticas*. Notas de curso para AGRA II, Cusco 2015. Accessible `https://webusers.imj-prg.fr/~harald.helfgott/agraweb/AGRAIIMarcMarusia.pdf`
[7] M. Hindry y J. Silverman, *Diophantine Geometry. An introduction.* .Springer-Verlag, 2000.
[8] G. R. Kempf, *Complex abelian varieties and theta functions*. Universitext, Springer-Verlag, Berlin, 1991.
[9] J. Milne, *Abelian varieties*. In [2], 103–150.
[10] J. Milne, *Jacobian varieties*. In [2], 167–212.
[11] J. Milnor, *Curvatures of left invariant metrics on Lie groups*, Adv. Math. **21**:3 (1976), 293–329. DOI: 10.1016/S0001-8708(76)80002-3.
[12] D. Mumford, *Abelian varieties*. Oxford U. Press, 1970.
[13] D. Mumford, *Curves and Jacobians*. Univ. of Michigan, 1975.
[14] M. Rosen, *Abelian varieties over* $\mathbb{C}$. In [2], 76–102.
[15] J. Silverman, *Arithmetic of Elliptic Curves*. Springer-Verlag, 1986.
[16] J. Silverman, *Advanced Topics on the Arithmetic of Elliptic Curves*. Springer-Verlag, 1994.
[17] H. P. F. Swinnerton-Dyer, *Analytic theory of abelian varieties*. London Mathematical Society Lecture Note Series, No. 14. Cambridge University Press, London-New York, 1974.

REFERENCIAS MÁS ESPECIALIZADAS

[18] The LMFDB Collaboration – *The L-function and Modular Forms Data Base*, `http://www.lmfdb.org`, 2013.
[19] *Numerical tables on elliptic curves*. In "Modular functions of one variable, IV (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972)", Lecture Notes in Math. **476**, 74–144, Springer, Berlin, 1975.
[20] A.R. Booker, J. Sijsling, A.V. Sutherland, J. Voight, D. Yasaki, *A database of genus-2 curves over the rational numbers*. LMS J. Comput. Math. **19**, 235–254 (2016). DOI 10.1112/S146115701600019X.
[21] L. Clozel, M. Harris, R. Taylor, *Automorphy for some l-adic lifts of automorphic mód l Galois representations*. Publ. Math. Inst. Hautes Études Sci. **108**, 1–181 (2008). DOI 10.1007/s10240-008-0016-1.
[22] J.E. Cremona, *Algorithms for modular elliptic curves*. Cambridge University Press, Cambridge, 1997.
[23] P. Deligne, J. S. Milne, A. Ogus, K-y Shih. *Hodge cycles, motives, and Shimura varieties*. Lecture Notes in Mathematics, 900. Springer-Verlag, Berlin-New York, 1982.
[24] S.A. DiPippo, E.W. Howe, *Real polynomials with all roots on the unit circle and abelian varieties over finite fields*. J. Number Theory **73**, 426–450 (1998). DOI 10.1006/jnth.1998.2302.
[25] T. Dokchitser, *Computing special values of motivic L-functions* Experiment. Math. **13**, no. 2, 137–149 (2004).
[26] T. Dokchitser and C. Doris, *3-torsion and conductor of genus 2 curves*. Arxiv 1706.06162 (2018).
[27] G. Faltings, *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*. Invent. Math. **73**, 349–366 (1983). DOI 10.1007/BF01388432.
[28] F. Fité, A.V. Kedlaya, K.S. Sutherland, *Sato-Tate distributions and Galois endomorphism modules in genus 2*. Compos. Math. **148**, 1390–1442 (2012). DOI 10.1112/S0010437X12000279.
[29] B.H. Gross, D.B. Zagier, *On singular moduli*. J. Reine Angew. Math. **355**, 191–220 (1985).
[30] T. Honda, *Isogeny classes of abelian varieties over finite fields*. J. Math. Soc. Japan **20**, 83–95 (1968). DOI 10.2969/jmsj/02010083.
[31] J-I. Igusa, *Arithmetic variety of moduli for genus two*. Ann. of Math. **72** (1960) 612–649.
[32] J.W. Jones, D.P. Roberts, *A database of local fields*. J. Symbolic Comput. **41**, 80–97 (2006). DOI 10.1016/j.jsc.2005.09.003.
[33] N.M. Katz, *Larsen's alternative, moments, and the monodromy of Lefschetz pencils*. In "Contributions to automorphic forms, geometry, and number theory", 521–560, Johns Hopkins Univ. Press, Baltimore, MD, 2004.
[34] J.-F. Mestre, *Construction de courbes de genre* 2 *à partir de leurs modules*. In "Effective methods in algebraic geometry (Castiglioncello, 1990)", 313–334, Progr. Math. **94**, Birkhäuser Boston, Boston, MA, 1991.
[35] J.-P. Serre, *Propriétés galoisiennes des points d'ordre fini des courbes elliptiques*. Invent. Math. **15**, 259–331 (1972). DOI 10.1007/BF01405086.

[36] J.-P. Serre, J. Tate, *Good reduction of abelian varieties*. Annals of Math. **88**, 492–517 (1968).
[37] J. Tate, *Endomorphisms of abelian varieties over finite fields*. Invent. Math. **2**, 134–144 (1966). DOI 10.1007/BF01404549.
[38] W.C. Waterhouse, *Abelian varieties over finite fields*. Ann. Sci. École Norm. Sup. (4) **2**, 521–560 (1969).
[39] A. Wiles, *Modular elliptic curves and Fermat's last theorem*. Ann. of Math. (2) **141**, no. 3, 443–551 (1995).
[40] Yu. G. Zarhin, *A finiteness theorem for unpolarized abelian varieties over number fields with prescribed places of bad reduction*. Invent. Math. **79** (1985), 309–321.

Université Paris Diderot Paris 7, France
*Email address*: marc.hindry@imj-prg.fr

Université Clermont Auvergne, France
*Email address*: marusia.rebolledo@uca.fr

University of Minnesota, Morris, Minnesota 56267, USA
*Email address*: roberts@morris.umn.edu

Las Publicaciones Matemáticas del Uruguay (PMU) tienen como objetivo reflejar parte de las actividades de investigación matemática que se lleva a cabo en Uruguay. Nuestro interés es publicar artículos de investigación, así como artículos de tipo survey, anuncios, y otros trabajos que el Consejo Editor considere adecuado.

Los volúmenes no necesariamente serán arbitrados. Esto se indicará cuidadosamente en cada volumen.

---

Este volumen contiene notas de cursos.

---

The goal of Publicaciones Matemáticas del Uruguay (PMU) is to reflect part of the mathematics research activities taking place in Uruguay. It is our interest to publish research articles, survey-type articles, research announcements and other papers considered suitable by the Editorial Board.

The editorial process may or may not involve a revision by referees. This will be carefully indicated in each volume.

---

This volume contains course notes.

---

Publicaciones Matemáticas del Uruguay

Volumen 18 Diciembre 2023