Dpto. de Matemáticas, Facultad de Ciencias
Universidad Autónoma de Madrid

Cuerpos de números y el teorema de las S-unidades

Trabajo de Fin de Máster
Máster en Matemáticas y Aplicaciones

Raúl Segurado Herrera

dirigido por
Fernando Chamizo Lorente

Julio de 2024

Resumen

Este trabajo pretende ser una prueba no efectiva del teorema de las S-unidades. Para un cuerpo de números cualquiera, este resultado afirma que la ecuación $x + y = 1$ tiene una cantidad finita de soluciones para $x, y \in \mathcal{O}_S^*$, donde $\mathcal{O}_S^*$ denota el grupo de S-unidades. La razón de ser de este teorema es su gran cantidad de aplicaciones en el terreno de las ecuaciones diofánticas. Para poder probarlo, necesitaremos introducir multitud de herramientas algebraicas, entre las que destacan el anillo de enteros de un cuerpo de números, los conceptos de S-entero y S-unidad, y la noción de altura en un cuerpo de números. Además del teorema de las S-unidades, otros resultados de gran relevancia que veremos serán el Teorema de las Unidades de Dirichlet, el Teorema de Ostrowski para valores absolutos no arquimedianos y el Teorema de Northcott.

Abstract

This dissertation aims to be a non-effective proof of the S-units theorem. For any number field, this result states that the equation $x + y = 1$ has a finite number of solutions for $x, y \in \mathcal{O}_S^*$, where $\mathcal{O}_S^*$ denotes the group of S-units. The only purpose of this theorem is its vast number of applications in the area of Diophantine equations. To be able to prove it, we will need to introduce a multitude of algebraic tools, among which are the ring of integers of a number field, the concepts of S-integer and S-unit, and the notion of height in a number field. In addition to the S-units theorem, other highly relevant results we will see include Dirichlet's Unit Theorem, Ostrowski's Theorem for non-archimedean absolute values, and Northcott's Theorem.

Índice general

Introducción

La Teoría de Números y el Álgebra son dos de los mayores campos de estudio de las Matemáticas, y serán también los dos protagonistas de este trabajo. Nuestro objetivo principal será demostrar el teorema de las S-unidades, resultado clave en Teoría de Números y con infinidad de aplicaciones en áreas más específicas como las ecuaciones diofánticas. Nuestro acercamiento a este campo será casi completamente algebraico, y por lo tanto emplearemos multitud de herramientas relativas a diferentes estructuras algebraicas (grupos, anillos, cuerpos y módulos, principalmente) y de Teoría de Galois.

Este trabajo pretende ser dos cosas: una revisión de los contenidos elementales de Teoría Algebraica de Números y una demostración del teorema de las S-unidades. Si bien nuestro objetivo principal es acabar probando dicho teorema, será necesario introducir multitud de objetos de naturaleza algebraica con gran interés intrínseco y en los cuales nos detendremos lo necesario para comprenderlos en profundidad.

En el primer capítulo introducimos todos los conceptos elementales de Teoría Algebraica de Números que usaremos durante todo el trabajo. El objeto más importante será el anillo de enteros de un cuerpo de números K, que denotaremos por $\mathcal{O}_K$. Veremos en profundidad muchas de sus propiedades, y enunciaremos y demostraremos el Teorema de las Unidades de Dirichlet, que nos permitirá entender la estructura de $\mathcal{O}_K^*$. Para acabar el capítulo veremos algunas aplicaciones en casos específicos. En concreto analizaremos la estructura de todos los cuerpos cuadráticos y estudiaremos un ejemplo particular de cuerpo cúbico.

El segundo capítulo estará dedicado a los ideales primos del anillo $\mathcal{O}_K$, lo que se conoce comúnmente como su espectro. Veremos que existe una correspondencia entre los ideales primos no nulos de $\mathcal{O}_K$ y los valores absolutos no arquimedianos de K (teorema de Ostrowski). Introduciremos también los conceptos de S-entero y S-unidad, y enunciaremos el teorema de las S-unidades. Si bien todavía no estaremos preparados para probarlo, veremos algunas de sus aplicaciones en el terreno de las ecuaciones diofánticas.

Finalmente llegaremos al tercer y último capítulo. Nuestro objetivo será probar el resultado estrella del trabajo: el teorema de las S-unidades. Para ello, introduciremos el concepto de altura en un cuerpo de números junto con

muchas de sus propiedades. Requeriremos de ellas para ser capaces de dar una demostración completa del teorema. La prueba pasará por cierto resultado fundamental cuya demostración dejaremos para el final por escaparse un poco de los contenidos del trabajo.

Cabe destacar que la prueba que veremos del teorema de las S-unidades no es efectiva. Esto quiere decir que solo probaremos la finitud del número de soluciones de la ecuación $x + y = 1$ en $\mathcal{O}_S^*$, sin dar ningún tipo de cota. Históricamente la prueba original era de este tipo, pero actualmente existen otras que sí afinan más este resultado y son efectivas en este sentido.

Capítulo 1

Cuerpos de números y unidades

En este primer capítulo veremos todas las nociones elementales de Teoría Algebraica de Números a las que apelaremos durante todo el trabajo.

1.1. Cuerpos de números

Definición 1.1. Un *cuerpo de números* es cualquier extensión finita del cuerpo de los números racionales $\mathbb{Q}$.

Si $K/\mathbb{Q}$ es una extensión, un elemento $\alpha \in K$ se dice *algebraico* sobre $\mathbb{Q}$ si es raíz de un polinomio con coeficientes racionales. Si K es un cuerpo de números entonces todo elemento es algebraico sobre $\mathbb{Q}$. Nos interesaremos exclusivamente por un caso particular de elemento algebraico que definimos a continuación.

Definición 1.2. Diremos que un número $\alpha \in \mathbb{C}$ es un *entero algebraico* si existe un polinomio mónico $f \in \mathbb{Z}[x]$ con $f(\alpha) = 0$.

Recordamos también la siguiente definición.

Definición 1.3. Dado $\alpha \in \mathbb{C}$ un entero algebraico, el *polinomio mínimo* de α sobre $\mathbb{Q}$, es un polinomio mónico $f \in \mathbb{Q}[x]$ del menor grado posible que satisfaga $f(\alpha) = 0$.

Sabemos también que para cualquier elemento algebraico $\alpha \in \mathbb{C}$, el polinomio mínimo siempre existe y es único, además de ser irreducible y dividir a cualquier otro polinomio con coeficientes racionales que tenga a α como raíz. Esto se aplica también al caso de los enteros algebraicos.

Lema 1.4. *Si $\alpha \in \mathbb{C}$ es un entero algebraico, su polinomio mínimo tiene coeficientes enteros.*

Demostración. Sea $f \in \mathbb{Q}[x]$ el polinomio mínimo de α y sea $g \in \mathbb{Z}[x]$ mónico con $g(\alpha) = 0$. Sabemos entonces que existe $h \in \mathbb{Q}[x]$ con $g = fh$. Si $f \notin \mathbb{Z}[x]$, entonces existe un p primo que divide al denominador de alguno de los coeficientes de f.

Sea p^i la potencia más grande de p que divide a algún denominador de algún coeficiente de f. Igualmente definimos p^j como la potencia más grande de p que divida a algún denominador de algún coeficiente de h. Esto quiere decir que los coeficientes de $p^i f$ y los de $p^j h$ pertenecen al anillo local

$$\mathbb{Z}_{(p)} = \left\{ \frac{a}{b} \in \mathbb{Q} : p \nmid b \right\},$$

y podemos reducir sus coeficientes módulo p mediante $a/b \mapsto ab^{-1}\,(p)$. Luego tenemos la igualdad $p^{i+j}g = (p^i f)(p^j h)$, que vista en $\mathbb{F}_p[x]$ nos da 0 en el lado izquierdo y el producto de dos polinomios no nulos en el lado derecho, lo que resulta ser una contradicción. □

Proposición 1.5. *Un número $\alpha \in \mathbb{C}$ es entero algebraico si y solo si $\mathbb{Z}[\alpha]$ está finitamente generado como $\mathbb{Z}$-módulo.*

Demostración. Supongamos que α es entero algebraico y sea $f \in \mathbb{Z}[x]$ el polinomio mínimo de α (Lema 1.4). Entonces $\mathbb{Z}[\alpha]$ está generado por $1, \alpha, \ldots, \alpha^{d-1}$, donde d es el grado de f.

Por otro lado, supongamos que $\mathbb{Z}[\alpha]$ está generado por $f_1(\alpha), \ldots, f_n(\alpha)$. Sea d cualquier entero mayor que el grado de todos los f_i. Entonces existen enteros a_i tales que $\alpha^d = \sum_{i=1}^n a_i f_i(\alpha)$. Por tanto α es entero algebraico porque es raíz del polinomio mónico $x^d - \sum_{i=1}^n a_i f_i(x) \in \mathbb{Z}[x]$. □

Definición 1.6. El *anillo de enteros algebraicos*, que denotamos por $\overline{\mathbb{Z}}$, es el anillo formado por todos los $\alpha \in \mathbb{C}$ que son raíz de un polinomio mónico en $\mathbb{Z}[x]$.

Proposición 1.7. *El anillo de enteros $\overline{\mathbb{Z}}$ es, en efecto, un subanillo de $\mathbb{C}$.*

Demostración. Sean $\alpha, \beta \in \overline{\mathbb{Z}}$, y sean m, n los grados de sus polinomios mínimos, respectivamente. Entonces, por la Proposición 1.5, los elementos $1, \alpha, \ldots, \alpha^{m-1}$ generan $\mathbb{Z}[\alpha]$ y $1, \beta, \ldots, \beta^{n-1}$ generan $\mathbb{Z}[\beta]$, y por tanto $\alpha^i \beta^j$ para $i < m, j < n$ generan $\mathbb{Z}[\alpha, \beta]$.

Como $\mathbb{Z}[\alpha + \beta]$ y $\mathbb{Z}[\alpha\beta]$ son submódulos de $\mathbb{Z}[\alpha, \beta]$, ambos están finitamente generados, y de nuevo aplicando la Proposición 1.5, se tiene que $\alpha + \beta, \alpha\beta \in \overline{\mathbb{Z}}$. □

Definición 1.8. Dado un cuerpo de números K, su *anillo de enteros*, que denotaremos por $\mathcal{O}_K$, es el subanillo de $\mathbb{C}$ formado por todos los enteros algebraicos contenidos en K. Es decir,

$$\mathcal{O}_K = K \cap \overline{\mathbb{Z}} = \{x \in K : x \text{ es entero algebraico}\}.$$

En estas primeras secciones nos interesaremos especialmente por este subanillo de un cuerpo de números genérico K, más concretamente por el grupo de unidades $\mathcal{O}_K^*$. Veamos algunas de sus propiedades.

Lema 1.9. *Sea $\mathcal{O}_K$ el anillo de enteros de un cuerpo de números. Entonces $\mathcal{O}_K \cap \mathbb{Q} = \mathbb{Z}$ y $\mathbb{Q}\mathcal{O}_K = K$.*

Demostración. Supongamos $\alpha \in \mathcal{O}_K \cap \mathbb{Q}$ con $\alpha = a/b$ irreducible y $b > 0$. Entonces el polinomio mínimo de α es $x - a/b$, y como α es entero algebraico se tiene que $b = 1$. Por tanto $\alpha \in \mathbb{Z}$. La otra inclusión es trivial.

El contenido $\mathbb{Q}\mathcal{O}_K \subset K$ también es evidente. Si $\alpha \in K$, sea $f \in \mathbb{Q}[x]$ su polinomio mínimo. Para cualquier entero d, el polinomio mínimo de $d\alpha$ es $d^{\deg(f)} f(x/d)$. Si d es el mínimo común múltiplo de los denominadores de los coeficientes de f, el polinomio mínimo de $d\alpha$ tiene coeficientes enteros, y por tanto $d\alpha \in \mathcal{O}_K$. Por consiguiente $\alpha = (1/d)d\alpha \in \mathbb{Q}\mathcal{O}_K$. □

Sea K un cuerpo de números. En particular, la extensión $K/\mathbb{Q}$ es finita y separable, así que por el Teorema del Elemento Primitivo existe $\gamma \in K$ con $K = \mathbb{Q}(\gamma)$. Digamos que $n = [K : \mathbb{Q}]$ y sean $\gamma_1, \ldots, \gamma_n$, con $\gamma = \gamma_1$, las raíces del polinomio mínimo de γ sobre $\mathbb{Q}$. Entonces existen exactamente n $\mathbb{Q}-$automorfismos del cierre normal de K (y no necesariamente de K si la extensión $K/\mathbb{Q}$ no es de Galois), $\sigma_1, \ldots, \sigma_n$, que satisfacen $\sigma_i(\gamma) = \gamma_i$ (en particular, $\sigma_1 = \mathrm{id}$). Esto nos permite dar la siguiente noción de conjugación, que depende de la extensión y no solo del elemento.

Definición 1.10. Dados $K = \mathbb{Q}(\gamma)$ un cuerpo de números y $\alpha \in K$ un elemento cualquiera, los $\mathbb{Q}-$*conjugados* de α (o simplemente *conjugados*) son las imágenes de α por los automorfismos $\sigma_1, \ldots, \sigma_n$. Los denotaremos por $\alpha^{(1)}, \ldots, \alpha^{(n)}$.

Observación. Si $\alpha \in K$ es entero algebraico y β es uno de sus conjugados, entonces β también es entero algebraico por satisfacer la misma ecuación entera que α.

Algunos de estos automorfismos $\sigma_1, \ldots, \sigma_n$ pueden ser emparejados como conjugados complejos uno del otro. De ahora en adelante, dado $\alpha \in K$ un entero algebraico, denotaremos por r_1 al número de $\mathbb{Q}$-conjugados reales, y por r_2 al número de pares de $\mathbb{Q}$-conjugados complejos. De esta forma, cualquier entero algebraico tendrá $n = r_1 + 2r_2$ $\mathbb{Q}$-conjugados.

Elegimos (y renombramos) los r_1 automorfismos reales, $\sigma^{(1)}, \ldots, \sigma^{(r_1)}$, y r_2 automorfismos complejos, $\sigma^{(r_1+1)}, \ldots, \sigma^{(r_1+r_2)}$, uno de cada par de conjugados. Conservamos el convenio $\sigma^{(1)} = \mathrm{id}$ incluso cuando $r_1 = 0$. Si $\alpha^{(j)} = \sigma^{(j)}(\alpha)$, para $j = 1, \ldots, r_1 + r_2$, definimos la aplicación

$$X : K \to \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} \cong \mathbb{R}^n$$

dada por $X(\alpha) = (\alpha^{(1)}, \dots, \alpha^{(r_1)}, \alpha^{(r_1+1)}, \dots, \alpha^{(r_1+r_2)})$. En particular, los conjugados de α son

$$\alpha^{(1)}, \dots, \alpha^{(r_1)}, \alpha^{(r_1+1)}, \dots, \alpha^{(r_1+r_2)}, \overline{\alpha^{(r_1+1)}}, \dots, \overline{\alpha^{(r_1+r_2)}}.$$

Como los $\sigma^{(j)}$ son morfismos, la función X es lineal vista como aplicación entre $\mathbb{Q}$-espacios vectoriales, y es trivialmente inyectiva pues $\sigma^{(1)} = \mathrm{id}$.

Definición 1.11. Si K es un cuerpo de números, dado $\alpha \in K$ un elemento cualquiera, la *norma* de α, que denotamos por $N_{K/\mathbb{Q}}(\alpha)$, es el producto de todos sus conjugados. Es decir, con la notación anterior,

$$N_{K/\mathbb{Q}}(\alpha) = \alpha^{(1)} \cdots \alpha^{(r_1)} |\alpha^{(r_1+1)}|^2 \cdots |\alpha^{(r_1+r_2)}|^2$$

Computacionalmente hablando, nos resulta más útil la definición anterior. No obstante, una forma equivalente de definir la norma (y más habitual en general) es considerar la aplicación *multiplicación por* α en K:

$$\begin{array}{rccc} m_\alpha : & K & \to & K \\ & \beta & \mapsto & \alpha\beta \end{array}$$

que es un endomorfismo del $\mathbb{Q}$-espacio vectorial K. Definimos entonces la norma $N_{K/\mathbb{Q}}(\alpha)$ como el determinante de la aplicación m_α.

Ejemplo 1.12. Sean $\alpha = \sqrt[3]{2}$ y $\omega = e^{i2\pi/3}$. Consideramos primero el cuerpo $K = \mathbb{Q}(\alpha)$, para el cual sabemos que $[K : \mathbb{Q}] = 3$. Como las raíces del polinomio mínimo de α son $\alpha, \alpha\omega, \alpha\omega^2$, en este caso tenemos $r_1 = r_2 = 1$, $X(\alpha) = (\alpha, \alpha\omega)$ y $N_{K/\mathbb{Q}}(\alpha) = \alpha|\alpha\omega|^2 = 2$.

Sin embargo, si consideramos ahora el cuerpo $\tilde{K} = \mathbb{Q}(\alpha, \omega)$ los conjugados y la norma cambian. En este caso, $[\tilde{K} : \mathbb{Q}] = 6$, y se tiene $r_1 = 0$, $r_2 = 3$, $X(\alpha) = (\alpha, \alpha\omega, \alpha\omega^2)$ y $N_{\tilde{K}/\mathbb{Q}}(\alpha) = |\alpha|^2|\alpha\omega|^2|\alpha\omega^2|^2 = \alpha^6 = 4$.

Observación. Para $\alpha \in K$, $N_{K/\mathbb{Q}}(\alpha)$ es, salvo un signo y elevar a una potencia, el término independiente del polinomio mínimo de α sobre $\mathbb{Q}$; en particular, $N_{K/\mathbb{Q}}(\alpha) \in \mathbb{Q}$, y si α es entero algebraico entonces $N_{K/\mathbb{Q}}(\alpha) \in \mathbb{Z}$.

Definición 1.13. Sea $\Gamma \subset \mathbb{R}^n$ un $\mathbb{Z}-$módulo (es decir, un grupo abeliano). Diremos que Γ es un *subretículo* (o *sublattice*) de $\mathbb{R}^n$ si es un $\mathbb{Z}-$módulo libre finitamente generado, esto es, si existe un conjunto de vectores linealmente independientes $\{v_1, \dots, v_m\}$ tales que

$$\Gamma = \{k_1 v_1 + \cdots + k_m v_m : k_j \in \mathbb{Z}\}.$$

Si la dimensión de Γ como $\mathbb{Z}-$módulo es igual que la dimensión de $\mathbb{R}^n$, esto es, si $m = n$, entonces diremos que es un *retículo* (o *lattice*) de $\mathbb{R}^n$.

Notación. Si $\{v_1, \dots, v_m\}$ es una base libre de Γ (y por tanto lo genera) escribiremos $\Gamma = [v_1, \dots, v_m]$.

Observación. De igual forma se pueden definir retículos y subretículos en un subespacio vectorial V de $\mathbb{R}^n$. Si $m < n$, claramente $\Gamma = [v_1, \dots, v_m] \subset \mathbb{R}^n$ no es un retículo de $\mathbb{R}^n$. Sin embargo, sí es un retículo de $\langle v_1, \dots, v_m \rangle$.

Introducimos esta noción porque nuestro siguiente objetivo es probar que $X(\mathcal{O}_K)$ es un retículo en $\mathbb{R}^n$, $n = r_1 + 2r_2$. Para ello comenzamos definiendo, para $\alpha, \alpha_1, \dots, \alpha_n \in K$, la aplicación *traza*

$$t_{K/\mathbb{Q}}(\alpha) := \alpha^{(1)} + \cdots + \alpha^{(n)},$$

que es la suma de todos los conjugados, y la aplicación *discriminante*

$$d_{K/\mathbb{Q}}(\alpha_1, \dots, \alpha_n) := \left(\det[\alpha_j^{(i)}] \right)^2,$$

que se relacionan mediante la expresión

$$d_{K/\mathbb{Q}}(\alpha_1, \dots, \alpha_n) = \det \left[t_{K/\mathbb{Q}}(\alpha_i \alpha_j) \right].$$

De nuevo, y al igual que con la norma, se tiene que $t_{K/\mathbb{Q}}(\alpha)$ es la traza de la aplicación multiplicar por α, m_α, y que $t_{K/\mathbb{Q}}(\alpha) \in \mathbb{Q}$ para todo $\alpha \in K$. Esto es por ser, salvo elevar a una potencia, uno de los coeficientes del polinomio mínimo de α sobre $\mathbb{Q}$; en particular, si $\alpha \in \mathcal{O}_K$, se tiene que $t_{K/\mathbb{Q}}(\alpha) \in \mathbb{Z}$, y si $\alpha_1, \dots, \alpha_n \in \mathcal{O}_K$, entonces $d_{K/\mathbb{Q}}(\alpha_1, \dots, \alpha_n) \in \mathbb{Z}$.

Por otro lado, si $\beta \in K$, sabemos, por el Lema 1.9, que existen $m'/m \in \mathbb{Q}$ y $\alpha \in \mathcal{O}_K$ tales que

$$\beta = (m'/m)\alpha \Leftrightarrow m\beta = m'\alpha \in \mathcal{O}_K,$$

por ser $\mathcal{O}_K$ un anillo. Esto implica que, dado cualquier $\beta \in K$, existe un $m \in \mathbb{Z}$ con $m\beta \in \mathcal{O}_K$. En particular, dada $\{\beta_1, \dots, \beta_n\}$ una base de K como $\mathbb{Q}-$espacio vectorial, podemos encontrar otra base $\{\alpha_1, \dots, \alpha_n\}$ con $\alpha_j = m_j \beta_j$, para algunos $m_j \in \mathbb{Z}$, tales que $\alpha_j \in \mathcal{O}_K$. Todo esto nos permite probar el siguiente resultado.

Lema 1.14. *El anillo $\mathcal{O}_K$ es un $\mathbb{Z}-$módulo libre finitamente generado, esto es, existen $\alpha_1, \dots, \alpha_n \in \mathcal{O}_K$ con $n = [K : \mathbb{Q}]$ tales que*

$$\mathcal{O}_K = \{k_1\alpha_1 + \cdots + k_n\alpha_n : k_j \in \mathbb{Z}\}.$$

Demostración. De todas las bases $\{\alpha_1, \dots, \alpha_n\}$ de K que satisfacen $\alpha_j \in \mathcal{O}_K$ (que sabemos que existen), elegimos una para la cual $|d_{K/\mathbb{Q}}(\alpha_1, \dots, \alpha_n)| \in \mathbb{N}$ sea mínimo.

Dado un $\alpha \in \mathcal{O}_K$ arbitrario, sabemos que existen unos únicos coeficientes $a_1, \dots, a_n \in \mathbb{Q}$ tales que $\alpha = a_1\alpha_1 + \cdots + a_n\alpha_n$. Queremos ver que todos los a_i son enteros. Supongamos que uno no lo es; sin pérdida de generalidad, asumimos $a_1 \notin \mathbb{Z}$. Entonces $a_1 = m + \theta$, donde $m \in \mathbb{Z}$ y $0 < \theta < 1$. Construimos una nueva base:

$$\beta_1 = \alpha - m\alpha_1, \beta_2 = \alpha_2, \dots, \beta_n = \alpha_n.$$

Como $\beta_1 = \theta\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n$, tenemos

$$(\beta_1\ \beta_2\ \ldots\ \beta_n) = (\alpha_1\ \alpha_2\ \ldots\ \alpha_n)\begin{pmatrix} \theta & 0 & 0 & \cdots & 0 \\ a_2 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_n & 0 & 0 & \cdots & 1 \end{pmatrix},$$

y por tanto

$$|d_{K/\mathbb{Q}}(\beta_1, \ldots, \beta_n)| = \theta^2 |d_{K/\mathbb{Q}}(\alpha_1, \ldots, \alpha_n)| < |d_{K/\mathbb{Q}}(\alpha_1, \ldots, \alpha_n)|,$$

lo que contradice la minimalidad de $\{\alpha_1, \ldots, \alpha_n\}$. □

Corolario 1.15. *$X(\mathcal{O}_K)$ es un retículo en $\mathbb{R}^n$, $n = r_1 + 2r_2$.*

Demostración. Como consecuencia del Lema 1.14, dado $\alpha \in \mathcal{O}_K$ existen enteros $k_1, \ldots, k_n$ tales que

$$\alpha = k_1\alpha_1 + \cdots + k_n\alpha_n.$$

Entonces $X(\alpha) = k_1X(\alpha_1) + \cdots + k_nX(\alpha_n)$, y se tiene que los vectores $X(\alpha_1), \ldots, X(\alpha_n)$ generan todo $X(\mathcal{O}_K)$. Como además X es inyectiva, son linealmente independientes, y concluimos $X(\mathcal{O}_K) = [X(\alpha_1), \ldots, X(\alpha_n)]$. □

Nuestro siguiente objetivo es caracterizar los retículos (y subretículos) de $\mathbb{R}^n$. Introducimos primero un par de definiciones.

Definición 1.16. Para un $\mathbb{Z}-$módulo Γ de $\mathbb{R}^n$, si existe un conjunto acotado $C \subset \mathbb{R}^n$ tal que $\mathbb{R}^n = C + \Gamma$ decimos que el grupo cociente $\mathbb{R}^n/\Gamma$ es *compacto*.

Observación. Esta definición es equivalente a la compacidad de $\mathbb{R}^n/\Gamma$ como espacio topológico cociente.

Definición 1.17. Un conjunto no vacío $M \subset \mathbb{R}^n$ se dice *discreto* si para todo $x \in M$ existe un $\delta > 0$ tal que $B_\delta(x) \cap M = \{x\}$, o equivalentemente, si la topología inducida en M es la discreta.

Lema 1.18. *Para Γ un $\mathbb{Z}-$módulo en $\mathbb{R}^n$, se tiene*

$$\Gamma \text{ es discreto} \Leftrightarrow B_r(0) \cap \Gamma \text{ es finito para todo } r > 0.$$

Demostración. ($\Leftarrow$) Sea $x \in \Gamma$, entonces existe un $r > 0$ tal que $x \in B_r(0)$. Como el conjunto $B_r(0) \cap \Gamma$ es finito, podemos tomar

$$\delta < \min\{\operatorname{dist}(x, y) : y \in B_r(0) \cap \Gamma, y \neq x\} \cup \{\operatorname{dist}(x, \partial B_r(0))\},$$

de tal manera que $B_\delta(x) \subset B_r(0)$ y $B_\delta(x) \cap \Gamma = \{x\}$. Por tanto, Γ es discreto. ($\Rightarrow$) Supongamos ahora que Γ es un conjunto discreto. Veamos primero que

es cerrado. Si no fuera así, existiría un $x \in \overline{\Gamma}\backslash\Gamma$ y una sucesión $(x_n) \subset \Gamma$ con $\lim_{n\to\infty} x_n = x$. Esto quiere decir que para $\varepsilon > 0$ existe un N entero positivo tal que si $n, m > N$ se tiene $|x_n - x_m| < \varepsilon$. Como Γ es un $\mathbb{Z}-$módulo, $x_n - x_m \in \Gamma$, y tenemos que $B_\varepsilon(0) \cap \Gamma \supsetneq \{0\}$, lo que contradice que Γ sea discreto.

Por tanto Γ es cerrado y el conjunto $\overline{B_r(0)} \cap \Gamma$ es discreto y compacto. Esto implica que es finito, y que también lo es $B_r(0) \cap \Gamma$. □

Proposición 1.19. *Sea Γ un $\mathbb{Z}-$módulo en $\mathbb{R}^n$.*

(1) *Γ es subretículo $\Leftrightarrow$ Γ es discreto.*

(2) *Γ es retículo $\Leftrightarrow$ Γ es discreto y $\mathbb{R}^n/\Gamma$ es compacto.*

Demostración. (1): ($\Rightarrow$) Si $m \leq n$ es la dimensión de Γ, sea $\{v_i\}, i = 1, \ldots, n$ una base de $\mathbb{R}^n$ de tal forma que $\Gamma = [v_1, \ldots, v_m]$. Como Γ es un subconjunto de $\Gamma' = [v_1, \ldots, v_n]$, si Γ' fuera discreto entonces Γ también lo sería. Así que, sin pérdida de generalidad, asumimos $m = n$.

Sea ahora $\{v_i^*\}$ la base dual de $\{v_i\}$ en $\mathbb{R}^n$ dada por $\langle v_i^*, v_j\rangle = \delta_{ij}$. Para $x \in \Gamma$ escribimos $x = \sum a_j v_j$ con $a_i \in \mathbb{Z}$, y tenemos $\langle v_i^*, x\rangle = a_i$, $|a_i| = |\langle v_i^*, x\rangle| \leq |v_i^*||x|$. Si además $x \in \Gamma \cap B_r(0)$ para $r > 0$, se tiene $|a_i| \leq r|v_i^*|$, y el número de $a_i's$ posible es finito. Por tanto, el conjunto $\Gamma \cap B_r(0)$ también es finito, y por el Lema 1.18 el conjunto Γ es discreto.

($\Leftarrow$) Sea Γ un $\mathbb{Z}$-módulo discreto en $\mathbb{R}^n$ y sea L el subespacio vectorial de $\mathbb{R}^n$ formado por todas las combinaciones lineales de elementos de Γ. Sea m la $\mathbb{R}$-dimensión de L y escribimos $L = \langle v_1, \ldots, v_m\rangle$, $v_i \in \Gamma$. Sea $\Gamma_0 = [v_1, \ldots, v_m] \subset \Gamma$, queremos ver que $[\Gamma : \Gamma_0] < \infty$. Para vectores independientes $v_1, \ldots, v_m$, definimos primero el *paralelotopo* generado por los v_i como

$$\Pi = \Pi(v_1, \ldots, v_m) := \left\{ v = \sum_{i=1}^{m} \alpha_i v_i \in \mathbb{R}^n : 0 \leq \alpha_i < 1 \right\}.$$

Como Π está acotado en $\mathbb{R}^n$, existe un $r > 0$ con $\Pi \subset B_r(0)$. Si $x \in \Gamma \subset L$, escribimos $x = \sum \beta_i v_i$, $\beta_i \in \mathbb{R}$. Por tanto

$$x = \sum_{i=1}^{m} (\lfloor \beta_i \rfloor + \{\beta_i\}) v_i = \sum_{i=1}^{m} \lfloor \beta_i \rfloor v_i + \sum_{i=1}^{m} \{\beta_i\} v_i \in \Gamma_0 + \Pi,$$

y usando que $\Gamma_0 \subset \Gamma$,

$$\sum_{i=1}^{m} \{\beta_i\} v_i = x - \sum_{i=1}^{m} \lfloor \beta_i \rfloor v_i \in \Gamma \cap \Pi \subset \Gamma \cap B_r(0).$$

Como el conjunto $\Gamma \cap B_r(0)$ es finito (Lema 1.18), se tiene que cualquier $x \in \Gamma$ es equivalente a un conjunto finito de vectores en Γ módulo Γ_0, por lo

que $[\Gamma : \Gamma_0] = N < \infty$. Si $x \in \Gamma$ tenemos que $Nx \in \Gamma_0$, y por tanto

$$\Gamma \subset N^{-1}\Gamma_0 = [v_1/N, \ldots, v_m/N].$$

Como $N^{-1}\Gamma_0$ es un $\mathbb{Z}-$módulo libre de dimensión m ($N^{-1}\Gamma_0 \cong \mathbb{Z}^m$), Γ es un $\mathbb{Z}-$módulo libre de dimensión $k \leq m$, $\Gamma = [u_1, \ldots, u_k]$. No obstante, $\Gamma \supset \Gamma_0 = [v_1, \ldots, v_m]$, necesariamente $k = m$, y Γ es un subretículo de dimensión m.

(2): ($\Rightarrow$) Sea $\Gamma = [v_1, \ldots, v_n]$ un retículo. Para $x \in \mathbb{R}^n$ escribimos

$$x = \sum_{i=1}^{n} \lambda_i v_i = \sum_{i=1}^{n} \lfloor \lambda_i \rfloor v_i + \sum_{i=1}^{n} \{\lambda_i\} v_i \in \Gamma + \Pi,$$

donde $\Pi = \Pi(v_1, \ldots, v_n)$ es un paralelotopo, que claramente es acotado, y por tanto el cociente $\mathbb{R}^n/\Gamma$ es compacto.

($\Leftarrow$) Supongamos que Γ es un $\mathbb{Z}-$módulo discreto y que $\mathbb{R}^n/\Gamma$ es compacto. Por (1) tenemos que Γ es subretículo, $\Gamma = [v_1, \ldots, v_m]$. Queremos probar que $m = n$, así que supongamos que $m < n$. Queremos ver que $\mathbb{R}^n \supsetneq C + \Gamma$ para cualquier conjunto acotado $C \subset \mathbb{R}^n$. Para un C arbitrario, existe un $r > 0$ con $C \subset B_r(0)$. Sea $L = \langle v_1, \ldots, v_m \rangle$, el $\mathbb{R}-$subespacio vectorial generado por los v_i, y consideramos su subespacio ortogonal

$$L^{\perp} = \{v \in \mathbb{R}^n : \langle v, x \rangle = 0, \forall x \in L\},$$

de dimensión $n - m > 0$. Entonces un vector $v \in L^{\perp}$ con $|v| > r$ no puede pertenecer al conjunto $C + \Gamma$. De hecho, si $v \in C + \Gamma$, $v = c + \gamma$ con $c \in C$, $\gamma \in \Gamma$, y

$$|v|^2 = \langle v, v \rangle = \langle v, c + \gamma \rangle = \langle v, c \rangle \leq |v||b| \leq r|v|,$$

o, equivalentemente, $|v| \leq r$, una contradicción, así que $\mathbb{R}^n/\Gamma$ no es compacto si $m < n$. □

En relación a esta prueba es habitual definir el *volumen* de un retículo $\Gamma = [v_1, \ldots, v_n]$ como el volumen del paralelotopo $\Pi(v_1, \ldots, v_n)$, es decir,

$$\mathrm{vol}(\Gamma) := \mathrm{vol}(\Pi(v_1, \ldots, v_n)) = \det[v_1, \ldots, v_n].$$

El último resultado clásico que necesitaremos en relación a los retículos es el siguiente:

Teorema 1.20 (Minkowski)**.** *Sea $\Gamma \subset \mathbb{R}^n$ un retículo y sea $B \subset \mathbb{R}^n$ un subconjunto simétrico*[1] *y convexo. Si $\mathrm{vol}(B) > 2^n \mathrm{vol}(\Gamma)$ entonces B contiene algún punto de Γ distinto de 0.*

Demostración. Ver [15, Theorem 3.1]. □

[1] Simétrico quiere decir que $x \in B$ si y solo si $-x \in B$.

1.2. Teorema de las Unidades de Dirichlet

Sea K un cuerpo de números. Nuestro objetivo es entender la estructura del grupo de unidades del anillo de enteros de K, esto es $\mathcal{O}_K^*$. En esta sección seguiremos los contenidos de [15, Ch. 3.3].

Consideramos el conjunto $E^{s,t} := \mathbb{R}^s \times \mathbb{C}^t$. Con la multiplicación coordenada a coordenada, $E^{s,t}$ tiene una estructura natural de $\mathbb{R}$-álgebra conmutativa de dimensión $s + 2t$, cuyo grupo de unidades es

$$(E^{s,t})^* = (\mathbb{R}^\times)^s \times (\mathbb{C}^\times)^t.$$

Definimos la aplicación $l : (E^{s,t})^* \to \mathbb{R}^{s+t}$ dada por

$$l(x) = (l_1(x), \dots, l_s(x), l_{s+1}(x), \dots, l_{s+t}(x)),$$

donde

$$\begin{cases} l_i(x) = \log |x_i|, & 1 \leq i \leq s, \\ l_i(x) = \log |x_i|^2, & s+1 \leq i \leq s+t. \end{cases}$$

Proposición 1.21. *La aplicación l es un homomorfismo de grupos sobreyectivo cuyo núcleo es $\{\pm 1\}^s \times \mathbb{D}^t$, donde $\mathbb{D} = \{z \in \mathbb{C} : |z| = 1\}$.*

Demostración. De manera inmediata observamos que $l(xy) = l(x)+l(y)$ para todo $x, y \in E^{s,t}$, por lo que l es un homomorfismo.

Dado $a = (a_i) \in \mathbb{R}^{s+t}$, tomamos $x = (x_i) \in (E^{s,t})^*$ con $x_i = e^{a_i}$ para $1 \leq i \leq s$, y $x_i = e^{a_i/2}$ para $s + 1 \leq i \leq s + t$. De esta manera tenemos $l(x) = a$, y por tanto l es sobreyectiva.

La afirmación acerca del núcleo resulta trivial, pues si $\log |x| = 0$ entonces $|x| = 1 \Leftrightarrow x = \pm 1$ para $x \in \mathbb{R}$, y si $\log |z|^2 = 0$ entonces $|z|^2 = 1 \Leftrightarrow z \in \mathbb{D}$ para $z \in \mathbb{C}$. □

Sea K un cuerpo de números, y sea $n = [K : \mathbb{Q}]$. Podemos considerar de nuevo la aplicación $X : K \to E^{r_1,r_2}$, y combinarla con la aplicación l, obteniendo una nueva aplicación $l^* := l \circ X : K^\times \to \mathbb{R}^{r_1+r_2}$ dada por

$$l^*(\alpha) = (\log |\alpha^{(1)}|, \dots, \log |\alpha^{(r_1)}|, \log |\alpha^{(r_1+1)}|^2, \dots, \log |\alpha^{(r_1+r_2)}|^2).$$

Si denotamos por $l_i^*(\alpha)$ a la coordenada i-ésima de $l^*(\alpha)$, $1 \leq i \leq r_1 + r_2$, tenemos

$$\sum_{i=1}^{r_1+r_2} l_i^*(\alpha) = \log |N_{K/\mathbb{Q}}\alpha|.$$

Consideramos el hiperplano de dimensión $r = r_1 + r_2 - 1$

$$H^r := \left\{ v = (v_1, \dots, v_{r_1+r_2}) \in \mathbb{R}^{r_1+r_2} : \sum_{i=1}^{r_1+r_2} v_i = 0 \right\}.$$

Como $N_{K/\mathbb{Q}}(\alpha) = \pm 1$ para todo $\alpha \in \mathcal{O}_K^*$, tenemos que $l^*(\mathcal{O}_K^*) \subset H^r$. Denotamos por $l^\#$ la restricción de l^* en el subgrupo $\mathcal{O}_K^*$ de $K^\times$. Esto nos da un homomorfismo

$$l^\# : \mathcal{O}_K^* \to H^r.$$

Proposición 1.22. $\ker l^\# = W_K := \{\alpha \in \mathcal{O}_K^* : \alpha \text{ tiene orden finito}\}$.

Demostración. Tomamos $\alpha \in W_K$. Como $\alpha^n = 1$ para algún $n \in \mathbb{N}$, tenemos que $|\alpha^{(i)}| = 1$ para todo $1 \leq i \leq r_1 + r_2$. Por tanto, $l^\#(\alpha) = 0$, y $\alpha \in \ker l^\#$.

Para la otra inclusión, tomamos ahora $\alpha \in \ker l^\#$, es decir $l^\#(\alpha) = 0$. Entonces $|\alpha^{(i)}| = 1$ para todo $1 \leq i \leq r_1 + r_2$, lo que implica que el conjunto $X(\ker l^\#)$ está acotado en $\mathbb{R}^n \cong E^{r_1,r_2}$, $n = r_1 + 2r_2$. Por otro lado, como $X(\mathcal{O}_K)$ es un retículo en $\mathbb{R}^n$ (Corolario 1.15), es discreto en $\mathbb{R}^n$, y por tanto también lo es el subconjunto $X(\ker l^\#)$. Como $X(\ker l^\#)$ es finito y X es inyectiva, $\ker l^\#$ es un grupo finito y por tanto $\ker l^\# \subset W_K$. □

Lema 1.23. *Sea K un cuerpo. Si G es un subgrupo finito de $K^\times$, G es cíclico.*

Demostración. Supongamos $|G| = n$ y sea $d \in \mathbb{N}$ el orden maximal de todos los elementos de G. Si $g \in G$ cumple $\mathrm{o}(g) = d$, entonces para cualquier otro $x \in G$ se tiene $\mathrm{o}(x) \mid d$. Si esto no fuera así, entonces existiría un x con $\mathrm{mcm}(\mathrm{o}(x), d) > d$, lo que implicaría $\mathrm{o}(xg) > d$, contradiciendo así la maximalidad de d.

Por tanto $x^d = 1$ para todo $x \in G$, y el polinomio $x^d - 1 \in K[x]$ tiene n raíces en K. Como K es un cuerpo y el polinomio tiene grado d, a lo sumo tiene d raíces, así que $d = n$ y G es cíclico. □

Observación. Por el Lema 1.23, el grupo W_K es cíclico. Como $-1 \in W_K$, el orden de W_K es par y W_K son en realidad todas las raíces de la unidad que haya en K.

Proposición 1.24. $\mathrm{Im}\, l^\#$ *es un subretículo de H^r.*

Demostración. Como $\mathrm{Im}\, l^\# \subset H^r$, por la Proposición 1.19 basta con probar que $\mathrm{Im}\, l^\#$ es discreto en $\mathbb{R}^{r_1+r_2}$. Para ello probaremos que, para todo $r > 0$, el conjunto $\mathrm{Im}\, l^\# \cap B_r(0)$ es finito. Tomamos un vector $l^\#(\alpha) \in \mathrm{Im}\, l^\# \cap B_r(0)$, entonces

$$(\log|\alpha^{(1)}|)^2 + \cdots + (\log|\alpha^{(r_1+r_2)}|^2)^2 < r^2,$$

y por tanto

$$|\alpha^{(1)}| < e^r, \ldots, |\alpha^{(r_1)}| < e^r, |\alpha^{(r_1+1)}|^2 < e^r, \ldots, |\alpha^{(r_1+r_2)}|^2 < e^r.$$

Esto implica que el conjunto

$$S = \left\{ X(\alpha) : \alpha \in \mathcal{O}_K^*, l^\#(\alpha) \in \mathrm{Im}\, l^\# \cap B_r(0) \right\}$$

está acotado en $E^{r_1,r_2} = \mathbb{R}^n$, $n = r_1 + 2r_2$. Por otro lado, $X(\mathcal{O}_K)$ es un retículo en $\mathbb{R}^n$ (Corolario 1.15) y el subconjunto S es discreto. Por tanto S es finito, y como X es inyectiva, tenemos finalmente que $\operatorname{Im} l^{\#} \cap B_r(0)$ es finito. □

Proposición 1.25. *Sea m un número natural. Existen $\alpha_1, \ldots, \alpha_N$ en $\mathcal{O}_K$ con $|N_{K/\mathbb{Q}}\alpha_i| = m, 1 \leq i \leq N$ tales que cualquier $\alpha \in \mathcal{O}_K$ con $|N_{K/\mathbb{Q}}\alpha| = m$ puede expresarse como $\alpha = \varepsilon\alpha_i$ para algún i fijo y algún $\varepsilon \in \mathcal{O}_K^*$.*

Demostración. Si escribimos $\mathcal{O}_K = [\omega_1, \ldots, \omega_n]$, entonces el conjunto $m\mathcal{O}_K$ es un subanillo de $\mathcal{O}_K$ y $[\mathcal{O}_K : m\mathcal{O}_K] = m^n$.

Consideramos ahora el conjunto $\{\alpha \in \mathcal{O}_K : |N_{K/\mathbb{Q}}\alpha| = m\}$. Si lo vemos módulo el subanillo $m\mathcal{O}_K$ tendrá una cantidad finita de representantes, digamos $\alpha_1, \ldots, \alpha_N$. Si tomamos un α genérico de nuestro conjunto, sabemos que existe un i tal que $\alpha \equiv \alpha_i\,(m\mathcal{O}_K)$.

Ahora queremos probar que α/α_i es una unidad de K. Sabemos que existe un $\beta \in \mathcal{O}_K$ tal que $\alpha - \alpha_i = \beta m$, y entonces

$$\alpha/\alpha_i = 1 + \beta(m/\alpha_i) = 1 \pm \beta(N_{K/\mathbb{Q}}\alpha_i/\alpha_i) = 1 \pm \beta\alpha_i^{(2)} \ldots \alpha_i^{(r_1+r_2)} \in \mathcal{O}_K.$$

Intercambiando el papel de α y α_i, obtenemos igualmente que $\alpha_i/\alpha \in \mathcal{O}_K$, así que $\alpha/\alpha_i \in \mathcal{O}_K^*$, y por tanto existe un $\varepsilon \in \mathcal{O}_K^*$ con $\alpha = \varepsilon\alpha_i$. □

Proposición 1.26. $\operatorname{Im} l^{\#}$ *es un retículo de H^r.*

Demostración. Para l la aplicación definida anteriormente, definimos

$$T := l^{-1}(H^r) = \{y \in (E^{r_1,r_2})^* : l(y) \in H^r\}.$$

Para $y = (y_1, \ldots, y_{r_1}, y_{r_1+1}, \ldots, y_{r_1+r_2}) \in E^{r_1,r_2} = \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$, definimos también

$$N_{E^{r_1,r_2}/\mathbb{R}}(y) := y_1 \cdots y_{r_1} |y_{r_1+1}|^2 \cdots |y_{r_1+r_2}|^2.$$

Entonces tenemos

$$\begin{aligned} y \in T &\Leftrightarrow l(y) \in H^r \Leftrightarrow \sum_{i=1}^{r_1+r_2} l_i(y) = 0 \Leftrightarrow \log|N_{E^{r_1,r_2}/\mathbb{R}}(y)| = 0 \\ &\Leftrightarrow N_{E^{r_1,r_2}/\mathbb{R}}(y) = \pm 1. \end{aligned}$$

Sea $\Gamma = X(\mathcal{O}_K)$, que ya sabemos que es un retículo (Corolario 1.15) en $E^{r_1,r_2} \cong \mathbb{R}^n$, con $n = r_1 + 2r_2$, y sea $\operatorname{vol}(\Gamma)$ el volumen del paralelotopo $\Pi(v_1, \ldots, v_n)$ para $\Gamma = [v_1, \ldots, v_n]$. Tomamos un $y \in T$ cualquiera. Para el retículo $y\Gamma = [yv_1, \ldots, yv_n]$, definimos la matriz $M(y) \in \mathrm{GL}_n(\mathbb{R})$ que satisface

$$(yv_1, \ldots, yv_n) = (v_1, \ldots, v_n)M(y),$$

donde, recordamos, la multiplicación de vectores en $\mathbb{R}^n$ yv_i es coordenada a coordenada. Tomando determinantes en ambos lados

$$N_{E^{r_1,r_2}}(y)\det(v_1,\dots,v_n) = \det(v_1,\dots,v_n)\det M(y),$$

y como $\{v_1,\dots,v_n\}$ es una base de $\mathbb{R}^n$, $\det(v_1,\dots,v_n) \neq 0$. Despejando obtenemos

$$\det M(y) = N_{E^{r_1,r_2}/\mathbb{R}}(y),$$

y como $y \in T$, $\det M(y) = \pm 1$ y por tanto $\mathrm{vol}(\Gamma) = \mathrm{vol}(y\Gamma)$. Elegimos ahora números $c_1,\dots,c_{r_1+r_2} \in \mathbb{R}$ que satisfagan

$$\gamma = c_1\cdots c_{r_1+r_2} > \left(\frac{4}{\pi}\right)^{r_2}\mathrm{vol}(\Gamma)$$

y definimos

$$B = \{x \in E^{r_1,r_2} : |x_1| \leq c_1,\dots,|x_{r_1}| \leq c_{r_1},\\ |x_{r_1+1}|^2 \leq c_{r_1+1},\dots |x_{r_1+r_2}|^2 \leq c_{r_1+r_2}\}.$$

Entonces B es un subconjunto de $E^{r_1,r_2} \cong \mathbb{R}^n$ que es simétrico y convexo, y cuyo volumen es

$$\mathrm{vol}(B) = 2^{r_1}\pi^{r_2}\prod_{i=1}^{r_1+r_2} c_i.$$

Por tanto tenemos

$$\mathrm{vol}(B) > 2^{r_1}\pi^{r_2}\left(\frac{4}{\pi}\right)^{r_2}\mathrm{vol}(\Gamma) = 2^n\,\mathrm{vol}(\Gamma) = 2^n\,\mathrm{vol}(y\Gamma),$$

y aplicando el Teorema 1.20, obtenemos que $B \cap (y\Gamma\setminus\{0\}) \neq \varnothing$, o lo que es lo mismo, que existe un $z \neq 0$ tal que $z \in B \cap y\Gamma$. Si escribimos $z = yX(\alpha)$, para $\alpha \in \mathcal{O}_K, \alpha \neq 0$, tenemos que $z \in (E^{r_1,r_2})^*$. Por otro lado, como $z \in B$ se tiene

$$|N_{E^{r_1,r_2}/\mathbb{R}}(z)| = \left|z_1\cdots z_{r_1}|z_{r_1+1}|^2\cdots|z_{r_1+r_2}|^2\right| \leq c_1\cdots c_{r_1+r_2} = \gamma,$$

y también $|N_{K/\mathbb{Q}}\alpha| \leq \gamma$, porque

$$N_{E^{r_1,r_2}/\mathbb{R}}(z) = N_{E^{r_1,r_2}/\mathbb{R}}(y)N_{K/\mathbb{Q}}\alpha = \pm N_{K/\mathbb{Q}}\alpha.$$

Aplicando la Proposición 1.25 a los enteros $m = 1,2,\dots,\lfloor\gamma\rfloor$, encontramos $\alpha_1,\dots,\alpha_N \in \mathcal{O}_K$ con $|N_{K/\mathbb{Q}}\alpha_i| \leq \gamma$ tales que $\varepsilon\alpha = \alpha_i$ para algún i y $\varepsilon \in \mathcal{O}_K^*$. Por tanto z puede ser escrito como $z = yX(\alpha_i)X(\varepsilon)^{-1}$, y cualquier $y \in T$ se puede escribir como $y = zX(\alpha_i^{-1})X(\varepsilon)$. Ahora definimos

$$B_0 := T \cap \left(\bigcup_{i=1}^{N} BX(\alpha_i^{-1})\right) \subset E^{r_1,r_2} \cong \mathbb{R}^n,$$

que es claramente un conjunto acotado por ser B acotado. Probamos ahora la siguiente igualdad:

$$T = B_0 X(\mathcal{O}_K^*).$$

Puesto que la inclusión $T \supset B_0 X(\mathcal{O}_K^*)$ es trivial, basta con demostrar $T \subset B_0 X(\mathcal{O}_K^*)$. Si $y \in T$, lo escribimos como $y = zX(\alpha_i^{-1})X(\varepsilon)$. Por un lado, $zX(\alpha_i^{-1}) \in BX(\alpha_i^{-1})$; por otro, la igualdad $zX(\alpha_i^{-1}) = yX(\varepsilon^{-1}) \in T$ implica que $zX(\alpha_i^{-1}) \in B_0$. Con todo lo anterior finalmente obtenemos $y = zX(\alpha_i^{-1})X(\varepsilon) \in B_0 X(\mathcal{O}_K^*)$, lo que demuestra el contenido. Tomando l en la igualdad

$$l(T) = l(B_0) + l^{\#}(\mathcal{O}_K^*) \subset H^r.$$

Como la aplicación $l : (E^{r_1,r_2})^* \to \mathbb{R}^{r_1+r_2}$ es sobreyectiva (Proposición 1.21), para cualquier $\lambda \in H^r$ existe un $y \in (E^{r_1,r_2})^*$ con $\lambda = l(y)$. Además, por la propia definición de T tenemos que $y \in T$, y $l(T) = H^r$ y por tanto

$$H^r = l(B_0) + l^{\#}(\mathcal{O}_K^*).$$

Como B_0 es acotado, $\overline{B_0}$ es compacto y también lo es la imagen $l(\overline{B_0})$, lo que implica que $l(B_0)$ es acotado también. Como $l^{\#}(\mathcal{O}_K^*)$ es un $\mathbb{Z}-$módulo discreto de H^r y $l(B_0)$ es acotado, por la Proposición 1.19 $l^{\#}(\mathcal{O}_K^*)$ es un retículo en H^r. □

Teorema 1.27 (Unidades de Dirichlet)**.** *Sea K un cuerpo de números y $\mathcal{O}_K^*$ el grupo de unidades de $\mathcal{O}_K$. Entonces se tiene*

$$\mathcal{O}_K^* \cong W_K \times \mathbb{Z}^r, \quad r = r_1 + r_2 - 1,$$

donde W_K es el grupo de las raíces de la unidad contenidas en K, r_1 es el número de $\mathbb{Q}-$conjugados reales de K, y r_2 es el número de pares de $\mathbb{Q}-$conjugados complejos de K.

Demostración. Por la Proposición 1.26, $l^{\#}(\mathcal{O}_K^*)$ es un retículo de H^r y tenemos $l^{\#}(\mathcal{O}_K^*) = \left[l^{\#}(\varepsilon_1), \ldots, l^{\#}(\varepsilon_r)\right]$ para algunos $\varepsilon_i \in \mathcal{O}_K^*$. Por tanto, para cualquier unidad $\varepsilon \in \mathcal{O}_K^*$, existen unos únicos enteros a_i tales que

$$l^{\#}(\varepsilon) = \sum_{i=1}^{r} a_i l^{\#}(\varepsilon_i) = l^{\#}\left(\prod_{i=1}^{r} \varepsilon_i^{a_i}\right).$$

Así, $l^{\#}\left(\varepsilon \prod_{i=1}^{r} \varepsilon_i^{-a_i}\right) = 0$, o equivalentemente, $\varepsilon \prod_{i=1}^{r} \varepsilon_i^{-a_i} \in \ker l^{\#} = W_K$ (Proposición 1.22). Por tanto existe $\zeta \in W_K$ con $\varepsilon = \zeta \prod_{i=1}^{r} \varepsilon_i^{a_i}$. □

1.3. Unidades en cuerpos cuadráticos

En esta sección trataremos los *cuerpos cuadráticos*. Para $d \in \mathbb{Z} \setminus \{0, 1\}$ libre de cuadrados, definimos el cuerpo $K_d = \mathbb{Q}(\sqrt{d})$; pedimos que sea libre de cuadrados porque si $d = p^2 d'$, entonces $K_d = K_{d'}$.

Nos preguntamos en primer lugar por la estructura de su anillo de enteros. Supongamos que $x = a + b\sqrt{d} \in K_d$, $a, b \in \mathbb{Q}$, es un entero algebraico. Entonces su polinomio mínimo ha de tener coeficientes enteros:

$$(x-a)^2 = b^2 d \Rightarrow x^2 - 2ax + a^2 - b^2 d = 0.$$

Y por tanto tenemos que imponer dos condiciones:

$$2a \in \mathbb{Z}, \quad a^2 - b^2 d \in \mathbb{Z}.$$

Una primera observación sencilla es que si a, b son ambos enteros, entonces las dos condiciones se satisfacen. Esto implica que $\mathbb{Z}[\sqrt{d}] \subset \mathcal{O}_{K_d}$. Veamos que esta inclusión es en realidad una igualdad salvo en el caso $d \equiv 1\,(4)$. Para ello observemos primero la siguiente cadena de equivalencias:

$$a \in \mathbb{Z} \Leftrightarrow a^2 \in \mathbb{Z} \Leftrightarrow b^2 d \in \mathbb{Z} \Leftrightarrow b^2 \in \mathbb{Z} \Leftrightarrow b \in \mathbb{Z}.$$

Quizás la única no completamente trivial es $b^2 d \in \mathbb{Z} \Rightarrow b^2 \in \mathbb{Z}$, que es consecuencia de que d sea libre de cuadrados. Por tanto, si existe una solución en la que alguno de los dos, a o b, no sea entero, entonces el otro tampoco lo es.

Vamos a tratar de encontrar dicha solución no entera. Como $2a \in \mathbb{Z}$, supongamos entonces que $a = m/2$, donde $m \in \mathbb{Z}$ es impar, de tal manera que $a \notin \mathbb{Z}$. Entonces,

$$\frac{m^2}{4} - b^2 d \in \mathbb{Z} \Leftrightarrow m^2 - 4b^2 d \in \mathbb{Z} \Leftrightarrow (2b)^2 d \in \mathbb{Z} \Leftrightarrow (2b)^2 \in \mathbb{Z} \Leftrightarrow 2b \in \mathbb{Z}.$$

Así que escribimos $b = n/2$ para $n \in \mathbb{Z}$. La segunda condición nos dice que

$$\frac{m^2}{4} - \frac{n^2}{4} d = \frac{m^2 - n^2 d}{4} \in \mathbb{Z} \Leftrightarrow m^2 - n^2 d \equiv 0\,(4).$$

Como d es libre de cuadrados, $d \not\equiv 0\,(4)$, y $m^2, n^2 \equiv 0, 1\,(4)$ según sean pares o impares. Por tanto, si m y n no tienen la misma paridad, la ecuación $m^2 - n^2 d \equiv 0\,(4)$ no puede satisfacerse y si ambos son pares, se satisface para todo d (recuperando la solución entera). Sin embargo, si m y n son impares, obtenemos una nueva solución siempre y cuando

$$1 - d \equiv 0\,(4) \Leftrightarrow d \equiv 1\,(4).$$

En resumen, un elemento $x \in K_d$ es entero algebraico si existen a, b enteros tales que

$$x = a + b\sqrt{d},$$

o, solo si $d \equiv 1\,(4)$,

$$x = \frac{a}{2} + \frac{b}{2}\sqrt{d},$$

para a, b ambos impares. Esto se traduce en que

$$\mathcal{O}_{K_d} = \begin{cases} \mathbb{Z}[\sqrt{d}], & \text{si } d \equiv 2, 3\,(4), \\ \mathbb{Z}[\frac{1+\sqrt{d}}{2}], & \text{si } d \equiv 1\,(4). \end{cases}$$

(El caso $d \equiv 0\,(4)$ directamente no lo consideramos porque entonces d no sería libre de cuadrados.)

Ahora veamos explícitamente qué es W_{K_d}, que recordamos era el núcleo de la aplicación $l^{\#} : \mathcal{O}_{K_d}^* \to H^r$. Los automorfismos del grupo de Galois de la extensión $K_d/\mathbb{Q}$ son la identidad y la conjugación $\sigma : \sqrt{d} \mapsto -\sqrt{d}$. Distinguimos dos casos:

- Si $d > 0$, $K_d \subset \mathbb{R}$, así que $r_1 = 2, r_2 = 0$ y la aplicación $l^{\#}$ viene dada por
$$l^{\#}(a + b\sqrt{d}) = (\log|a + b\sqrt{d}|, \log|a - b\sqrt{d}|),$$
para $a, b \in \mathbb{Z}$, y si $d \equiv 1\,(4)$,
$$l^{\#}\left(\frac{a}{2} + \frac{b}{2}\sqrt{d}\right) = \left(\log\left|\frac{a}{2} + \frac{b}{2}\sqrt{d}\right|, \log\left|\frac{a}{2} - \frac{b}{2}\sqrt{d}\right|\right),$$
para a, b impares. Si pedimos entonces que $l^{\#}(x) = (0, 0)$, es sencillo ver que los únicos $x \in \mathcal{O}_{K_d}^*$ que lo cumplen son 1 y -1, así que en estos casos $W_{K_d} = \{\pm 1\}$. En particular, el Teorema de las Unidades de Dirichlet nos dice
$$\mathcal{O}_{K_d}^* \cong \{\pm 1\} \times \mathbb{Z}.$$

- Si $d < 0$, la identidad y σ resultan ser conjugadas, así que $r_1 = 0, r_2 = 1$ y la aplicación $l^{\#}$ es simplemente
$$l^{\#}(a + b\sqrt{d}) = \log|a + b\sqrt{d}|,$$
para $a, b \in \mathbb{Z}$, y si $d \equiv 1\,(4)$,
$$l^{\#}\left(\frac{a}{2} + \frac{b}{2}\sqrt{d}\right) = \log\left|\frac{a}{2} + \frac{b}{2}\sqrt{d}\right|,$$
para a, b impares.

  Veamos primero el caso $d \equiv 2, 3\,(4)$. Observamos primero que, como $d < 0$, se tiene $d = -|d|$, así que en particular, $a + b\sqrt{d} = a + bi\sqrt{|d|}$. Por tanto,
$$\log|a + b\sqrt{d}| = 0 \Leftrightarrow |a + b\sqrt{d}| = 1 \Leftrightarrow a^2 + b^2|d| = 1.$$
Si $d < -1$ las únicas soluciones son $a = \pm 1, b = 0$. En el caso $d = -1$ tenemos dos soluciones adicionales, $a = 0, b = \pm 1$.

Para $d \equiv 1\,(4)$ tenemos otra posibilidad. Si $a, b \in \mathbb{Z}$ son ambos impares

$$\log\left|\frac{a}{2}+\frac{b}{2}\sqrt{d}\right| = 0 \Leftrightarrow \frac{a^2}{4}+\frac{b^2}{4}|d| = 1 \Leftrightarrow a^2+b^2|d| = 4.$$

Si $d < -3$ no hay ninguna solución más. Sin embargo, para $d = -3$ tenemos cuatro soluciones adicionales, que corresponden a los valores $a = \pm 1$, $b = \pm 1$.

Como $r = r_1 + r_2 - 1 = 0$, el Teorema de las Unidades de Dirichlet nos da que $\mathcal{O}_{K_d}^* = W_{K_d}$. Finalmente, para $\omega = e^{2\pi i/3} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$, a modo de resumen tenemos

$$\mathcal{O}_{K_d}^* = W_{K_d} = \begin{cases} \{\pm 1, \pm i\}, & \text{si } d = -1, \\ \{\pm 1, \pm\omega, \pm\omega^2\}, & \text{si } d = -3, \\ \{\pm 1\}, & \text{en otro caso.} \end{cases}$$

Hemos visto por tanto que si $d < 0$ el grupo de unidades $\mathcal{O}_{K_d}^*$ es finito y podemos describirlo completamente. Sin embargo, si $d > 0$ el grupo de unidades es infinito; concretamente, $\mathcal{O}_{K_d}^* = \{\pm 1\} \times \mathbb{Z}$. Para acabar esta sección daremos un algoritmo que permita encontrar un generador de la parte *libre* del grupo, es decir, del subgrupo isomorfo a $\mathbb{Z}$.

En esencia estamos buscando un elemento u tal que $\mathcal{O}_{K_d}^* = \langle \pm u \rangle$. A este u lo llamamos *unidad fundamental*.

Comenzamos con el caso $d \equiv 2, 3\,(4)$, para el cual $\mathcal{O}_{K_d} = \mathbb{Z}[\sqrt{d}]$. Por tanto, el grupo de unidades es

$$\mathcal{O}_{K_d}^* = \left\{ a + b\sqrt{d} \in \mathbb{Z}[\sqrt{d}] : a^2 - b^2 d = \pm 1 \right\}.$$

Luego estamos buscando (salvo un signo) soluciones de la ecuación de Pell

$$a^2 - db^2 = 1.$$

Para resolver esta ecuación diofántica se utiliza la teoría de las fracciones continuas. Todo lo que usaremos puede consultarse en [4]. En esencia, todo número real $\alpha \in \mathbb{R}$ puede escribirse como

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \dots}}$$

donde los a_n son todos enteros y están unívocamente determinados. Un número $\alpha \in \mathbb{R}$ es racional si y solo si existe un $n_0 \in \mathbb{N}$ para el cual $a_n = 0$ para todo $n > n_0$ (es decir, el proceso es finito). Cada secuencia de números enteros (todos positivos salvo el primero, que es el único que puede tener signo) determina un número real, y viceversa. Es habitual utilizar la notación

$$\alpha = [a_0, a_1, a_2, \dots].$$

En el caso particular de la raíz cuadrada de un $d \in \mathbb{Z}^+$ libre de cuadrados, se tiene

$$\sqrt{d} = [a_0, \overline{a_1, a_2, \ldots, a_{k-1}, 2a_0}],$$

donde la barra denota que la secuencia $a_1, a_2, \ldots, a_{k-1}, 2a_0$ es periódica. Llamamos entonces a k el *periodo* de la fracción continua.

Podemos recuperar un número racional a partir de su fracción continua gracias al siguiente algoritmo: si $p/q = [a_0, a_1, \ldots, a_N] \in \mathbb{Q}$ escribimos la siguiente tabla

	a_0	a_1	a_2	$\cdots$	a_N	
0	1	p_0	p_1	$\cdots$	p_{N-1}	p_N
1	0	q_0	q_1	$\cdots$	q_{N-1}	q_N

Los números p_k y q_k se obtienen mediante las recurrencias

$$\begin{cases} p_k = a_k p_{k-1} + p_{k-2}, \\ p_0 = a_1, p_{-1} = 1, \end{cases} \qquad \begin{cases} q_k = a_k q_{k-1} + q_{k-2}, \\ q_0 = 1, q_{-1} = 0 \end{cases}$$

En esencia, lo que estamos haciendo es multiplicar a_k por los números que tiene debajo y sumarles los que tienen a su izquierda, obteniendo así los siguientes.

Este proceso se conoce como *algoritmo de la fracción continua* y en el último paso recuperamos la fracción original, esto es, $p_N/q_N = p/q$. Lo interesante de este procedimiento es que podemos aplicarlo también con la fracción continua de un número irracional. En este caso no obtendremos el número original, sino una sucesión de racionales que convergen a él. A las fracciones p_n/q_n que obtenemos usando este algoritmo las llamamos las *convergentes* del número irracional α.

El hecho imprescindible que relaciona la ecuación de Pell con las unidades de $\mathcal{O}_{K_d}$ es (ver [4, Corolario 3.3.7]):

$$(x, y) \textit{ satisface } x^2 - dy^2 = \pm 1 \;\Rightarrow\; x/y \textit{ es una convergente de } \sqrt{d}.$$

De hecho, las convergentes concretas que satisfacen la ecuación son las p_m/q_m tal que $k \mid m+1$, donde k es la longitud del periodo de la fracción continua de $\sqrt{d}$. Concretamente,

$$\textit{Si } \; k \mid m+1 \; \textit{ entonces } \; p_m^2 - dq_m^2 = (-1)^{m+1}.$$

En particular, las soluciones de la ecuación de Pell son las convergentes p_m/q_m con $k \mid m+1$ y m impar.

Volviendo a lo que nos ocupaba, recordemos que buscábamos un generador (salvo signo) de $\mathcal{O}_{K_d}^*$, que se traduce entonces en buscar la solución de

$$x^2 - dy^2 = \pm 1$$

con $x, y \in \mathbb{Z}^+$ mínimos. La razón de buscarlos positivos es porque todas las combinaciones de signos posibles se pueden obtener multiplicando por -1 o calculando su inverso, y pedir que sean lo más pequeños posibles es para que, en efecto, sea un generador. Si $x + y\sqrt{d} \in \mathcal{O}_{K_d}^*$ con $x, y > 0$, entonces

$$\left(x + y\sqrt{d}\right)^2 = (x^2 + dy^2) + 2xy\sqrt{d},$$

y claramente $x^2 + dy^2 > x$, $2xy > y$, luego el generador debe ser la solución mínima. Por lo que hemos visto del algoritmo de la fracción continua, y si $\sqrt{d} = [a_0, \overline{a_1, a_2, \ldots, a_{k-1}, 2a_0}]$ (con periodo k), entonces basta con calcular la convergente p_{k-1}/q_{k-1}:

	a_0	a_1	a_2	$\cdots$	a_{k-1}	$2a_0$	$\cdots$
0	1	p_0	p_1	$\cdots$	p_{k-2}	p_{k-1}	$\cdots$
1	0	q_0	q_1	$\cdots$	q_{k-2}	q_{k-1}	$\cdots$

Esto quiere decir que $\mathcal{O}_{K_d}^* = \left\langle \pm \left(p_{k-1} + q_{k-1}\sqrt{d}\right) \right\rangle$.

Cuando $d \equiv 1\,(4)$ es prácticamente idéntico. En este caso, $\mathcal{O}_{K_d} = \mathbb{Z}[\eta_d]$, con $\eta_d = \frac{1+\sqrt{d}}{2}$. Siguiendo un razonamiento similar a lo anterior (esencialmente, encontrando sus convergentes), y si escribimos $\eta_d^* = \frac{1-\sqrt{d}}{2}$, se puede comprobar que

$$p_{k-1} - q_{k-1}\eta_d^*$$

es unidad fundamental, donde p_n/q_n son las convergentes de η_d. En particular, $\mathcal{O}_{K_d}^* = \left\langle \pm \left(p_{k-1} - q_{k-1}\eta_d^*\right) \right\rangle$

1.4. Un ejemplo cúbico

En este último apartado estudiaremos el cuerpo cúbico

$$K = \mathbb{Q}(r) \quad \text{con} \quad r = \sqrt[3]{6}.$$

Es un hecho (no trivial) que para este cuerpo se tiene $\mathcal{O}_K = \mathbb{Z}[r]$. Si denotamos por $\omega = e^{2\pi i/3} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$, la raíz cúbica de la unidad, las raíces del polinomio mínimo de r, $f(x) = x^3 - 6$, son $r, r\omega, r\omega^2$. Tenemos por tanto tres inmersiones:

$$\mathrm{id} : r \mapsto r, \quad \sigma : r \mapsto r\omega, \quad \sigma^2 : r \mapsto r\omega^2.$$

Pero como $\omega^2 = \overline{\omega}$, las inmersiones σ y σ^2 son conjugadas. Por tanto, si $x + yr + zr^2 \in \mathcal{O}_K$ (esto es, $x, y, z \in \mathbb{Z}$) tenemos, según la notación introducida después de la Definición 1.10,

$$\begin{aligned} X(x + yr + xr^2) &= (\mathrm{id}(x + yr + zr^2), \sigma(x + yr + zr^2)) \\ &= (x + yr + zr^2, x + yr\omega + zr^2\omega^2). \end{aligned}$$

Si un elemento $x + yr + zr^2$ está en W_K entonces

$$|x + yr + zr^2| = 1, \quad |x + yr\omega + zr^2\omega^2| = 1,$$

y solo con la primera condición se tiene que $W_K = \{\pm 1\}$ puesto que $r, r^2 \notin \mathbb{Q}$.

Para calcular la norma de un elemento sustituimos el valor de ω en la segunda coordenada de X:

$$x + yr\omega + zr^2\omega^2 = \left(x - \frac{1}{2}yr - \frac{1}{2}zr^2\right) + i\left(\frac{\sqrt{3}}{2}yr - \frac{\sqrt{3}}{2}zr^2\right).$$

Su módulo al cuadrado entonces es

$$\begin{aligned} \left|x + yr\omega + zr^2\omega^2\right|^2 &= \left|\left(x - \frac{1}{2}yr - \frac{1}{2}zr^2\right) + i\left(\frac{\sqrt{3}}{2}yr - \frac{\sqrt{3}}{2}zr^2\right)\right|^2 \\ &= \left(x - \frac{1}{2}yr - \frac{1}{2}zr^2\right)^2 + \left(\frac{\sqrt{3}}{2}yr - \frac{\sqrt{3}}{2}zr^2\right)^2 \\ &= x^2 + y^2r^2 + 6z^2r - xyr - xzr^2 - 6yz, \end{aligned}$$

y la norma de $x + yr + zr^2 \in \mathcal{O}_K$ viene dada por

$$\begin{aligned} N(x + yr + zr^2) &= (x + yr + zr^2)|x + yr\omega + zr^2\omega^2|^2 \\ &= x^3 + xy^2r^2 + 6xz^2r - x^2yr - x^2zr^2 - 6xyz \\ &\quad + x^2yr + 6y^3 + 6yz^2r^2 - xy^2r^2 - 6xyz - 6y^2zr \\ &\quad + x^2zr^2 + 6y^2zr + 36z^3 - 6xyz - 6xz^2r - 6yz^2r^2 \\ &= x^3 + 6y^3 + 36z^3 - 18xyz. \end{aligned}$$

Por tanto, el polinomio homogéneo cúbico $P(x, y, z) \in \mathbb{Z}[x, y, z]$ que da la función norma en $\mathcal{O}_K$ es

$$P(x, y, z) = N(x + yr + zr^2) = x^3 + 6y^3 + 36z^3 - 18xyz.$$

Consideramos ahora los elementos

$$o_1 = 215 - 42r - 42r^2, \quad o_2 = 971 - 2088r + 855r^2,$$

que cumplen

$$N(o_1) = -1, \quad N(o_2) = -1.$$

En particular, $o_1, o_2 \in \mathcal{O}_K^*$. Estos números han sido encontrados utilizando un programa que, para un rango razonable de enteros x e y, buscaba los z también enteros para los cuales $P(x, y, z) = \pm 1$. En este caso, la aplicación $l^{\#} : \mathcal{O}_K^* \to H^1$ viene dada por

$$l^{\#}(x + yr + zr^2) = \left(\log\left|x + yr + zr^2\right|, \log\left|x + yr\omega + zr^2\omega^2\right|^2\right).$$

Podemos hacer un cálculo aproximado (redondeando a la cuarta cifra decimal) de las imágenes de o_1 y o_2, en concreto:

$$l^{\#}(o_1) = (-11'5799, 11'5799), \quad l^{\#}(o_2) = (-17'3698, 17'3698).$$

A la vista de estos valores inferimos que son linealmente dependientes, en particular

$$3 \cdot l^{\#}(o_1) - 2 \cdot l^{\#}(o_2) = (0, 0).$$

Quizá pudiera parecer que la dependencia lineal es una casualidad debida al error cometido con la aproximación, pero en realidad es una consecuencia inmediata de que ambos vectores, $l^{\#}(o_1)$ y $l^{\#}(o_2)$, pertenezcan al hiperplano H^1. En nuestro caso

$$\dim H^1 = r_1 + r_2 - 1 = 1 + 1 - 1 = 1,$$

de hecho,

$$H^1 = \left\{(x, y) \in \mathbb{R}^2 : x + y = 0\right\} = \left\{(x, -x) \in \mathbb{R}^2 : x \in \mathbb{R}\right\},$$

y los vectores $l^{\#}(o_1)$ y $l^{\#}(o_2)$ son linealmente dependientes.

Ya sabemos que ambos vectores pertenecen a un retículo de H^1. Si quisiéramos hallar el menor retículo que contuviera a ambos, basta con ver (utilizando los cálculos aproximados) que $3l^{\#}(o_1) = 2l^{\#}(o_2)$. De una forma más rigurosa se puede comprobar que, en efecto,

$$(215 - 42r - 42r^2)^3 = -(971 - 2088r + 855r^2)^2,$$

y que por tanto

$$3 \log \left|215 - 42r - 42r^2\right| = 2 \log \left|971 - 2088r + 855r^2\right|,$$

de lo que se deduce $3l^{\#}(o_1) = 2l^{\#}(o_2)$. Sabiendo esto, tomamos $\vec{v} \in \mathbb{R}^2$ tal que $l^{\#}(o_1) = n\vec{v}$ y $l^{\#}(o_2) = m\vec{v}$ con $n, m \in \mathbb{Z}$ lo más pequeños posible. Como $2m = 3n$, tomamos $n = 2$ y $m = 3$ y el vector resulta ser

$$\vec{v} = \frac{l^{\#}(o_1)}{2} = l^{\#}(\pm(1 - 6r + 3r^2)).$$

Si nos creemos que el proceso ha sido exitoso (es decir, que el retículo $l^{\#}(\mathcal{O}_K^*)$ está generado por $\vec{v}$, algo nada trivial), el elemento $u = \pm(1 - 6r + 3r^2)$ es lo que se conoce, recordemos, como *unidad fundamental*, un generador de la parte libre de $\mathcal{O}_K^*$. Es decir, todos los elementos $\alpha \in \mathcal{O}_K^*$ satisfacen:

$$\alpha = \pm u^k, \quad \text{para algún } k \in \mathbb{Z}.$$

De esta forma, también es unidad fundamental el elemento

$$\pm u^{-1} = \pm \frac{1}{1 - 6r + 3r^2} = \pm(109 + 60r + 33r^2).$$

Por último, veamos una aplicación nada inmediata de nuestro estudio del cuerpo K:

No existe ningún cubo perfecto de la forma $6n^3+1$ *con* $n \in \mathbb{Z}^+$.

De hecho, probaremos algo ligeramente más fuerte:

$$\left\{(x,y) \in \mathbb{Z}^2 : x^3 + 6y^3 = 1\right\} = \{(1,0)\}.$$

Fijamos para ello como unidad fundamental $u_0 = 1 - 6r + 3r^2 \geq 0$. Para cada $(x,y) \in \mathbb{Z}^2$ con $x^3 + 6y^3 = 1$ consideramos el número $x + yr \in \mathcal{O}_K$. Si calculamos su norma,

$$N(x+yr) = P(x,y,0) = x^3 + 6y^3 = 1,$$

así que $x + yr \in \mathcal{O}_K^*$. Por el Teorema de las Unidades de Dirichlet, existe un $k \in \mathbb{Z}$ tal que $x + yr = \pm u_0^k$. Además, como $N(x+yr) = N(u_0) = 1$, podemos descartar el caso negativo, de tal manera que

$$x + yr = u_0^k, \qquad \text{para algún } k \in \mathbb{Z}.$$

Queremos concluir que el único caso posible es cuando $k = 0$, que nos da la solución $(1,0)$ que ya teníamos. Distinguimos primero entre $k \geq 0$ y $k < 0$. El caso negativo es análogo al positivo utilizando la unidad fundamental $u_0^{-1} = 109 + 60r + 33r^2$, así que solo consideraremos $k \geq 0$.

Como decíamos, si $k = 0$ recuperamos la solución trivial, y $k = 1$ es imposible. Supongamos por tanto que $k \geq 2$. El binomio de Newton nos da

$$\begin{aligned} x + yr &= (1 - 6r + 3r^2)^k \\ &= \sum_{j=0}^{k} \binom{k}{j} 1^{k-j} (3r^2 - 6r)^j \\ &= 1 + 3k(r^2 - 2r) + \sum_{j=2}^{k} 3^j (r^2 - 2r)^j \binom{k}{j}. \end{aligned}$$

Si expresamos $(r^2 - 2r)^j$ como una combinación lineal entera $a + br + cr^2$ y nos fijamos ahora en el coeficiente de r^2 a ambos lados de la igualdad:

$$0 = 3k + \sum_{j=2}^{k} 3^j b_j \binom{k}{j},$$

para algunos enteros b_j, o, equivalentemente,

$$k = \sum_{j=2}^{k} 3^{j-1} c_j \binom{k}{j},$$

para algunos enteros c_j.

Para concluir que esto no se puede dar, consideremos primero la aplicación $v_3 : \mathbb{N} \to \mathbb{N}$, que asocia a cada natural n el exponente del primo 3 en la factorización de n (siendo $v_3(n) = 0$ cuando $3 \nmid n$). De la igualdad

$$\binom{k}{j} = \frac{k!}{j!(k-j!)} = \frac{k}{j}\frac{(k-1)!}{(j-1)!(k-1-(j-1))!} = \frac{k}{j}\binom{k-1}{j-1}$$

deducimos que

$$v_3\left(\binom{k}{j}\right) = v_3(k) - v_3(j) + v_3\left(\binom{k-1}{j-1}\right) \geq v_3(k) - v_3(j).$$

Por otro lado,

$$j - v_3(j) \geq 2 \quad \text{cuando } j \geq 2.$$

Si $v_3(j) = 0$ no hay nada que probar. Si $v_3(j) \geq 1$ (es decir, $3 \mid j$), se tiene

$$j - v_3(j) \geq 3^{v_3(j)} - v_3(j) \geq 2.$$

Esto es consecuencia inmediata de que la función $f(x) = 3^x - x$ sea creciente en $[1, \infty)$ y que por tanto alcance su mínimo en $x = 1$, cuyo valor es $f(1) = 2$.

Con todos estos ingredientes es sencillo encontrar una contradicción. Recordamos que

$$k = \sum_{j=2}^{k} 3^{j-1} c_j \binom{k}{j},$$

es evidente entonces que $3^{v_3(k)+1}$ no divide el lado izquierdo. Sin embargo, para $2 \leq j \leq k$,

$$3^{v_3(k)+1} \leq 3^{v_3(k)+1+(j-v_3(j)-2)} = 3^{(j-1)+(v_3(k)-v_3(j))} \leq 3^{(j-1)+v_3\left(\binom{k}{j}\right)},$$

y como tenemos

$$3^{j-1}3^{v_3\left(\binom{k}{j}\right)} \mid 3^{j-1} c_j \binom{k}{j}$$

para todo $j \geq 2$, concluimos que $3^{v_3(k)+1}$ divide al lado derecho, y llegamos a la contradicción deseada.

Capítulo 2

Ideales primos y el grupo de S-unidades

En este capítulo hablaremos de la estructura de los ideales de $\mathcal{O}_K$. Habitualmente utilizaremos las letras I, J para denotar un ideal genérico, y reservaremos $\mathfrak{p}, \mathfrak{q}$ para ideales primos. El conjunto de todos los ideales primos lo denotaremos por

$$\mathrm{Spec}(\mathcal{O}_K) = \{\mathfrak{p} \leq \mathcal{O}_K : \mathfrak{p} \text{ es primo}\},$$

el *espectro* de $\mathcal{O}_K$. Además, en $\mathcal{O}_K$ todo ideal primo es maximal.

Notación. Dados $I, J \leq \mathcal{O}_K$, escribiremos $I \mid J$ si $J \subset I$.

2.1. Factorización única en los ideales de $\mathcal{O}_K$

En general, el anillo $\mathcal{O}_K$ no es un dominio de factorización única. Esto nos fuerza a considerar, no las factorizaciones de los elementos en sí, sino de los ideales principales que generan. Afortunadamente, sí tenemos una factorización única en ideales primos para todo ideal en $\mathcal{O}_K$, cosa que nos ocupará esta primera sección y que aprovecharemos durante el resto del capítulo. Comenzamos dando un resultado general en anillos conmutativos.

Teorema 2.1. *En un anillo conmutativo A, $\mathfrak{p} \leq A$ es primo si y solo si para todos $I, J \leq A$ se tiene*

$$\mathfrak{p} \mid IJ \Rightarrow \mathfrak{p} \mid I \text{ ó } \mathfrak{p} \mid J.$$

Demostración. ($\Rightarrow$) Sea $\mathfrak{p} \leq A$ un ideal primo. Si $IJ \subset \mathfrak{p}$ pero $I \not\subset \mathfrak{p}$, sea $x \in I$ tal que $x \notin \mathfrak{p}$. Para todo $y \in J$, $xy \in IJ \subset \mathfrak{p}$, y como $\mathfrak{p}$ es primo y $x \notin \mathfrak{p}$, $y \in \mathfrak{p}$. Como y era genérico, $J \subset \mathfrak{p}$.

($\Leftarrow$) Sean $x, y \in A$ tales que $xy \in \mathfrak{p}$, entonces $(x)(y) = (xy) \subset \mathfrak{p}$, y por tanto $(x) \subset \mathfrak{p}$ ó $(y) \subset \mathfrak{p}$. Por tanto $x \in \mathfrak{p}$ ó $y \in \mathfrak{p}$, por lo que $\mathfrak{p}$ es primo. $\square$

Corolario 2.2. *Sea K un cuerpo de números. En $\mathcal{O}_K$, si $\mathfrak{p} \mid \mathfrak{p}_1 \cdots \mathfrak{p}_r$, donde todos los ideales son primos no nulos, entonces $\mathfrak{p} = \mathfrak{p}_i$ para algún i.*

Demostración. Por el Teorema 2.1, $\mathfrak{p} \mid \mathfrak{p}_i$ para algún i. Como todo ideal primo en $\mathcal{O}_K$ es maximal, entonces $\mathfrak{p} = \mathfrak{p}_i$. □

Necesitamos otros dos resultados previos referentes a los ideales del espectro de $\mathcal{O}_K$. El primero de ellos es el siguiente:

Lema 2.3. *Todo ideal no nulo de $\mathcal{O}_K$ contiene un producto de ideales primos no nulos.*

Demostración. Supongamos que el enunciado es falso y tomemos I un ideal no nulo que tenga índice minimal[1] y que no contenga un producto de ideales primos no nulos. Evidentemente, $I \neq \mathcal{O}_K$, puesto que $\mathcal{O}_K$ contiene ideales primos no nulos, así que $[\mathcal{O}_K : I] \geq 2$. Como I no puede ser primo, deben existir $x, y \in \mathcal{O}_K \backslash I$ tales que $xy \in I$. Por tanto los ideales $(x) + I, (y) + I$ contienen propiamente a I y tienen índice más pequeño, por lo que existen $\mathfrak{p}_1, \ldots, \mathfrak{p}_r, \mathfrak{q}_1, \ldots, \mathfrak{q}_s \in \operatorname{Spec}(\mathcal{O}_K)$ no nulos tales que

$$\mathfrak{p}_1 \cdots \mathfrak{p}_r \subset (x) + I, \qquad \mathfrak{q}_1 \cdots \mathfrak{q}_s \subset (y) + I.$$

Por tanto,

$$\mathfrak{p}_1 \cdots \mathfrak{p}_r \, \mathfrak{q}_1 \cdots \mathfrak{q}_s \subset ((x) + I)\,((y) + I) = (xy) + xI + yI + I^2 \subset I,$$

donde usamos que $xy \in I$. Por consiguiente, I contiene un producto de ideales primos no nulos, lo que resulta ser una contradicción. □

Para el segundo necesitamos introducir, para $\mathfrak{p} \in \operatorname{Spec}(\mathcal{O}_K)$ no nulo, el $\mathcal{O}_K$-módulo

$$\tilde{\mathfrak{p}} = \{x \in K : x\,\mathfrak{p} \subset \mathcal{O}_K\}\,.$$

Este módulo hace las veces de *inverso* del ideal, en el sentido de que se tiene

$$\mathfrak{p}\,\tilde{\mathfrak{p}} = \mathcal{O}_K$$

como $\mathcal{O}_K$-módulos (ver [8, Theorem 3.2]).

Lema 2.4. *Para $I, J \leq \mathcal{O}_K, \mathfrak{p} \in \operatorname{Spec}(\mathcal{O}_K)$ no nulos, $\mathfrak{p}\,I = \mathfrak{p}\,J$ si y solo si $I = J$.*

Demostración. La implicación de derecha a izquierda es trivial. Para la otra, supongamos que $\mathfrak{p}\,I = \mathfrak{p}\,J$. Multiplicamos por $\tilde{\mathfrak{p}}$ a ambos lados, y como la multiplicación de $\mathcal{O}_K$-módulos es asociativa, se tiene $I = J$. □

[1] Este índice es siempre finito. Lo veremos como consecuencia de la Proposición 2.12 más adelante.

Con esto ya podemos probar la factorización única en ideales primos.

Teorema 2.5. *Todo ideal propio no nulo de $\mathcal{O}_K$ es un producto único de ideales primos no nulos.*

Demostración. Para la existencia, probaremos por inducción en $r \geq 1$ que si un ideal propio no nulo $I \leq \mathcal{O}_K$ contiene un producto de r ideales primos no nulos (Lema 2.3), entonces factoriza como un producto de ideales primos no nulos. Cuando $r = 1$, $\mathfrak{p} \subset I$, y como $\mathfrak{p}$ es maximal, $I = \mathfrak{p}$. Asumiendo el resultado para r, supongamos $\mathfrak{p}_1 \cdots \mathfrak{p}_{r+1} \subset I$. Como I es propio, existe un ideal maximal $\mathfrak{p}$ tal que $I \subset \mathfrak{p}$. Entonces $\mathfrak{p}_1 \cdots \mathfrak{p}_{r+1} \subset \mathfrak{p}$, y por el Corolario 2.2, $\mathfrak{p} = \mathfrak{p}_i$ para algún i. De la cadena $\mathfrak{p}_1 \cdots \mathfrak{p}_{r+1} \subset I \subset \mathfrak{p}_i$ multiplicando por $\tilde{\mathfrak{p}}_i$ obtenemos

$$\mathfrak{p}_1 \cdots \mathfrak{p}_{i-1}\, \mathfrak{p}_{i+1} \cdots \mathfrak{p}_{r+1} \subset \tilde{\mathfrak{p}}_i I \subset \mathcal{O}_K.$$

Entonces por hipótesis inductiva $\tilde{\mathfrak{p}}_i I$ es un producto de ideales primos no nulos, y por tanto I también multiplicando por $\mathfrak{p}_i$ de vuelta.

Para la unicidad, supongamos $\mathfrak{p}_1 \cdots \mathfrak{p}_r = \mathfrak{q}_1 \cdots \mathfrak{q}_s$. Podemos cancelar los ideales primos comunes en ambos lados (Lema 2.4), de tal forma que podemos suponer que $\mathfrak{p}_i \neq \mathfrak{q}_j$ para todo i, j. Como $\mathfrak{p}_1 \mid \mathfrak{p}_1 \cdots \mathfrak{p}_r = \mathfrak{q}_1 \cdots \mathfrak{q}_s$, por el Corolario 2.2, $\mathfrak{p}_1 = \mathfrak{q}_j$ para algún j, lo cual resulta ser una contradicción. □

2.2. Valores absolutos y el teorema de Ostrowski

Pasamos ahora a hablar de valoraciones.

Definición 2.6. Para $\mathfrak{p} \in \operatorname{Spec}(\mathcal{O}_K)$ no nulo, definimos la aplicación

$$v_{\mathfrak{p}} : \mathcal{O}_K \setminus \{0\} \to \mathbb{Z}_{\geq 0},$$

dada por $v_{\mathfrak{p}}(x) = m$ si $x\mathcal{O}_K = \mathfrak{p}^m I$ con $\mathfrak{p} \nmid I$.

Como consecuencia de la factorización única en ideales primos de $\mathcal{O}_K$, para cualesquiera $x, y \in \mathcal{O}_K$ no nulos se tiene

$$v_{\mathfrak{p}}(xy) = v_{\mathfrak{p}}(x) + v_{\mathfrak{p}}(y).$$

Esto nos permite extender $v_{\mathfrak{p}}$ a todo $K^\times$: dado $\alpha \in K^\times$ escribimos $\alpha = x/y$ para $x, y \in \mathcal{O}_K$ no nulos (Lema 1.9) y definimos

$$v_{\mathfrak{p}}(\alpha) := v_{\mathfrak{p}}(x) - v_{\mathfrak{p}}(y) \in \mathbb{Z}.$$

Para comprobar que está bien definido, si $x/y = x'/y'$ entonces $xy' = x'y$. Por tanto, $v_{\mathfrak{p}}(x) + v_{\mathfrak{p}}(y') = v_{\mathfrak{p}}(x') + v_{\mathfrak{p}}(y)$, y $v_{\mathfrak{p}}(x) - v_{\mathfrak{p}}(y') = v_{\mathfrak{p}}(x') - v_{\mathfrak{p}}(y')$. En $K^\times$, $v_{\mathfrak{p}}(\alpha\beta) = v_{\mathfrak{p}}(\alpha) + v_{\mathfrak{p}}(\beta)$ para todo $\alpha, \beta \in K^\times$, así que la aplicación

$$v_{\mathfrak{p}} : K^\times \to \mathbb{Z}$$

es un homomorfismo de grupos sobreyectivo (para $x \in \mathfrak{p} \setminus \mathfrak{p}^2$, $v_\mathfrak{p}(x) = 1$). Definimos $v_\mathfrak{p}(0) = \infty$, donde $\infty > n$ para todo entero n.

En $\mathcal{O}_K$, $v_\mathfrak{p}(x+y) \geq \min(v_\mathfrak{p}(x), v_\mathfrak{p}(y))$. Si x, y ó $x+y$ son alguno 0, entonces la desigualdad es trivial. Si no es el caso y $m := \min(v_\mathfrak{p}(x), v_\mathfrak{p}(y))$, entonces $\mathfrak{p}^m \mid x\mathcal{O}_K$, $\mathfrak{p}^m \mid y\mathcal{O}_K$, y por tanto $x, y \in \mathfrak{p}^m$. Como $\mathfrak{p}$ es ideal, $x+y \in \mathfrak{p}^m$, así que $v_\mathfrak{p}(x+y) \geq m$. Esta desigualdad se extiende a todo K de manera natural: $v_\mathfrak{p}(\alpha+\beta) \geq \min(v_\mathfrak{p}(\alpha), v_\mathfrak{p}(\beta))$, para todo $\alpha, \beta \in K$. Por consiguiente, $v_\mathfrak{p}$ es una *valoración* en K.

Definición 2.7. Fijada una constante $c > 1$, definimos la aplicación

$$\begin{array}{rccc} |\cdot|_\mathfrak{p}: & K & \to & \mathbb{R} \\ & \alpha & \mapsto & |\alpha|_\mathfrak{p} \end{array}$$

dada por $|\alpha|_\mathfrak{p} = c^{-v_\mathfrak{p}(\alpha)}$ para $\alpha \in K^\times$ y $|0|_\mathfrak{p} = 0$. Esta aplicación es un *valor absoluto* $\mathfrak{p}$*-ádico* en K.

Cumple las propiedades de cualquier valor absoluto sobre un dominio de integridad:

- No negativa: $|\alpha|_\mathfrak{p} \geq 0$.
- Definida positiva: $|\alpha|_\mathfrak{p} = 0$ si y solo si $\alpha = 0$.
- Multiplicativa: $|\alpha\beta|_\mathfrak{p} = |\alpha|_\mathfrak{p}|\beta|_\mathfrak{p}$.
- Desigualdad triangular: $|\alpha+\beta|_\mathfrak{p} \leq |\alpha|_\mathfrak{p} + |\beta|_\mathfrak{p}$.

Además, también satisface la propiedad más fuerte

$$|\alpha+\beta|_\mathfrak{p} \leq \max(|\alpha|_\mathfrak{p}, |\beta|_\mathfrak{p}),$$

que se conoce como *desigualdad ultramétrica*. Los valores absolutos que satisfacen esta condición adicional se llaman *no arquimedianos* o *ultramétricos*. En caso contrario, se llaman *arquimedianos*.

Cambiar la constante c produce una norma equivalente (en el sentido de que la topología inducida es la misma), así que tenemos una topología $\mathfrak{p}$-ádica bien definida en K independiente del valor de c. Para $\mathfrak{p}, \mathfrak{q} \in \operatorname{Spec}(\mathcal{O}_K)$ no nulos con $\mathfrak{p} \neq \mathfrak{q}$ los valores absolutos $\mathfrak{p}$-ádico y $\mathfrak{q}$-ádico no son equivalentes: el teorema chino del resto nos permite encontrar un $x \in \mathcal{O}_K$ tal que

$$x \equiv 0 \mod \mathfrak{p}, \quad x \equiv 1 \mod \mathfrak{q},$$

así que el valor absoluto $\mathfrak{p}$-ádico de x es menor que 1 mientras que el $\mathfrak{q}$-ádico es igual a 1; por tanto, no son equivalentes.

Los valores absolutos arquimedianos en K se definen en términos de las inmersiones $\sigma : K \to \mathbb{R}$ y $\sigma : K \to \mathbb{C}$:

$$|\alpha|_\sigma := |\sigma(\alpha)|_\infty^{1/[K:\mathbb{Q}]},$$

donde $|\cdot|_\infty$ es el valor absoluto estándar en $\mathbb{R}$ y $\mathbb{C}$. Con la notación previa, tenemos r_1+2r_2 inmersiones, pero solamente r_1+r_2 valores absolutos arquimedianos en K (bajo equivalencia) puesto que las inmersiones conjugadas complejas producen el mismo valor absoluto ($|a+bi|_\infty = |a-bi|_\infty$ en $\mathbb{C}$).

Nuestro primer objetivo es demostrar el Teorema de Ostrowski para valores absolutos no arquimedianos, que afirma que cualquiera de estos valores absolutos es equivalente a uno $\mathfrak{p}$-ádico para un único $\mathfrak{p} \in \mathrm{Spec}(\mathcal{O}_K)$ no nulo.

Lema 2.8. *Sea $\mathfrak{p} \in \mathrm{Spec}(\mathcal{O}_K)$ no nulo. Si $\alpha \in K^\times$ y $v_\mathfrak{p}(\alpha) \geq 0$ entonces $\alpha = x/y$ con $x, y \in \mathcal{O}_K$ no nulos y $v_\mathfrak{p}(y) = 0$.*

Demostración. Escribimos $\alpha\mathcal{O}_K = IJ^{-1}$ para I, J ideales coprimos en $\mathcal{O}_K$. Como $v_\mathfrak{p}(\alpha) \geq 0$, $\mathfrak{p} \nmid J$. Tomamos un $x \in I \setminus \mathfrak{p}\, I$ y definimos un nuevo ideal $L = xI^{-1}$. Entonces tenemos

$$(\alpha) = IJ^{-1} = IL(JL)^{-1}.$$

Claramente $IL = (x)$ es principal. Por otro lado

$$JL = IL(\alpha)^{-1} = (x)(\alpha)^{-1} = (x\alpha^{-1}),$$

así que JL también es principal, digamos $JL = (y)$. Los ideales L y $\mathfrak{p}$ son coprimos, ya que si no lo fueran $\mathfrak{p} \mid L$ y por tanto $L \subset \mathfrak{p}$, pero que $xI^{-1} \subset \mathfrak{p}$ implicaría que $x \in \mathfrak{p}\, I$, lo que contradice la elección de x. Como J y L no son divisibles por $\mathfrak{p}$, $v_\mathfrak{p}(y) = 0$. Por tanto

$$(\alpha) = (x)(y)^{-1} = (x/y),$$

y, reescalando por una unidad, $\alpha = x/y$ con $v_\mathfrak{p}(y) = 0$. □

Teorema 2.9. *Sea $v : K^\times \to \mathbb{R}$ un homomorfismo no nulo con*

$$v(\alpha+\beta) \geq \min(v(\alpha), v(\beta)). \quad (*)$$

Entonces $v = tv_\mathfrak{p}$ para un único $\mathfrak{p} \in \mathrm{Spec}(\mathcal{O}_K)$ no nulo y $t > 0$.

Demostración. La unicidad es sencilla. Si $v = tv_\mathfrak{p}$ para $\mathfrak{p}$ no nulo,

$$\{\alpha \in \mathcal{O}_K : tv_\mathfrak{p} > 0\} = \{\alpha \in \mathcal{O}_K : v_\mathfrak{p} > 0\} = \{\alpha \in \mathcal{O}_K : \alpha \in \mathfrak{p}\} = \mathfrak{p}\,.$$

Como $tv_\mathfrak{p}$ toma valores positivos exclusivamente en $\mathfrak{p}$, podemos recuperar $\mathfrak{p}$ de las propiedades de $tv_\mathfrak{p}$. Como t es el valor positivo más pequeño que $tv_\mathfrak{p}$ toma en $K^\times$, el valor de t está determinado también.

Para la existencia de un $\mathfrak{p}$ y un t tal que $v = tv_\mathfrak{p}$, probaremos que el conjunto

$$\mathfrak{p} := \{\alpha \in \mathcal{O}_K \setminus \{0\} : v(\alpha) > 0\} \cup \{0\}$$

satisface $\mathfrak{p} \in \mathrm{Spec}(\mathcal{O}_K)$ y que $v = tv_\mathfrak{p}$ para algún $t > 0$.

Primero veamos que $v(\alpha) \geq 0$ para todo $\alpha \in \mathcal{O}_K \setminus \{0\}$. Como $v(1 \cdot 1) = v(1) + v(1)$, $v(1) = 0$. Por tanto $0 = v((-1)^2) = 2v(-1)$, así que también $v(-1) = 0$. Por inducción, y usando $(*)$, se tiene que $v(n) \geq 0$ para todo $n \in \mathbb{Z}$ no nulo. Para $\alpha \in \mathcal{O}_K$ no nulo, podemos escribir una ecuación de dependencia sobre $\mathbb{Z}$

$$\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1\alpha + a_0 = 0,$$

con $a_j \in \mathbb{Z}$. Elegimos n lo más pequeño posible, de tal forma que $a_0 \neq 0$. Reescribimos la ecuación de arriba como

$$\alpha^n = -a_{n-1}\alpha^{n-1} - \cdots - a_1\alpha + a_0.$$

Si $a_j \neq 0$, $v(-a_j\alpha^j) = v(a_j) + jv(\alpha) \geq jv(\alpha)$. Cuando $a_j = 0$, el término $a_j\alpha^j$ es 0 y podemos ignorarlo.

Ahora si $v(\alpha) < 0$, entonces $v(-a_j\alpha^j) \geq (n-1)v(\alpha)$ ya que $j \leq n-1$. Por tanto, extendiendo $(*)$ a una suma de varios términos,

$$v(-a_{n-1}\alpha^{n-1} - \cdots - a_1\alpha + a_0) \geq (n-1)v(\alpha)$$

Como $v(\alpha^n) = nv(\alpha)$, tenemos $nv(\alpha) \geq (n-1)v(\alpha)$, así que $v(\alpha) \geq 0$. Esto contradice nuestra asunción inicial $v(\alpha) < 0$. Por tanto $v(\alpha) \geq 0$.

Como v no es idénticamente 0 en $K^\times$, tampoco lo es en $\mathcal{O}_K \setminus \{0\}$, lo que significa que debe tomar algún valor positivo en $\mathcal{O}_K \setminus \{0\}$. Por tanto, el conjunto $\mathfrak{p}$ no es $\{0\}$. Como v es un homomorfismo y cumple la condición $(*)$, el conjunto $\mathfrak{p}$ es un subgrupo de $\mathcal{O}_K$, y también es un $\mathcal{O}_K$-módulo debido a la no negatividad de v en $\mathcal{O}_K \setminus \{0\}$. Por tanto $\mathfrak{p}$ es un ideal de $\mathcal{O}_K$. Como $v(1) = 0$, $1 \neq \mathfrak{p}$, así que es un ideal propio de $\mathcal{O}_K$. Veamos que es un ideal primo. Supongamos que $\alpha, \beta \notin \mathfrak{p}$ y $\alpha\beta \in \mathfrak{p}$. Por tanto, $v(\alpha) = v(\beta) = 0$, y $v(\alpha\beta) = v(\alpha) + v(\beta) = 0$, lo que contradice que $\alpha\beta \in \mathfrak{p}$. Concluimos finalmente que $\mathfrak{p} \in \operatorname{Spec}(\mathcal{O}_K)$.

Ahora probemos que $v = tv_\mathfrak{p}$ para algún $t > 0$.

Primero veamos que si $v_\mathfrak{p}(\alpha) = 0$, entonces $v(\alpha) = 0$. Por el Lema 2.8 podemos escribir $\alpha = x/y$ con $x, y \in \mathcal{O}_K$ y $v_\mathfrak{p}(y) = 0$. Por tanto $v_\mathfrak{p}(x) = v_\mathfrak{p}(\alpha y) = 0 + 0 = 0$. Como $x, y \in \mathcal{O}_K$ y no están en $\mathfrak{p}$, la definición de $\mathfrak{p}$ nos dice que $v(x) = v(y) = 0$, de lo que concluimos que $v(\alpha) = 0$.

Finalmente veamos que $v = tv_\mathfrak{p}$ para algún $t > 0$. Para $\alpha \in K^\times$, sea $n = v_\mathfrak{p}(\alpha) \in \mathbb{Z}$. Elegimos $\gamma \in \mathfrak{p} \setminus \mathfrak{p}^2$, de tal manera que $v_\mathfrak{p}(\gamma) = 1$ y $v(\gamma) > 0$. Por tanto $v_\mathfrak{p}(\alpha/\gamma^n) = 0$, y $v(\alpha/\gamma^n) = 0$, y entonces

$$v(\alpha) = nv(\gamma) = v_\mathfrak{p}(\alpha)v(\gamma).$$

La elección de γ no tiene nada que ver con α. Esta ecuación se cumple para todo $\alpha \in K^\times$, así que $v = tv_\mathfrak{p}$ con $t = v(\gamma) > 0$. □

Corolario 2.10 (Teorema de Ostrowski)**.** *Todo valor absoluto no arquimediano en K es equivalente a un valor absoluto $\mathfrak{p}$-ádico para un único $\mathfrak{p} \in \mathrm{Spec}(\mathcal{O}_K)$ no nulo.*

Demostración. Sea $|\cdot|$ un valor absoluto no arquimediano en K. Consideramos la aplicación

$$\begin{array}{rccc} v: & K^\times & \to & \mathbb{R} \\ & \alpha & \mapsto & -\log|\alpha| \end{array}$$

Es sencillo ver que v es un homomorfismo de grupos, y, por ser $|\cdot|$ no arquimediano, también satisface la condición $(*)$, así que por el Teorema 2.9 existen un único $\mathfrak{p} \in \mathcal{O}_K$ no nulo y una constante $t > 0$ tal que

$$-\log|\alpha| = t v_{\mathfrak{p}}(\alpha).$$

Reescribiendo la ecuación como $|\alpha| = (e^t)^{-v_{\mathfrak{p}}}$, vemos que $|\cdot|$ es un valor absoluto $\mathfrak{p}$-ádico con constante $c = e^t > 1$. □

2.3. Normas de ideales y fórmula producto

Ahora ya tenemos una clasificación de todos los valores absolutos en K en función de la topología inducida. A cada clase de equivalencia de valores absolutos la llamamos *lugar*.

Denotamos por M_K, o simplemente M, al conjunto de lugares de K salvo una excepción: para los valores absolutos arquimedianos complejos (si los hay), elegimos los *dos* representantes conjugados, a pesar de que producen la misma topología. El motivo de tomar ambos valores absolutos aparecerá de forma natural más adelante. Escribiremos M^∞ para denotar los valores absolutos arquimedianos, que por tanto tendrá $r_1 + 2r_2 = n$ elementos, y M^0 para $M \backslash M^\infty$.

Recordamos que, para los valores absolutos no arquimedianos, teníamos la elección de la constante $c > 1$ como grado de libertad. Queremos elegir unas constantes c adecuadas de tal forma que podamos obtener una expresión cerrada en función de los valores absolutos. Para ello, introducimos la siguiente noción de norma.

Definición 2.11. Para $I \leq \mathcal{O}_K$ un ideal no nulo, definimos la *norma absoluta* de I como el cardinal del anillo cociente $\mathcal{O}_K/I$, esto es,

$$N_{K/\mathbb{Q}}(I) = |\mathcal{O}_K/I|.$$

Por convención, definimos $N((0)) = 0$.

En primer lugar veamos por qué esta definición extiende nuestro concepto inicial de norma.

Proposición 2.12. *Sea $x \in \mathcal{O}_K$. Entonces*

$$N_{K/\mathbb{Q}}((x)) = |N_{K/\mathbb{Q}}(x)|.$$

Observación. En el lado izquierdo tenemos la norma absoluta del ideal generado por x mientras que en el derecho tenemos el valor absoluto de la norma del elemento x que vimos en la Definición 1.11.

Demostración. Para $x = 0$ el resultado es trivial, así que asumimos $x \neq 0$. Por el Lema 1.14, sabemos que el anillo de enteros $\mathcal{O}_K$ es un $\mathbb{Z}$-módulo de dimensión n. Por otro lado, el ideal $x\mathcal{O}_K$ también es un $\mathbb{Z}$-módulo de dimensión n porque la aplicación $m_x : \mathcal{O}_K \to x\mathcal{O}_K$ es un isomorfismo de $\mathbb{Z}$-módulos (es inyectiva por ser $\mathcal{O}_K$ un dominio de integridad y es claramente sobreyectiva).

Sabemos por tanto que existe una $\mathbb{Z}$-base $\{e_1, \dots, e_n\}$ de $\mathcal{O}_K$ y enteros positivos $c_i \in \mathbb{Z}^+$ tales que $\{c_1e_1, \dots, c_ne_n\}$ es una $\mathbb{Z}$-base de $x\mathcal{O}_K$. Por consiguiente, el cociente $\mathcal{O}_K/x\mathcal{O}_K$ es isomorfo a $\prod_{i=1}^n \mathbb{Z}/c_i\mathbb{Z}$; en particular, $N_{K/\mathbb{Q}}((x)) = \prod_{i=1}^n c_i$.

Consideramos la función $\mathbb{Z}$-lineal $u : \mathcal{O}_K \to x\mathcal{O}_K$ definida en la base anterior como $u(e_i) = c_ie_i$, con determinante $\det(u) = \prod_{i=1}^n c_i$. También sabemos que $\{xe_1, \dots, xe_n\}$ es una $\mathbb{Z}$-base de $x\mathcal{O}_K$, así que podemos considerar el automorfismo de cambio de base $v : x\mathcal{O}_K \to x\mathcal{O}_K$ dado por $v(c_ie_i) = xe_i$. Por tener v inversa, $\det(v)$ es invertible en $\mathbb{Z}$, así que $\det(v) = \pm 1$. Por otro lado, la composición $v \circ u$ es claramente la multiplicación por x, m_x, cuyo determinante es $N_{K/\mathbb{Q}}(x)$ por definición.

Juntando todo esto obtenemos

$$N_{K/\mathbb{Q}}(x) = \det(m_x) = \det(v)\det(u) = \pm\prod_{i=1}^n c_i = \pm N_{K/\mathbb{Q}}((x)),$$

lo que concluye la prueba. □

Como corolario inmediato, $N_{K/\mathbb{Q}}(I) < +\infty$ para cualquier ideal $I \leq \mathcal{O}_K$ no nulo. Veamos ahora que la aplicación $N_{K/\mathbb{Q}}$ es completamente multiplicativa.

Proposición 2.13. *Sean $I, J \leq \mathcal{O}_K$. Entonces*

$$N_{K/\mathbb{Q}}(IJ) = N_{K/\mathbb{Q}}(I)N_{K/\mathbb{Q}}(J).$$

Demostración. Veamos primero que la norma es multiplicativa en ideales primos. Sean $\mathfrak{p}, \mathfrak{q} \in \mathrm{Spec}(\mathcal{O}_K)$ distintos y no nulos (si alguno es cero, el resultado es trivial). Entonces, por el Teorema Chino del Resto,

$$\mathcal{O}_K/\mathfrak{p}\,\mathfrak{q} = \mathcal{O}_K/\mathfrak{p} \times \mathcal{O}_K/\mathfrak{q}\,.$$

Así que, tomando cardinales, $N_{K/\mathbb{Q}}(\mathfrak{p}\,\mathfrak{q}) = N_{K/\mathbb{Q}}(\mathfrak{p})N_{K/\mathbb{Q}}(\mathfrak{q})$.

Comprobemos ahora que $N_{K/\mathbb{Q}}(\mathfrak{p}^n) = N_{K/\mathbb{Q}}(\mathfrak{p})^n$ por inducción en n. El caso $n = 1$ es evidente, así que supongámoslo cierto para n. Consideramos la aplicación

$$\begin{array}{rccc} \pi: & \mathcal{O}_K/\,\mathfrak{p}^{n+1} & \to & \mathcal{O}_K/\,\mathfrak{p}^n \\ & x+\mathfrak{p}^{n+1} & \mapsto & x+\mathfrak{p}^n \end{array}$$

que es claramente sobreyectiva y cuyo núcleo es $\ker\pi = \mathfrak{p}^n\,/\,\mathfrak{p}^{n+1}$. Por el Primer Teorema de Isomorfía tenemos

$$|\mathcal{O}_K/\,\mathfrak{p}^{n+1}\,| = |\,\mathfrak{p}^n\,/\,\mathfrak{p}^{n+1}\,|\cdot|\mathcal{O}_K/\,\mathfrak{p}^n\,|,$$

es decir,

$$N_{K/\mathbb{Q}}(\mathfrak{p}^{n+1}) = |\,\mathfrak{p}^n\,/\,\mathfrak{p}^{n+1}\,|\cdot N_{K/\mathbb{Q}}(\mathfrak{p}^n).$$

Para calcular $|\,\mathfrak{p}^n\,/\,\mathfrak{p}^{n+1}\,|$ tomamos un $p \in \mathfrak{p}\setminus\mathfrak{p}^2$ y consideramos el homomorfismo de grupos

$$\begin{array}{ccc} \mathcal{O}_K/\,\mathfrak{p} & \to & \mathfrak{p}^n\,/\,\mathfrak{p}^{n+1} \\ x+\mathfrak{p} & \mapsto & xp^n+\mathfrak{p}^{n+1} \end{array}$$

Veamos que es isomorfismo:

- Para la inyectividad, si $xp^n \in \mathfrak{p}^{n+1}$, como $p \notin \mathfrak{p}^2$ necesariamente $x \in \mathfrak{p}$, es decir, $x = 0$ en $\mathcal{O}_K/\,\mathfrak{p}$.

- Para la sobreyectividad, tenemos la cadena de ideales

  $$\mathfrak{p}^{n+1} \subsetneq (p^n)+\mathfrak{p}^{n+1} \subseteq \mathfrak{p}^n,$$

  donde la primera es estricta porque $p \notin \mathfrak{p}^2$. Como no hay ideales intermedios entre $\mathfrak{p}^n$ y $\mathfrak{p}^{n+1}$, necesariamente se tiene $(p^n)+\mathfrak{p}^{n+1} = \mathfrak{p}^n$, por lo que la aplicación es sobreyectiva.

Por ser isomorfismo, tenemos

$$|\,\mathfrak{p}^n\,/\,\mathfrak{p}^{n+1}\,| = |\mathcal{O}_K/\,\mathfrak{p}\,| = N_{K/\mathbb{Q}}(\mathfrak{p}),$$

y por tanto,

$$N_{K/\mathbb{Q}}(\mathfrak{p}^{n+1}) = N_{K/\mathbb{Q}}(\mathfrak{p})\cdot N_{K/\mathbb{Q}}(\mathfrak{p}^n) = N_{K/\mathbb{Q}}(\mathfrak{p})^{n+1}.$$

Finalmente, sean $I, J \leq \mathcal{O}_K$ ideales no nulos (si alguno de ellos es cero el resultado es trivial). Los descomponemos en ideales primos:

$$I = \mathfrak{p}_1^{\alpha_1}\ldots\mathfrak{p}_r^{\alpha_r}, \quad J = \mathfrak{q}_1^{\beta_1}\ldots\mathfrak{q}_s^{\beta_s}.$$

Utilizando lo anterior, es trivial comprobar que se cumple $N_{K/\mathbb{Q}}(IJ) = N_{K/\mathbb{Q}}(I)N_{K/\mathbb{Q}}(J)$. □

Definimos ahora qué constante tomamos para cada valor absoluto no arquimediano. Si $\mathfrak{p} \in \operatorname{Spec}(\mathcal{O}_K)$ es no nulo, definimos

$$|x|_{\mathfrak{p}} := N_{K/\mathbb{Q}}(\mathfrak{p})^{-v_{\mathfrak{p}}(x)/[K:\mathbb{Q}]}$$

para $x \in K^\times$. Es decir, tomamos como constante la norma absoluta del ideal primo $\mathfrak{p}$ (que sabemos que es mayor que 1 por definición) elevada a $1/[K:\mathbb{Q}]$. De esto obtenemos el siguiente resultado.

Teorema 2.14 (Fórmula producto)**.** *Si $x \in K^\times$,*

$$\prod_{v \in M} |x|_v = 1.$$

Demostración. Como los valores absolutos son multiplicativos, basta con comprobar la igualdad para $x \in \mathcal{O}_K \setminus \{0\}$.

Supongamos primero que $x \in \mathcal{O}_K^*$. En este caso, $|x|_{\mathfrak{p}} = 1$ para todo $\mathfrak{p} \in \operatorname{Spec}(\mathcal{O}_K)$ no nulo (o, equivalentemente, $|x|_v = 1$ para todo $v \in M^0$). Por tanto, solo sobreviven los valores absolutos arquimedianos $v \in M^\infty$. Si σ es una inmersión de K en $\mathbb{R}$ o $\mathbb{C}$, tenemos

$$\prod_{v \in M} |x|_v = \prod_{v \in M^\infty} |x|_v = \prod_{\sigma} |\sigma(x)| = \left|N_{K/\mathbb{Q}}(x)\right|^{1/[K:\mathbb{Q}]}.$$

La igualdad de arriba se cumple porque permitimos a cada par de inmersiones complejas conjugadas ser consideradas como elementos distintos de M a pesar de inducir el mismo lugar. Como $x \in \mathcal{O}_K^*$, $N_{K/\mathbb{Q}}(x) = \pm 1$, así que $\left|N_{K/\mathbb{Q}}(x)\right|^{1/[K:\mathbb{Q}]} = 1$.

Sea ahora $x \notin \mathcal{O}_K^*$. El ideal $x\mathcal{O}_K$ es propio. Consideramos su descomposición en ideales primos:

$$x\mathcal{O}_K = \mathfrak{p}_1^{\alpha_1} \dots \mathfrak{p}_r^{\alpha_r},$$

donde $\alpha_j \geq 1$. Los únicos términos de $\prod_{v \in M} |x|_v$ que no son necesariamente 1 provienen de los valores absolutos $\mathfrak{p}$-ádicos asociados a $\mathfrak{p}_1, \dots, \mathfrak{p}_r$ y de los arquimedianos. En otras palabras, si $\mathfrak{q} \neq \mathfrak{p}_j$, entonces $|x|_{\mathfrak{q}} = 1$. Por tanto, tenemos lo siguiente:

$$\prod_{v \in M} |x|_v = \prod_{v \in M^\infty} |x|_v \prod_{j=1}^{r} |x|_{\mathfrak{p}_j}.$$

Como en el caso anterior, $\prod_{v \in M^\infty} |x|_v = \left|N_{K/\mathbb{Q}}(x)\right|^{1/[K:\mathbb{Q}]}$, así que solo nos queda por calcular el segundo productorio. Utilizando la definición:

$$\begin{aligned} \prod_{j=1}^{r} |x|_{\mathfrak{p}_j} &= \left(\prod_{j=1}^{r} N_{K/\mathbb{Q}}(\mathfrak{p}_j)^{-\alpha_j} \right)^{1/[K:\mathbb{Q}]} \\ &\overset{\text{Prop 2.13}}{=} N_{K/\mathbb{Q}}(x\mathcal{O}_K)^{-1/[K:\mathbb{Q}]} \\ &\overset{\text{Prop 2.12}}{=} \left|N_{K/\mathbb{Q}}(x)\right|^{-1/[K:\mathbb{Q}]}, \end{aligned}$$

y por tanto concluimos

$$\prod_{v\in M} |x|_v = \left|N_{K/\mathbb{Q}}(x)\right|^{1/[K:\mathbb{Q}]} \cdot \left|N_{K/\mathbb{Q}}(x)\right|^{-1/[K:\mathbb{Q}]} = 1. \qquad \square$$

Observación. En realidad, no es necesario separar los dos casos como hacemos en la demostración. El segundo caso es aplicable a $x \in \mathcal{O}_K$ si relajamos la hipótesis de $\alpha_j \geq 1$ a $\alpha_j \geq 0$.

Ejemplo 2.15. Veamos que la fórmula producto se verifica en $K = \mathbb{Q}$. Ya sabemos que $\mathcal{O}_K = \mathbb{Z}$ y que

$$\mathrm{Spec}(\mathbb{Z}) = \{(p) : p \text{ es primo}\} \cup \{0\}\,.$$

Además del usual, tenemos un valor absoluto no arquimediano para cada número primo p. La valoración asociada a p viene dada por $v_p(x) = m$ si $x = p^m \frac{a}{b}$ con $a, b \in \mathbb{Z}$ y $p \nmid a$, $p \nmid b$. Por otro lado,

$$N_{K/\mathbb{Q}}(p\mathbb{Z}) = |\mathbb{Z}/p\mathbb{Z}| = p,$$

así que el valor absoluto asociado (teniendo en cuenta que $[K : \mathbb{Q}] = 1$) es:

$$|x|_p = p^{-v_p(x)}.$$

Comprobamos ahora la fórmula. En realidad, por ser los valores absolutos multiplicativos, solo hace falta probarla para un $x \in \mathbb{Z}$ no nulo. En tal caso, consideramos su descomposición en primos

$$x = \pm \prod_{j=1}^{r} p_j^{\alpha_j} = \pm p_1^{\alpha_1} \dots p_r^{\alpha_r},$$

donde los $\alpha_j \geq 1$. Esto implica que si q es primo y $q \neq p_j$ entonces $|x|_q = 1$. por tanto, tenemos:

$$\prod_{v\in M} |x|_v = |x| \prod_{j=1}^{r} |x|_{p_j} = \prod_{j=1}^{r} p_j^{\alpha_j} \prod_{j=1}^{r} p_j^{-\alpha_j} = 1.$$

2.4. El grupo de S-unidades

Antes de introducir los conceptos de S-entero y S-unidad, veámoslo primero en $\mathbb{Q}$. El conjunto S va a ser una colección finita de primos, y un S-entero será un número racional cuyo denominador factorice completamente con los primos de S. Una S-unidad, por tanto, será un S-entero cuyo inverso sea también un S-entero. Si $S = \{p_1, p_2, \dots, p_k\}$, y denotamos por $\mathcal{O}_S$ el anillo de los S-enteros, entonces tenemos

$$\mathcal{O}_S = \mathbb{Z}[(p_1 p_2 \dots p_k)^{-1}] \quad \text{y} \quad \mathcal{O}_S^* = \left\{\pm p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k} : \alpha_j \in \mathbb{Z}\right\}.$$

En principio, esta notación podría solaparse con el anillo de enteros $\mathcal{O}_K$, pero empleando letras diferentes (S y K) no habrá confusión.

Para generalizar este concepto a un cuerpo de números K, no podemos emplear directamente la misma idea porque $\mathcal{O}_K$ no tiene por qué ser un dominio de factorización única. Para ello hemos introducido previamente los ideales primos y los valores absolutos $\mathfrak{p}$-ádicos. A partir de ahora S será una colección finita de ideales primos en $\mathcal{O}_K$, es decir,

$$S = \left\{\mathfrak{p}_1, \mathfrak{p}_2, \ldots, \mathfrak{p}_k : \mathfrak{p}_j \in \operatorname{Spec}(\mathcal{O}_K) \text{ no nulos}\right\}.$$

Recordamos que si $\alpha \in K^\times$, entonces existen $x, y \in \mathcal{O}_K$ tales que $\alpha = x/y$ (Lema 1.9); de hecho, se puede suponer que $y \in \mathbb{Z}^+$. Como sí tenemos factorización única en ideales primos, diremos que un $\alpha = x/y \in K^\times$ es *S-entero* si al factorizar el ideal $(\alpha) = (x)(y)^{-1}$ en $\mathcal{O}_K$, las únicas potencias negativas corresponden a ideales primos de S, y diremos que es *S-unidad* si todas las potencias (positivas y negativas) son de elementos de S.

Podemos formalizar algo más nuestra definición usando valores absolutos $\mathfrak{p}$-ádicos:

Definición 2.16. Sea $\alpha \in K^\times$. Diremos que α es *S-entero* si

$$|\alpha|_{\mathfrak{q}} \leq 1 \text{ para todo } \mathfrak{q} \notin S,$$

y diremos que es *S-unidad* si

$$|\alpha|_{\mathfrak{q}} = 1 \text{ para todo } \mathfrak{q} \notin S.$$

Denotaremos por $\mathcal{O}_S$ el anillo de los elementos S-enteros, y por $\mathcal{O}_S^*$ su grupo multiplicativo de las S-unidades.

Para cualquier cuerpo de números K y cualquier elección de S, el grupo de S-unidades $\mathcal{O}_S^*$ satisface las siguientes propiedades:

P1. $\mathcal{O}_S^*$ es un grupo con la multiplicación.

P2. Para cualquier $\alpha \in K - \{0\}$ existe un S tal que $\alpha \in \mathcal{O}_S^*$.

P3. $\mathcal{O}_S^*$ está finitamente generado.

Las dos primeras son consecuencia inmediata de la definición, mientras que la tercera proviene del hecho de que el anillo de enteros es noetheriano.

El resultado principal, que tomaremos como punto de partida y que por ahora solo enunciamos, es el teorema de las S-unidades.

Teorema 2.17 (S-unidades)**.** *La ecuación $x + y = 1$ tiene un número finito de soluciones con $x, y \in \mathcal{O}_S^*$.*

Este será la clave para probar la finitud en las soluciones de numerosas ecuaciones diofánticas. Un primer ejemplo, si $K = \mathbb{Q}$, es el siguiente:

Ejemplo 2.18. Solo hay un número finito de potencias de 2 que, sumadas a una potencia de 5, tienen como resultado una potencia de 3 (como, por ejemplo, $2^2 + 5^1 = 3^2$). En efecto, si tomamos $S = \{(2), (3), (5)\}$, entonces

$$\mathcal{O}_S^* = \{\pm 2^{\alpha_1} 3^{\alpha_2} 5^{\alpha_3} : \alpha_j \in \mathbb{Z}\},$$

y existen un número finito de pares (x, y) tales que $x, y \in \mathcal{O}_S^*$ y $x + y = 1$. De todos ellos seleccionamos aquellos tales que $x = 2^\alpha 3^{-\gamma}$, $y = 5^\beta 3^{-\gamma}$, con $\alpha, \beta, \gamma \in \mathbb{Z}^+$, que nos dan todas las soluciones posibles al problema inicial.

Sin embargo, no nos centraremos en este tipo de ecuaciones exponenciales, sino en las polinómicas. El resultado que consideraremos será el siguiente.

Teorema 2.19. *Sea K un cuerpo de números. Si $P \in K[x, y]$ es homogéneo con $\deg P \geq 3$ y sin factores múltiples, entonces para cualquier $c \in K\backslash\{0\}$ la ecuación $P(x, y) = c$ tiene un número finito de soluciones con $x, y \in \mathcal{O}_S$.*

Observación. Si $c = 0$ es fácil encontrar un contraejemplo. Por ejemplo, si $P(x, y) = x^3 - y^3 \in \mathbb{Q}[x, y]$ y S es cualquier conjunto de primos, entonces $P(x, x) = 0$ para todo $x \in \mathcal{O}_S$.

De este teorema podemos deducir el siguiente resultado.

Corolario 2.20. *Sea $P \in \mathbb{Z}[x, y]$ homogéneo con algún factor irreducible de grado al menos 3, entonces para $c \in \mathbb{Z}\backslash\{0\}$ la ecuación diofántica $P(x, y) = c$ tiene solo un número finito de soluciones enteras.*

Demostración. Sea $P \in \mathbb{Z}[x, y]$ como en el enunciado. Entonces existen polinomios homogéneos $Q, R \in \mathbb{Z}[x, y]$ con $P = QR$ y $\deg Q \geq 3$. Como $c \neq 0$, si $(x, y) \in \mathbb{Z} \times \mathbb{Z}$ tenemos

$$\#\{(x, y) : P(x, y) = c\} = \sum_{d|c} \#\{(x, y) : Q(x, y) = d, R(x, y) = c/d\},$$

donde d recorre los divisores (positivos y negativos) de c. Por el Teorema 2.19, existe un número finito de soluciones $Q(x, y) = d$ para x, y en algún $\mathcal{O}_S$. En particular, sea cual sea el S, existe una cantidad finita de soluciones enteras. Esto implica que también son finitas la soluciones enteras de $P(x, y) = c$. □

Ejemplo 2.21. Para $n, d, c \in \mathbb{Z}^+$, consideramos el polinomio de coeficientes enteros $P(x, y) = x^n - dy^n$. Veamos que para $n \geq 3$ podemos aplicar el resultado anterior. Si llamamos $t = x/y$, podemos considerar el polinomio

$$f(t) = P(t, 1) = t^n - d.$$

Para ver que f no tiene factores múltiples, recordamos el siguiente resultado elemental de Teoría de Galois:

$$f \text{ no tiene factores múltiples} \Leftrightarrow \operatorname{mcd}(f, f') = 1,$$

donde f' denota la derivada (formal) de f. Utilizando esto, nuestro ejemplo se vuelve trivial, pues $f'(t) = nt^{n-1}$, y $\mathrm{mcd}(t^n - d, nt^{n-1}) = 1$.

Como f no tiene factores múltiples, P tampoco los tiene, luego podemos aplicar el teorema y concluir que existe una cantidad finita de soluciones enteras a la ecuación $x^n - dy^n = c$ siempre que $n \geq 3$.

Por tanto, solo puede haber infinitas soluciones en los casos $n = 1$ y $n = 2$. De hecho, si $n = 2$ y d no es un cuadrado perfecto, obtenemos la ecuación de Pell $x^2 - dy^2 = c$, que tiene infinitas soluciones.

Nuestro objetivo será probar la equivalencia entre los Teoremas 2.17 y 2.19, que a priori hablan de cosas distintas. La implicación verdaderamente interesante es cómo del teorema de las S-unidades se deduce la finitud de soluciones S-enteras en un polinomio homogéneo, aunque veremos ambas por establecer la equivalencia.

Demostración (del Teorema 2.19 usando el Teorema 2.17). Sea $P \in K[x, y]$ con las hipótesis del teorema. Escribimos

$$P = \sum_{j=1}^{d} \lambda_j x^j y^{d-j},$$

donde $\lambda_j \in K$. Para factorizar el polinomio, sacamos factor común y^d

$$P = y^d \sum_{j=1}^{d} \lambda_j \frac{x^j}{y^j}.$$

Si llamamos $t = x/y$ entonces tenemos un polinomio en una variable

$$f(t) = P(t, 1) = \sum_{j=1}^{d} \lambda_j t^j.$$

Como $K \subset \mathbb{C}$, el polinomio factoriza

$$f(t) = \prod_{j=1}^{d} (a_j t + b_j),$$

donde $a_j, b_j \in \mathbb{C}$. Finalmente reescribimos P:

$$P = y^d \prod_{j=1}^{d} \left(a_j \frac{x}{y} + b_j \right) = \prod_{j=1}^{d} (a_j x + b_j y) \in \mathbb{C}[x, y],$$

y de hecho podemos suponer para la prueba que $a_j, b_j \in K$; si no están en K, estarán en alguna extensión finita suya, y podemos ampliar el cuerpo para incluirlos. Esto hace más grande el anillo $\mathcal{O}_S$, pero demostrar la finitud

de soluciones aquí es suficiente para probarlo en el anillo original. Por P2, podemos extender S de manera que $a_j, b_j, c \in \mathcal{O}_S^*$.

Supongamos que $(x, y) \in \mathcal{O}_S \times \mathcal{O}_S$ es solución de $P(x, y) = c$. Si definimos $c_j = a_j x + b_j y$, entonces

$$\prod_{j=1}^{d}(a_j x + b_j y) = \prod_{j=1}^{d} c_j = c.$$

Si $\mathfrak{q} \notin S$,

$$|c_j|_{\mathfrak{q}} = |a_j x + b_j y|_{\mathfrak{q}} \leq \max(|a_j x|_{\mathfrak{q}}, |b_j y|_{\mathfrak{q}}) \leq 1,$$

así que $c_j \in \mathcal{O}_S$.

Como $d \geq 3$, definimos

$$d_1 = a_2 b_3 - a_3 b_2, \quad d_2 = a_1 b_3 - a_3 b_1 \quad \text{y} \quad d_3 = a_1 b_2 - a_2 b_1.$$

Los tres satisfacen $d_1, d_2, d_3 \in K \setminus \{0\}$, ya que si alguno fuera nulo entonces P tendría factores múltiples. De nuevo, podemos ampliar S de tal forma que d_1, d_2 y d_3 sean S-unidades. Como $\mathcal{O}_S^*$ es un grupo (P1), los elementos

$$X = \frac{c_1 d_1}{c_2 d_2}, \quad Y = \frac{c_3 d_3}{c_2 d_2},$$

también son S-unidades, que además satisfacen

$$\begin{aligned} X + Y &= \frac{c_1 d_1 + c_3 d_3}{c_2 d_2} \\ &= \frac{(a_1 x + b_1 y)(a_2 b_3 - a_3 b_2) + (a_3 x + b_3 y)(a_1 b_2 - a_2 b_1)}{c_2 d_2} \\ &= \frac{a_1 a_2 b_3 x - b_1 b_2 a_3 y + a_1 b_2 b_3 y - b_1 a_2 a_3 x}{c_2 d_2} \\ &= \frac{(a_2 x + b_2 y)(a_1 b_3 - a_3 b_1)}{c_2 d_2} \\ &= \frac{c_2 d_2}{c_2 d_2} \\ &= 1. \end{aligned}$$

Como el conjunto de pares $(X, Y) \in \mathcal{O}_S^* \times \mathcal{O}_S^*$ tales que $X + Y = 1$ es finito, escribimos

$$X = \frac{c_1 d_1}{c_2 d_2} = \frac{a_1 x + b_1 y}{a_2 x + b_2 y}\frac{d_1}{d_2} = \frac{a_1(x/y) + b_1}{a_2(x/y) + b_2}\frac{d_1}{d_2}.$$

Así que el conjunto de valores que toma x/y es finito, digamos $\{t_1, \ldots, t_N\}$. Para un $j = 1, \ldots, N$ fijo tenemos

$$P(x, y) = c, \qquad \frac{x}{y} = t_j,$$

y por tanto los posibles valores de y corresponden a las soluciones de la ecuación $P(t_j y, y) = c$, que son finitas por tratarse de una ecuación polinómica en una variable. Si consideramos todos los t_j, concluimos que los posibles valores de y también son finitos, y como $x/y = t_j$ para algún j, los de x también lo son, lo que concluye la prueba. □

Demostración (del Teorema 2.17 usando el Teorema 2.19). Por (P3), el grupo de unidades $\mathcal{O}_S^*$ está finitamente generado, digamos

$$\mathcal{O}_S^* = \langle g_1, \ldots, g_r \rangle.$$

Esto quiere decir que

$$\mathcal{O}_S^* = \{ g_1^{\alpha_1} \ldots g_r^{\alpha_r} : \alpha_j \in \mathbb{Z} \}.$$

Consideramos el subgrupo $(\mathcal{O}_S^*)^3 = \left\{ g^3 : g \in \mathcal{O}_S^* \right\}$ que es normal trivialmente por ser $\mathcal{O}_S^*$ conmutativo. El grupo cociente $\mathcal{O}_S^*/(\mathcal{O}_S^*)^3$ es finito, de hecho

$$\mathcal{O}_S^*/(\mathcal{O}_S^*)^3 = \{ [g_1^{\alpha_1} \ldots g_r^{\alpha_r}] : \alpha_j \in \{0,1,2\} \}.$$

Por tanto, cualquier elemento $g \in \mathcal{O}_S^*$ puede escribirse como $g_1^{\alpha_1} \ldots g_r^{\alpha_r} h^3$ donde $\alpha_j \in \{0,1,2\}$ y $h \in \mathcal{O}_S^*$.

Llamamos $h_1 \ldots, h_N$ a un conjunto completo de representantes de las clases de $\mathcal{O}_S^*/(\mathcal{O}_S^*)^3$ ($N \leq r^3$). Supongamos entonces que $(x,y) \in \mathcal{O}_S^* \times \mathcal{O}_S^*$ satisface $x + y = 1$. Por tanto, existen h_i, h_j y $X, Y \in \mathcal{O}_S^*$ tales que

$$x = h_i X^3, \quad y = h_j Y^3.$$

Consideramos los polinomios $P_{i,j} = h_i X^3 + h_j Y^3, 1 \leq i, j \leq N$. No tienen factores múltiples, puesto que su factorización es

$$h_i X^3 + h_j Y^3 = \prod_{k=0}^{2} (\sqrt[3]{h_i} X + \sqrt[3]{h_j} \omega^k Y),$$

donde $\omega = -\frac{1}{2} + i\frac{\sqrt{3}}{2} \in \mathbb{C}$ (una raíz cúbica primitiva de la unidad). Como para los polinomios $P_{i,j}$ existe una cantidad finita de $X, Y \in \mathcal{O}_S$ tales que $P_{i,j}(X,Y) = 1$, tenemos también una cantidad finita de $(x,y) \in \mathcal{O}_S^*$ tales que $x + y = 1$. □

Capítulo 3

El teorema de las S-unidades

El objetivo final del trabajo y el contenido de este capítulo es la demostración del teorema de las S-unidades, ya enunciado en el capítulo anterior. Para ello necesitaremos introducir el concepto de *altura* en cuerpos de números y explorar muy a fondo sus propiedades.

3.1. Alturas

Sea K un cuerpo de números. Antes de entrar a hablar de alturas, recordemos la definición del espacio proyectivo y su construcción.

Consideramos el conjunto $K^{n+1}\setminus\{0\}$. En él definimos la relación de equivalencia

$$(x_0,\dots,x_n)\sim(x_0',\dots,x_n')\Leftrightarrow\exists\lambda\in K^\times,(x_0',\dots,x_n')=(\lambda x_0,\dots,\lambda x_n).$$

El cociente,

$$\mathbb{P}^n:=\mathbb{P}^n(K):=\left(K^{n+1}\setminus\{0\}\right)/\sim,$$

es el *espacio proyectivo* n-dimensional sobre K. Denotamos por $[x_0:\dots:x_n]$ a la clase de $(x_0,\dots,x_n)$.

Definición 3.1. Sea $P=[x_0:\dots:x_n]\in\mathbb{P}^n(K)$. Definimos la aplicación

$$H_K:\mathbb{P}^n(K)\to\mathbb{R}$$

dada por la expresión

$$H_K(P):=\prod_{v\in M}\sup_j|x_j|_v.$$

Diremos que $H_K(P)$ es la *altura* de P.

Lo primero que debemos comprobar es que esta aplicación está bien definida. En efecto, si $\lambda \in K^\times$ y $P = [\lambda x_0 : \ldots : \lambda x_n]$, tenemos

$$\sup_j |\lambda x_j|_v = |\lambda|_v \sup_j |x_j|_v,$$

de tal manera que

$$\prod_{v\in M} \sup_j |\lambda x_j|_v = \left(\prod_{v\in M} |\lambda|_v\right)\left(\prod_{v\in M} \sup_j |x_j|_v\right) = \prod_{v\in M} \sup_j |x_j|_v,$$

donde utilizamos la fórmula producto aplicada a λ. Esto demuestra que la altura de $P \in \mathbb{P}^n(K)$ está bien definida, puesto que no depende del representante elegido. En particular, podemos asumir que una coordenada de P es 1, de tal manera que $H_K(P) \geq 1$ siempre.

Notación. $H_K(x_0 : \ldots : x_n) := H_K([x_0 : \ldots : x_n])$.

Sea ahora L/K una extensión finita, de manera que L también es un cuerpo de números. Tomamos un $P = [x_0 : \ldots : x_n] \in \mathbb{P}^n(K) \subset \mathbb{P}^n(L)$. Queremos comparar las alturas de P con respecto a K y a L. Se puede probar (ver [17]) que todo valor absoluto $w \in M_L$ puede ser entendido como una extensión de otro $v \in M_K$, y que, si escribimos $w \mid v$ para denotar que w extiende a v, se tiene la identidad

$$\prod_{w|v} |x|_w = |x|_v$$

para todo $x \in K$. Por tanto,

$$H_L(P) = \prod_{w\in M_L} \sup_j |x_j|_w = \prod_{v\in M_K} \prod_{w|v} \sup_j |x_j|_w = \prod_{v\in M_K} \sup_j |x_j|_v = H_K(P).$$

Esto implica que, de hecho, tenemos una *altura absoluta* definida para todo $P \in \mathbb{P}^n(\overline{\mathbb{Q}})$, y que no depende del cuerpo de números K al cual x pertenece. Eso nos permite olvidarnos del subíndice K de ahora en adelante.

Es habitual también trabajar con la altura logarítmica:

$$h := \log H.$$

Como $H(P) \geq 1$, satisface $h(P) \geq 0$. Esta altura logarítmica proviene de una versión aditiva de la fórmula producto, más natural en otros contextos. También definimos la altura de un elemento $x \in \overline{\mathbb{Q}}$ como

$$H(x) := H(1 : x), \quad h(x) := h(1 : x) = \log H(x)$$

En particular, $H(0) = H(1) = 1$ y $h(0) = h(1) = 0$. Veamos algunas propiedades generales.

Proposición 3.2. *Sean $x, x_1, \ldots, x_r \in \overline{\mathbb{Q}}$:*

(1) $h(x) \geq 0$.

(2) $h(1/x) = h(x)$ *y, más generalmente,* $h(x^m) = |m|h(x)$ *para* $m \in \mathbb{Z}$.

(3) $\sup_j h(x_j) \leq h(1 : x_1 : \ldots : x_r) \leq h(x_1) + \cdots + h(x_r)$.

(4) $h(x_1 \ldots x_r) \leq h(x_1) + \cdots + h(x_r)$.

(5) $h(x_1 + \cdots + x_r) \leq h(1 : x_1 : \ldots : x_r) + \log r$.

(6) $h(\sigma(x)) = h(x)$ *para todo* $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.

Demostración. (1): Trivial sabiendo que $H(x) \geq 1$.

(2): $H(1/x) = H(1 : 1/x) = H(x : 1) = H(x)$. Esto prueba la primera parte, y nos permite suponer para la segunda que $m \geq 0$. Para probarlo basta con observar que

$$H(x^m) = \prod_{v \in M} \sup\{1, |x^m|_v\} = \prod_{v \in M} \sup\{1, |x|_v\}^m = H(x)^m$$

y tomar logaritmos.

(3): Tenemos la cadena de desigualdades

$$\begin{aligned} \sup_j H(x_j) &= \sup_j H(1 : x_j) = \sup_j \prod_{v \in M} \sup\{1, |x_j|_v\} \\ &\leq \prod_{v \in M} \sup\{1, |x_1|_v, \ldots, |x_r|_v\} = H(1 : x_1 : \ldots : x_r) \\ &\leq \prod_{j=1}^{r} \prod_{v \in M} \sup\{1, |x_j|_v\} = H(x_1) \ldots H(x_r), \end{aligned}$$

y tomando logaritmos obtenemos

$$\sup_j h(x_j) \leq h(1 : x_1 : \ldots : x_r) \leq h(x_1) + \cdots + h(x_r).$$

(4): De nuevo, basta con tomar logaritmos en

$$\begin{aligned} H(x_1 \cdots x_r) &= \prod_{v \in M} \sup\{1, |x_1 \cdots x_r|_v\} \\ &\leq \prod_{v \in M} \prod_{j=1}^{r} \sup\{1, |x_j|_v\} \\ &= H(x_1) \cdots H(x_r). \end{aligned}$$

(5): Distinguimos entre si v es arquimediano o no. En el primer caso, si $v \in M^\infty$,

$$|x_1 + \cdots + x_r|_v \leq \sup_j |r x_j|_v = |r|_v \sup_j |x_j|_v.$$

En el segundo caso, si $v \in M^0$, podemos usar la desigualdad ultramétrica, de tal forma que

$$|x_1 + \cdots + x_r| \leq \sup_j |x_j|_v.$$

Además, $\prod_{v \in M^\infty} |r|_v = r$ puesto que $r \in \mathbb{Z}^+$, así que

$$\begin{aligned} H(x_1 + \cdots + x_r) &= \prod_{v \in M} \sup\{1, |x_1 + \cdots + x_r|_v\} \\ &\leq r \prod_{v \in M} \sup\left\{1, \sup_j |x_j|_v\right\} \\ &= rH(1 : x_1 : \ldots : x_r), \end{aligned}$$

y tomando logaritmos se obtiene

$$h(x_1 + \cdots + x_r) \leq h(1 : x_1 : \ldots : x_r) + \log r.$$

(6): Es inmediato observando que, por definición, $|\sigma(x)|_v = |x|_{v^\sigma}$, de tal manera que la conjugación actúa como una permutación en los valores absolutos, dejando invariantes tanto $H(x)$ como $h(x)$. □

Ejemplo 3.3. Sea $K = \mathbb{Q}$. Vamos a buscar una fórmula explícita para la altura de un $a/b \in \mathbb{Q}^\times$. Recordamos que para cualquier número racional, y en particular para un $x \in \mathbb{Z}$ no nulo, tenemos el único valor absoluto arquimediano, que resulta ser el usual, y uno no arquimediano para cada primo p dado por

$$|x|_p = p^{-v_p(x)},$$

donde $v_p(x) = m$ si $x = p^m c$ con $p \nmid c$. En particular, si x es entero, $|x|_p \leq 1$ para todo primo p.

Tomamos ahora un $a/b \in \mathbb{Q}^\times$ arbitrario y con la fracción irreducible. Se tiene

$$h(a/b) = h(1 : a/b) = h(b : a) = \sum_v \log \sup\{|a|_v, |b|_v\}.$$

Ahora bien, si v es no arquimediano, necesariamente $\sup\{|a|_v, |b|_v\} = 1$, pues si no fuera así existiría un primo p tal que $p \mid a$ y $p \mid b$, contradiciendo así la irreducibilidad de a/b. Por tanto todos los términos de la suma correspondientes a valores absolutos no arquimedianos son nulos, y obtenemos

$$h(a/b) = \sup\{\log|a|, \log|b|\}.$$

Un resultado fundamental que será clave en la demostración del teorema de las S-unidades es el siguiente.

Teorema 3.4 (Northcott)**.** *Para cualesquiera enteros positivos B y d, solo existe una cantidad finita de números algebraicos $\alpha \in \overline{\mathbb{Q}}$ tales que $h(\alpha) \leq B$ y $\deg(\alpha) \leq d$.*[1]

En consecuencia, para cualquier n dado, solo existe una cantidad finita de puntos $P \in \mathbb{P}^n(\overline{\mathbb{Q}})$ con altura acotada y grado acotado.

Demostración. Para $d = 1$ la afirmación anterior se reduce al hecho de que solo existe un número finito de números racionales con altura acotada por B. Esto se deduce del ejemplo anterior de manera trivial gracias a la fórmula $h(a/b) = \sup\{\log|a|, \log|b|\}$ para $a, b \in \mathbb{Z}$ coprimos.

El caso general se puede reducir a esto. Sea $\alpha \in \overline{\mathbb{Q}}$ con $\deg(\alpha) \leq d$ y $h(\alpha) \leq B$. Por la Proposición 3.2 la altura es invariante por los automorfismos de Galois, de manera que $h(\sigma(\alpha)) \leq B$ para todos los como mucho d conjugados $\sigma(\alpha)$ de α. Si $\deg(\alpha) = d_0 \leq d$, para $0 \leq r \leq d_0$ consideramos los *polinomios simétricos elementales*

$$e_r(X_1, X_2, \ldots, X_{d_0}) = \sum_{1 \leq j_1 < \ldots < j_r \leq d_0} X_{j_1} \cdots X_{j_r}.$$

Por la Proposición 3.2,

$$\begin{aligned} h(e_r(\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(d_0)})) &\leq \sum_{1 \leq j_1 < \ldots < j_r \leq d_0} h(\alpha^{(j_1)} \cdots \alpha^{(j_r)}) + \log \binom{d_0}{r} \\ &\leq \binom{d_0}{r} r h(\alpha) + \log \binom{d_0}{r} \\ &\leq \binom{d}{r} (r h(\alpha) + 1) \\ &\leq 2^d (dB + 1). \end{aligned}$$

Por otro lado, como $e_r(\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(d_0)})$ es invariante por cualquier permutación de los $\alpha^{(j)}$ y por tanto por cualquier automorfismo del grupo de Galois que generan, se tiene que $e_r(\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(d_0)}) \in \mathbb{Q}$ para cualquier r, por lo que su altura, y en particular la de α, están acotadas.

Por último, la afirmación de los puntos proyectivos es consecuencia de las propiedades de la Proposición 3.2. □

Veamos ahora un par de lemas previos a la prueba del teorema de las S-unidades.

Lema 3.5. *Sean $P_j = a_{0j} + a_{1j}x + \cdots + a_{nj}x^n \in \mathbb{Z}[x]$ con $j = 1, \ldots, m$ y $x \in K$ que no sea un cero común a todos los P_j. Entonces*

$$h(P_1(x) : \ldots : P_m(x)) \leq n h(x) + \log(n+1) + \log \sup_{k,j} |a_{kj}|.$$

[1] Recordamos que el grado de un número algebraico es el grado de su polinomio mínimo.

Demostración. Si v es un valor absoluto no arquimediano, por la desigualdad ultramétrica tenemos

$$|P_j(x)|_v \leq \sup\left\{|a_{0j}|_v, |a_{1j}x|_v, \ldots, |a_{nj}x^n|_v\right\} = |a_{k_0 j}|_v |x|_v^{k_0}$$

para algún $0 \leq k_0 \leq n$. Además, $|a_{k_0 j}|_v \leq 1$ porque los coeficientes de P_j son enteros y se sigue

$$\log |P_j(x)|_v \leq n \log^+ |x|_v$$

donde $\log^+ = \log$ en $\mathbb{R}_{>1}$ y es nulo en $\mathbb{R}_{\leq 1}$. Lo introducimos por si se da el caso $k_0 = 0$.

Por otro lado, para cualquier valor absoluto (arquimediano o no), se cumple por la desigualdad triangular

$$|x_0 + x_1 + \cdots + x_n|_v \leq (n+1) \sup_j |x_j|_v .$$

Con la normalización que estamos usando, para cada v arquimediana existe un $q \geq 1$ tal que $|x|_v$ sigue siendo valor absoluto al elevar a q y coincide con el usual en $\mathbb{Q}$. Por consiguiente

$$|P_j(x)|_v^q \leq (n+1)\left|a_{k_0 j}x^{k_0}\right|_v^q = \left|(n+1)a_{k_0 j}x^{k_0}\right|_v^q$$

para algún $0 \leq k_0 \leq n$. Elevando a $1/q$ y tomando logaritmos

$$\log |P_j(x)|_v \leq \log |(n+1)a_{k_0 j}|_v + n \log^+ |x|_v .$$

De ambas cotas para $\log |P_j(x)|_v$ concluimos

$$\begin{aligned} h(P_1(x) : \ldots : P_m(x)) &= \sum_v \sup_j \log |P_j(x)|_v \\ &\leq \sum_{v \text{ arqu.}} \sup_j \log |(n+1)|a_{k_0 j}||_v + nh(x) \\ &\leq \sum_{v \text{ arqu.}} \log \left|(n+1)\sup_{k,j} |a_{kj}|\right|_v + nh(x) \\ &\leq \log(n+1) + \log \sup_{k,j} |a_{kj}| + nh(x), \end{aligned}$$

lo que termina la prueba. □

Lema 3.6. *Para $x \in K^\times$, $h(x) = 0$ si y solo si x es una raíz de la unidad.*

Demostración. ($\Rightarrow$) Supongamos que $h(x) = 0$. Entonces, por la Proposición 3.2, $h(x^n) = nh(x) = 0$ para $n = 0, 1, 2, \ldots$ Por el teorema de Northcott, el conjunto $\left\{1, x, x^2, \ldots\right\}$ ha de ser necesariamente finito, así que existen $n, m \in \mathbb{Z}_{\geq 0}$ con $x^n = x^m$. Sin pérdida de generalidad, podemos asumir que $n < m$, de tal manera que $x^{m-n} = 1$ y x es una raíz de la unidad.

($\Leftarrow$) Si $x^n = 1$ para $n \in \mathbb{Z}^+$, de nuevo por la Proposición 3.2 se tiene

$$0 = h(1) = h(x^n) = nh(x),$$

y por tanto $h(x) = 0$. □

Ahora queremos extender la definición de altura a puntos no proyectivos en $(K^\times)^n$. El motivo es que, eventualmente, querremos asignarle un tamaño a una solución de $x + y = 1$. Definimos

$$h(x_1, x_2, \ldots, x_n) = h(x_1) + h(x_2) + \cdots + h(x_n).$$

En la línea de lo dicho, estaremos interesados en el caso $n = 2$. En particular, si S es una colección finita de ideales primos de $\mathcal{O}_K$ (tal y como definimos en el Capítulo 2), consideramos

$$\Gamma = \mathcal{O}_S^* \times \mathcal{O}_S^*$$

Este es un grupo con la operación coordenada a coordenada y sabemos que está finitamente generado. Si eliminamos las raíces de la unidad, es decir, cocientamos por su subgrupo de torsión, tendremos un grupo abeliano libre de rango r y por tanto existirá un isomorfismo

$$\varphi : \Gamma/\Gamma_{\text{tor}} \to \mathbb{Z}^r.$$

A partir de este isomorfismo, podemos medir el tamaño de vectores enteros.

Proposición 3.7. *Para $\vec{n} \in \mathbb{Z}^r$, la expresión*

$$\|\vec{n}\| := h(\varphi^{-1}(\vec{n}))$$

define una norma en el retículo $\mathbb{Z}^r$.

Demostración. Por el Lema 3.6, la norma está bien definida, $\|\vec{n}\| \geq 0$ y $\|\vec{n}\| = 0$ si y solo si $\vec{n} = \vec{0}$.

Para la homogeneidad y la desigualdad triangular utilizamos las propiedades de la Proposición 3.2. Si $k \in \mathbb{Z}$ tenemos

$$\|k\vec{n}\| = h(\varphi^{-1}(k\vec{n})) = h(\varphi^{-1}(\vec{n})^k) = |k|h(\varphi^{-1}(\vec{n})) = |k|\|\vec{n}\|$$

Por otro lado, si $\vec{n}_1, \vec{n}_2 \in \mathbb{Z}^r$,

$$\begin{aligned} \|\vec{n}_1 + \vec{n}_2\| &= h(\varphi^{-1}(\vec{n}_1 + \vec{n}_2)) = h(\varphi^{-1}(\vec{n}_1)\varphi^{-1}(\vec{n}_2)) \\ &\leq h(\varphi^{-1}(\vec{n}_1)) + h(\varphi^{-1}(\vec{n}_2)) = \|\vec{n}_1\| + \|\vec{n}_2\|, \end{aligned}$$

y por tanto $\|\cdot\|$ es una norma. □

Con esta notación de normas, el teorema de Northcott se enunciaría, en una versión restringida, como:

Teorema 3.8 (Northcott)**.** *Para $M \in \mathbb{R}^+$ una constante fijada, el conjunto $\{\vec{n} \in \mathbb{Z}^r : \|\vec{n}\| \leq M\}$ es finito.*

Aunque solo lo utilizaremos en una ocasión, es importante ver que esta norma puede ser extendida a todo $\mathbb{R}^r$. Extenderla a $\mathbb{Q}^r$ es sencillo: si $\vec{v} \in \mathbb{Q}^r$ siempre lo podemos escribir como $\lambda\vec{n}$ con $\vec{n} \in \mathbb{Z}^r$ y decretar

$$\|\vec{v}\| = |\lambda|\|\vec{n}\|.$$

Ahora para extenderla a $\mathbb{R}^r$ lo natural sería utilizar la densidad de $\mathbb{Q}^r$ en $\mathbb{R}^r$ y definir, para $\vec{u} \in \mathbb{R}^r$,

$$\|\vec{u}\| = \lim \|\vec{v}_n\|, \text{ para } \vec{v}_n \in \mathbb{Q}^r \text{ con } \lim \vec{v}_n = \vec{u}.$$

Esto, en general, no induce una norma en $\mathbb{R}^r$ porque en este proceso la propiedad de ser definida positiva puede perderse (aunque siempre se obtiene una *seminorma*). En nuestro caso particular esto no ocurre, esencialmente porque las normas de los vectores $\vec{n} \in \mathbb{Z}^r$ no nulos son siempre mayores que una constante. Para ver más detalles se puede consultar [1, Lemma 3.1].

3.2. Demostración del teorema de las S-unidades

Volvamos a enunciar el teorema.

Teorema 3.9 (S-unidades)**.** *La ecuación $x + y = 1$ tiene un número finito de soluciones con $x, y \in \mathcal{O}_S^*$.*

Para poder demostrarlo, necesitaremos un resultado fundamental previo cuya prueba dejamos para después. Lo enunciamos también.

Teorema 3.10 (Resultado fundamental)**.** *Si $\vec{u}, \vec{v} \in \mathbb{Z}^r$ son vectores correspondientes a soluciones de la ecuación del teorema de las S-unidades, existe una constante c_0 tal que*

$$\|\vec{u}\| \leq \frac{1}{4}c_0 + \frac{2}{n-1}\|\vec{v} - 2n\vec{u}\|, \text{ para todo } n \in \mathbb{Z}_{>1}.$$

Por supuesto, la constante $1/4$ que acompaña a c_0 es prescindible y solo la ponemos para que futuros resultados queden más limpios. El por qué este resultado implica el teorema de las S-unidades requiere una explicación. Su importancia radica en que nos permitirá deducir que hay pocos vectores correspondientes a soluciones en direcciones parecidas. La idea de la prueba pasará por considerar una cantidad suficientemente grande de direcciones de tal forma que todas las del espacio estén cerca de alguna de ellas.

Primero veremos cómo deducir el teorema de las S-unidades de este resultado. Para ello, necesitaremos un lema previo.

Lema 3.11. *Para $\vec{u}, \vec{v} \in \mathbb{Z}^r$ cualesquiera y para $n \in \mathbb{Z}_{>1}$ se tiene*

$$\|\vec{v} - 2n\vec{u}\|\|\vec{u}\| \leq \big|\|\vec{v}\|\vec{u}\| - \vec{u}\|\vec{v}\|\big| + \|\vec{u}\|\big|\|\vec{v}\| - 2n\|\vec{u}\|\big|.$$

Demostración. Se tiene

$$\begin{aligned}\|\vec{v}-2n\vec{u}\|\|\vec{u}\| &= \big\|\|\vec{u}\|\,(\vec{v}-2n\vec{u})\,\big\| = \big\|\vec{v}\|\vec{u}\| - 2n\vec{u}\|\vec{u}\|\big\| \\ &= \big\|\vec{v}\|\vec{u}\| - \vec{u}\|\vec{v}\| + \vec{u}\|\vec{v}\| - 2n\vec{u}\|\vec{u}\|\big\| \\ &\leq \big\|\vec{v}\|\vec{u}\| - \vec{u}\|\vec{v}\|\big\| + \big\|\vec{u}\|\vec{v}\| - 2n\vec{u}\|\vec{u}\|\big\| \\ &= \big\|\vec{v}\|\vec{u}\| - \vec{u}\|\vec{v}\|\big\| + \|\vec{u}\|\big|\|\vec{v}\| - 2n\|\vec{u}\|\big|,\end{aligned}$$

lo que concluye la prueba. □

Este resultado es genérico, en el sentido de que es válido para $\vec{u},\vec{v}\in\mathbb{Z}^r$ cualesquiera. Volvamos al caso en el que $\vec{u},\vec{v}\in\mathbb{Z}^r$ son vectores correspondientes a soluciones de la ecuación del teorema de las S-unidades, para los cuales tenemos el siguiente resultado.

Proposición 3.12. *Con la notación del Teorema 3.10, si además $\vec{u},\vec{v}\in\mathbb{Z}^r$ satisfacen*

$$\big\|\vec{v}\|\vec{u}\| - \vec{u}\|\vec{v}\|\big\| \leq \frac{1}{20}\|\vec{u}\|\|\vec{v}\|,$$

entonces o bien $\|\vec{v}\|\leq 20\|\vec{u}\|$ o bien $\|\vec{u}\|\leq c_0$.

Demostración. Por el Teorema 3.10 y el Lema 3.11,

$$\begin{aligned}\|\vec{u}\| &\leq \frac{1}{4}c_0 + \frac{2}{n-1}\|\vec{v}-2n\vec{u}\| \\ &\leq \frac{1}{4}c_0 + \frac{2}{n-1}\left(\frac{1}{\|\vec{u}\|}\left(\big\|\vec{v}\|\vec{u}\| - \vec{u}\|\vec{v}\|\big\| + \|\vec{u}\|\big|\|\vec{v}\| - 2n\|\vec{u}\|\big|\right)\right) \\ &\leq \frac{1}{4}c_0 + \frac{2}{n-1}\left(\frac{1}{\|\vec{u}\|}\left(\frac{1}{20}\|\vec{u}\|\|\vec{v}\| + \|\vec{u}\|\big|\|\vec{v}\| - 2n\|\vec{u}\|\big|\right)\right) \\ &= \frac{1}{4}c_0 + \frac{2}{n-1}\left(\frac{1}{20}\|\vec{v}\| + \|\vec{v}\| - 2n\|\vec{u}\|\right) \\ &= \frac{1}{4}c_0 + \frac{2}{n-1}\left(\frac{21}{20}\|\vec{v}\| - 2n\|\vec{u}\|\right).\end{aligned}$$

Por tanto,

$$20(n-1)\|\vec{u}\| \leq 5(n-1)c_0 + 42\|\vec{v}\| - 20(4n)\|\vec{u}\|,$$

de lo que finalmente obtenemos

$$20(5n-1)\|\vec{u}\| \leq 5(n-1)c_0 + 42\|\vec{v}\|.$$

Supongamos ahora que $\|\vec{v}\| > 20\|\vec{u}\|$ y que $\|\vec{u}\| > c_0$. Entonces, si tomamos $2n$ como el mayor número par que satisface $\|\vec{v}\|\geq 2n\|\vec{u}\|$, de tal forma que $\|\vec{v}\| < (2n+2)\|\vec{u}\|$, se tiene

$$\begin{aligned}20(5n-1)\|\vec{u}\| &\leq 5(n-1)c_0 + 42\|\vec{v}\| \\ &< 5(n-1)\|\vec{u}\| + 42(2n+2)\|\vec{u}\| \\ &= (89n+79)\|\vec{u}\|,\end{aligned}$$

es decir,

$$100n - 20 < 89n + 79 \Leftrightarrow n < 9.$$

Pero recordemos que $\|\vec{v}\| > 20\|\vec{u}\|$, así que n debía ser al menos 10, lo que nos lleva a la contradicción deseada. □

Ya casi tenemos todo lo necesario para demostrar el teorema de las S-unidades. El último ingrediente que necesitamos es considerar, para cada vector unitario $\vec{t} \in \mathbb{R}^r$, el *cono*

$$C(\vec{t}) = \left\{ \vec{x} \in \mathbb{R}^r \setminus \{0\} : \left\| \frac{\vec{x}}{\|\vec{x}\|} - \vec{t} \right\| < \frac{1}{40} \right\}.$$

Es, en efecto, un cono pues si $\vec{x} \in C(\vec{t})$, entonces $\lambda\vec{x} \in C(\vec{t})$ para todo $\lambda \in \mathbb{R}^+$. En particular, es conveniente pensar que un vector $\vec{x}$ pertenece al cono $C(\vec{t})$ si y solo si pertenece dicho vector normalizado. Considerando solo los representantes de cada vector en la esfera $(n-1)$-dimensional $\mathbb{S}^{n-1}$, es inmediato ver que cada uno de estos conos representa un abierto en la topología inducida de la esfera. Como $\mathbb{S}^{n-1}$ es compacta, existe una colección finita de conos que la cubren por completo, y que también cubren todo el espacio $\mathbb{R}^r \setminus \{0\}$.

En particular, podemos afirmar que existe una colección finita de vectores $\vec{t}_1, \ldots, \vec{t}_J \in \mathbb{S}^{n-1}$ tales que $\bigcup C(\vec{t}_j)$ contiene todos los vectores $\vec{u}$ correspondientes a soluciones de la ecuación del teorema de las S-unidades con $\|\vec{u}\| > c_0$ y cada $C(\vec{t}_j)$ contiene al menos uno de ellos.

Con esto, somos capaces por fin de dar la demostración del teorema de las S-unidades.

Demostración del Teorema 3.9 (de las S-unidades). Consideramos la colección finita de conos $C(\vec{t}_j)$ descrita anteriormente. Escogemos un $\vec{u}_j \in \mathbb{Z}^r$ vector solución en cada $C(\vec{t}_j) \bigcap \{\|\vec{x}\| > c_0\}$ (que existe por construcción).

Sea ahora $\vec{v} \in \mathbb{Z}^r$ cualquier otro vector correspondiente a una solución de la ecuación. Entonces $\vec{v} \in C(\vec{t}_j)$ para algún j. Esto implica que

$$\begin{aligned} \left\| \frac{\vec{v}}{\|\vec{v}\|} - \frac{\vec{u}_j}{\|\vec{u}_j\|} \right\| &= \left\| \frac{\vec{v}}{\|\vec{v}\|} - \vec{t}_j + \vec{t}_j - \frac{\vec{u}_j}{\|\vec{u}_j\|} \right\| \\ &\leq \left\| \frac{\vec{v}}{\|\vec{v}\|} - \vec{t}_j \right\| + \left\| \frac{\vec{u}_j}{\|\vec{u}_j\|} - \vec{t}_j \right\| \\ &< \frac{1}{40} + \frac{1}{40} = \frac{1}{20}. \end{aligned}$$

En particular,

$$\left\| \|\vec{v}\|\vec{u}_j\| - \vec{u}_j\|\vec{v}\| \right\| \leq \frac{1}{20}\|\vec{u}_j\|\|\vec{v}\|.$$

Por la Proposición 3.12, y dado que $\|\vec{u}_j\| > c_0$, se tiene que $\|\vec{v}\| \leq 20\|\vec{u}_j\|$.

Sea ahora $M = \max_j \|\vec{u}_j\|$. Entonces $\|\vec{v}\| \leq 20M$ para todo $\vec{v} \in \mathbb{Z}^r$ vector correspondiente a una solución de la ecuación del teorema de las S-unidades. Por el Teorema 3.8 (Northcott), el conjunto de dichos vectores es finito, lo que concluye la prueba. □

Un comentario importante que debemos hacer es que esta demostración no es efectiva en el sentido de que no da una cota al número de soluciones posibles, solo prueba su finitud. Esto se debe a que, a pesar de que c_0 se puede hacer efectivo, la prueba depende de unos $\vec{u}_j$ que sabemos que son finitos, pero no tenemos ningún control sobre su tamaño.

3.3. Demostración del resultado fundamental

Para terminar, el último cabo suelto que nos queda es probar el Teorema 3.10. Consideraremos, para $n \in \mathbb{Z}^+$, los polinomios

$$\begin{aligned} P(x) &= c_n \int_0^1 \big(p(t,x)\big)^n \, dt, \quad p(t,x) = t(1-t)(t-x), \\ Q(x) &= c_n \int_0^1 \big(q(v,x)\big)^n \, dv, \quad q(v,x) = v(1-v)\big(1-(1-x)v\big), \\ R(x) &= c_n \int_0^1 \big(r(u,x)\big)^n \, du, \quad r(u,x) = u(1-u)(xu-1), \end{aligned}$$

donde $c_n = (3n+1)!/(n!)^3$.

Veamos en primer lugar que Q es un polinomio de grado n con coeficientes enteros. Para ello, utilizaremos la *función Beta*

$$B(z_1, z_2) = \int_0^1 s^{z_1-1}(1-s)^{z_2-1} \, ds,$$

para $z_1, z_2 \in \mathbb{C}$ con $\mathrm{Re}(z_1), \mathrm{Re}(z_2) > 0$. Sabemos que está relacionada con la *función Gamma* mediante la expresión

$$B(z_1, z_2) = \frac{\Gamma(z_1)\Gamma(z_2)}{\Gamma(z_1+z_2)},$$

en particular

$$B(m,n) = \frac{(m-1)!(n-1)!}{(m+n-1)!}$$

para $m, n \in \mathbb{Z}^+$. Calculemos explícitamente los coeficientes de Q hallando los valores de $Q^{(k)}(0)/k!$ para $k \in \mathbb{Z}_{\geq 0}$. Derivando tenemos

$$Q^{(k)}(x) = c_n \int_0^1 n(n-1)\cdots(n-k+1)v^{2k}(1-v)^k\big(q(v,x)\big)^{n-k} \, dv,$$

y por tanto,

$$\begin{aligned}
\frac{Q^{(k)}(0)}{k!} &= \frac{c_n}{k!}\int_0^1 \frac{n!}{(n-k)!}v^{n+k}(1-v)^{2n-k}\,dv \\
&= \frac{c_n}{k!}\frac{n!}{(n-k)!}B(n+k+1,2n-k+1) \\
&= \frac{(3n+1)!}{(n!)^3k!}\frac{n!}{(n-k)!}\frac{(n+k)!(2n-k)!}{(3n+1)!} \\
&= \frac{(n+k)!}{n!k!}\frac{(2n-k)!}{n!(n-k)!} \\
&= \binom{n+k}{n}\binom{2n-k}{n} \in \mathbb{Z}.
\end{aligned}$$

En particular, $Q^{(k)}(0)/k! = 0$ si $k > n$, así que $Q \in \mathbb{Z}[x]$ y $\deg Q = n$.

Por una parte, es fácil comprobar que $q(v, 1-x) = -r(v,x)$, de lo que deducimos

$$Q(1-x) = (-1)^n R(x).$$

Por otro lado, también se tiene que $q(v, 1-x^{-1}) = -x^{-1}p(v,x)$, y por tanto,

$$Q(1-x^{-1}) = (-1)^n x^{-n} P(x).$$

Tenemos entonces que

$$P(x) = (-1)^n x^n Q(1-x^{-1}) \quad \text{y} \quad R(x) = (-1)^n Q(1-x),$$

así que $P, R \in \mathbb{Z}[x]$ y $\deg P = \deg R = n$.

Con el cambio $t = xu$ se tiene

$$\begin{aligned}
P &= c_n\int_0^1 \big(p(t,x)\big)^n\,dt = c_n\int_0^1 \big(t(1-t)(t-x)\big)^n\,dt \\
&= c_n\int_0^{x^{-1}} \big(xu(1-xu)(xu-x)\big)^n x\,du \\
&= c_n\int_0^{x^{-1}} x^{2n+1}\big(u(xu-1)(1-u)\big)^n\,du \\
&= x^{2n+1}c_n\int_0^{x^{-1}} \big(r(u,x)\big)^n\,du \\
&= x^{2n+1}c_n\left(\int_0^1 \big(r(u,x)\big)^n\,du - \int_{x^{-1}}^1 \big(r(u,x)\big)^n\,du\right),
\end{aligned}$$

es decir,

$$P = x^{2n+1}R - x^{2n+1}c_n\int_{x^{-1}}^1 \big(r(u,x)\big)^n\,du.$$

Ahora con el cambio $u = (1 - x^{-1})v + x^{-1}$ se tiene

$$\begin{aligned} P &= x^{2n+1}R - x^{2n+1}c_n \int_{x^{-1}}^{1} \big(u(1-u)(xu-1)\big)^n \, du \\ &= x^{2n+1}R - x^{2n+1}c_n \int_{0}^{1} (1-x^{-1})^{2n+1} \big(v(1-v)(1-(1-x)v)\big)^n \, du \\ &= x^{2n+1}R + (1-x)^{2n+1}c_n \int_{0}^{1} \big(q(v,x)\big)^n \, du. \end{aligned}$$

Tenemos entonces la relación

$$P = x^{2n+1}R + (1-x)^{2n+1}Q.$$

En cierto momento necesitaremos acotar de manera uniforme los coeficientes de xR, $(1-x)Q$ y P.

- Por el binomio de Newton, $\binom{n}{k} < (1+1)^n = 2^n$, lo que nos permite acotar los coeficientes de Q:

  $$\binom{n+k}{n}\binom{2n-k}{n} < 2^{n+k}2^{2n-k} = 8^n.$$

  En particular, existe una constante $A_1 > 0$ tal que todos los coeficientes de $(1-x)Q$ están acotados por A_1^n.

- Para xR podemos utilizar que $R(x) = (-1)^n Q(1-x)$, así que existe una constante $A_2 > 0$ tal que todos los coeficientes de xR son menores que A_2^n.

- Por último, de la relación $P = x^{2n+1}R + (1-x)^{2n+1}Q$ que acabamos de deducir, concluimos que también existe una constante $A_3 > 0$ tal que todos los coeficientes de P están acotados por A_3^n.

Tomando el máximo de las tres, concluimos entonces que existe una constante $A > 0$ tal que todos los valores absolutos de los coeficientes de xR, $(1-x)Q$ y P son menores que A^n.

Tan solo necesitamos un resultado más antes de probar el Teorema 3.10. Eventualmente usaremos que si tenemos un sistema compatible determinado 2×2 cuyos coeficientes tengan altura pequeña, entonces su solución también la tiene. Concretamente:

Lema 3.13. *Sean*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}(2, \overline{\mathbb{Q}}) \qquad y \qquad C = \begin{pmatrix} e \\ f \end{pmatrix} \in \overline{\mathbb{Q}}^2.$$

Si $x, y \in \overline{\mathbb{Q}}$ satisfacen

$$A \begin{pmatrix} x \\ y \end{pmatrix} = C$$

entonces

$$h(1:x:y) \leq \log 2 + h(a:b:e) + h(c:d:f)$$

Demostración. Resolviendo para x, y encontramos

$$x = \frac{bf-de}{ad-bc}, \qquad y = \frac{af-ce}{ad-bc}.$$

Entonces por definición de altura tenemos

$$\begin{aligned} h(1:x:y) &= h(ad-bc : bf-de : af-ce) \\ &= \sum_v \log \sup \left\{ |ad-bc|_v, |bf-de|_v, |af-ce|_v \right\}. \end{aligned}$$

Observamos que $|ad-bc|_v \leq \sup\{1,|2|_v\} \sup\{|a|_v|d|_v, |b|_v|c|_v\}$, y ocurre algo similar para los otros términos. Por tanto obtenemos

$$\begin{aligned} h(1:x:y) &\leq h(2) + \sum_v \log \sup \{|a|_v, |b|_v, |e|_v\} \\ &\quad + \sum_v \log \sup \{|c|_v, |d|_v, |f|_v\} \\ &= \log 2 + h(a:b:e) + h(c:d:f), \end{aligned}$$

lo que concluye la prueba. □

Con esto, ya tenemos todos los ingredientes para demostrar el resultado fundamental.

Demostración del Teorema 3.10. Sean $(x_1,y_1),(x_2,y_2)$ las soluciones de la ecuación $x+y=1$ que dan lugar a $\vec{u},\vec{v}$ respectivamente. El enunciado que queremos probar es entonces equivalente a

$$h(x_1)+h(y_1) \leq \frac{1}{4}c_0 + \frac{2}{n-1}\left(h(x_2x_1^{-2n}) + h(y_2y_1^{-2n})\right).$$

Consideramos ahora el sistema lineal que, en la notación del Lema 3.13, viene dado por

$$A = \begin{pmatrix} x_2x_1^{-2n} & y_2y_1^{-2n} \\ x_1R(x_1) & y_1Q(x_1) \end{pmatrix}, \qquad C = \begin{pmatrix} 1 \\ P(x_1) \end{pmatrix}.$$

Es inmediato, utilizando la relación de P, Q y R, que $x = x_1^{2n}, y = y_1^{2n}$ es solución del sistema. Concretamente:

$$\begin{cases} x_2x_1^{-2n}x_1^{2n} + y_2y_1^{-2n}y_1^{2n} = x_2 + y_2 = 1 \\ x_1R(x_1)x_1^{2n} + y_1Q(x_1)y_1^{2n} = x_1^{2n+1}R(x_1) + (1-x_1)^{2n+1}Q(x_1) = P(x_1) \end{cases}$$

Aplicando el Lema 3.13 y la Proposición 3.2 obtenemos

$$\begin{aligned} h(x_1^{2n}) &\leq h(1 : x_1^{2n} : y_1^{2n}) \\ &\leq \log 2 + h(x_2 x_1^{-2n} : y_2 y_1^{-2n} : 1) + h(x_1 R(x_1) : y_1 Q(x_1) : P(x_1)) \\ &\leq \log 2 + h(x_2 x_1^{-2n}) + h(y_2 y_1^{-2n}) + h(x_1 R(x_1) : y_1 Q(x_1) : P(x_1)). \end{aligned}$$

Queremos acotar el último sumando. Observamos ahora que los polinomios $xR(x), (1-x)Q(x), P(x)$ tienen grado menor o igual que $n+1$, y que por lo dicho anteriormente todos sus coeficientes están acotados por A^n con A una constante positiva. El Lema 3.5 nos da

$$h(x_1 R(x_1) : (1-x_1)Q(x_1) : P(x_1)) \leq (n+1)h(x_1) + \log(n+2) + \log A^n,$$

y como $\log(n+2) \leq n$ para $n \geq 2$, se tiene

$$h(x_1 R(x_1) : (1-x_1)Q(x_1) : P(x_1)) \leq (n+1)h(x_1) + Bn$$

para $B = 1 + \log A$. Finalmente, la cota nos queda

$$2nh(x_1) \leq \log 2 + h(x_2 x_1^{-2n}) + h(y_2 y_1^{-2n}) + (n+1)h(x_1) + Bn,$$

o, equivalentemente,

$$(n-1)h(x_1) \leq \log 2 + h(x_2 x_1^{-2n}) + h(y_2 y_1^{-2n}) + Bn.$$

Utilizando la simetría de la solución de la ecuación $x+y=1$, podemos hacer el cambio $(x_1, y_1) \leftrightarrow (y_1, x_1)$, $(x_2, y_2) \leftrightarrow (y_2, x_2)$, lo que nos da la cota

$$(n-1)h(y_1) \leq \log 2 + h(y_2 y_1^{-2n}) + h(x_2 x_1^{-2n}) + Bn.$$

Sumando ambas cotas obtenemos

$$(n-1)\left(h(x_1) + h(y_1)\right) \leq 2\log 2 + 2Bn + 2\left(h(y_2 y_1^{-2n}) + h(x_2 x_1^{-2n})\right),$$

es decir

$$h(x_1) + h(y_1) \leq \frac{2\log 2}{n-1} + \frac{2nB}{n-1} + \frac{2}{n-1}\left(h(y_2 y_1^{-2n}) + h(x_2 x_1^{-2n})\right).$$

Como $n \geq 2$, podemos acotar

$$\frac{2\log 2}{n-1} \leq 2\log 2, \qquad \frac{2nB}{n-1} \leq 4B,$$

y tomando $c_0 = 8\log 2 + 16B$ tenemos

$$h(x_1) + h(y_1) \leq \frac{1}{4}c_0 + \frac{2}{n-1}\left(h(x_2 x_1^{-2n}) + h(y_2 y_1^{-2n})\right),$$

lo que concluye la prueba. □

Como comentario final, en cierto momento de la demostración hemos asumido que el sistema era compatible determinado para poder aplicar el Lema 3.13. En realidad, un sistema elegido al azar es compatible determinado con probabilidad 1, pero en caso de que no ocurra hay formas de solucionarlo. La forma más fácil de solventarlo es derivar en la expresión que relaciona a P, Q y R, obteniendo así

$$P' = x^{2n}(xR' + (2n+1)R) + (1-x)^{2n}\left((1-x)Q' - (2n+1)Q\right).$$

Entonces reemplazaríamos los polinomios xR, $(1-x)Q$ y P de la prueba por $xR' + (2n+1)R$, $(1-x)Q' - (2n+1)Q$ y P' respectivamente. Para más detalles se puede consultar [17, p.188].

Bibliografía

[1] Beukers, F., & Schlickewei, H. (1996). *The equation $x + y = 1$ in finitely generated groups*. Acta Arithmetica, 78(2), 189-199.

[2] Beukers, F., & Tijdeman, R. (1984). *On the multiplicities of binary complex recurrences*. Compositio Mathematica, 51(2), 193-213.

[3] Bombieri, E., & Gubler, W. (2006). *Heights in Diophantine geometry* (No. 4). Cambridge university press.

[4] Chamizo. F. (2022). *Teoría combinatoria y analítica de números curso* 2022/23 (Sec. 3.3). Universidad Autónoma de Madrid.

[5] Cohen, H. (2013). *A course in computational algebraic number theory* (Vol. 138). Springer Science.

[6] Cohen, H., Axler, S., & Ribet, K. A. (2007). *Number theory: Volume I: Tools and diophantine equations* (Vol. 560). Springer New York.

[7] Conrad, K. (2020). *Dirichlet's unit theorem*. Expository papers on Algebraic Number Theory.

[8] Conrad, K. *Ideal factorization*. Expository papers on Algebraic Number Theory.

[9] Conrad, K. (2010). *Ostrowski for number fields*. Expository papers on Algebraic Number Theory.

[10] Conrad, K. (2023). *Totally ramified primes and Eisenstein polynomials*. Expository papers on Algebraic Number Theory.

[11] Hardy, G. H., & Wright, E. M. (1979). *An introduction to the theory of numbers*. Oxford university press.

[12] Lang, S. (1994). *Algebraic number theory* (Vol. 110). Springer Science.

[13] Narkiewicz, W. (1974). *Elementary and analytic theory of algebraic numbers* (Vol. 57). Warszawa: Pwn.

[14] Neukirch, J. (2013). *Algebraic number theory* (Vol. 322). Springer Science.

[15] Ono, T. (2012). *An introduction to algebraic number theory*. Springer Science.

[16] Stewart, I., & Tall, D. (2001). *Algebraic number theory and Fermat's last theorem*. AK Peters/CRC Press.

[17] Zannier, U. (2009). *Lecture notes on Diophantine analysis* (Vol. 8). Edizioni della Normale, Pisa.