DIRICHLET'S UNIT THEOREM

KEITH CONRAD

1. Introduction

Dirichlet's unit theorem describes the structure of the unit group of orders in a number field.

Theorem 1.1 (Dirichlet, 1846)**.** *Let K be a number field with r_1 real embeddings and $2r_2$ pairs of complex conjugate embeddings. The unit group of an order in K is finitely generated with $r_1 + r_2 - 1$ independent generators of infinite order.*

More precisely, letting $r = r_1 + r_2 - 1$, each order $\mathcal{O}$ in K contains multiplicatively independent units $\varepsilon_1, \ldots, \varepsilon_r$ of infinite order such that every unit in $\mathcal{O}$ can be written uniquely in the form

$$\zeta \varepsilon_1^{m_1} \cdots \varepsilon_r^{m_r},$$

where ζ is a root of unity in $\mathcal{O}$ and the m_i's are in $\mathbf{Z}$. Abstractly, $\mathcal{O}^\times \cong \mu(\mathcal{O}) \times \mathbf{Z}^{r_1+r_2-1}$, where $\mu(\mathcal{O})$ is the finite cyclic group of roots of unity in $\mathcal{O}$.

Units $u_1, \ldots, u_k$ are called *multiplicatively independent*, or just independent, when they satisfy no multiplicative relations except the trivial one: $u_1^{m_1} \cdots u_k^{m_k} = 1 \Rightarrow m_i = 0$ for all i. It then follows that exponents in such a product are unique: if $u_1^{m_1} \cdots u_k^{m_k} = u_1^{n_1} \cdots u_k^{n_k}$ then $m_i = n_i$ for all i. This looks like linear independence, and that is exactly what it is: when we view $\mathcal{O}^\times$ as a $\mathbf{Z}$-module using its group law, multiplicative independence means $\mathbf{Z}$-linear independence.

If $r_1 > 0$ then $\mu(\mathcal{O}) = \{\pm 1\}$ since ± 1 are the only roots of unity in $\mathbf{R}$. If $r_1 = 0$ we might also have $\mu(\mathcal{O}) = \{\pm 1\}$, *e.g.*, $\mathcal{O} = \mathbf{Z}[\sqrt{d}]$ for $d < -1$.

It is important that the unit groups of all orders in K have the same number of independent generators of infinite order: $r_1 + r_2 - 1$. Therefore $[\mathcal{O}_K^\times : \mathcal{O}^\times]$ is finite. A choice of generators $\varepsilon_1, \ldots, \varepsilon_r$ for $\mathcal{O}^\times$ (really, for the quotient group $\mathcal{O}^\times/\mu(\mathcal{O})$) is called a system of *fundamental units*. We call $r_1 + r_2 - 1$ the rank of the unit group.

The unit groups of orders in number fields were, historically, the first important examples of finitely generated abelian groups. Finding algorithms to produce explicit generators for unit groups is one of the tasks of computational number theory.

In Section 2 we will look at some examples of the unit theorem. The theorem will be proved in Section 3 and some more examples are described in Sections 4 and 5.

2. Examples

Example 2.1. For $\mathbf{Q}(\sqrt{2})$ we have $r_1 + r_2 - 1 = 1$, so the unit group of each order in $\mathbf{Q}(\sqrt{2})$ has the form $\pm\varepsilon^{\mathbf{Z}}$ for some unit ε. In particular, $\mathbf{Z}[\sqrt{2}]^\times = \pm(1+\sqrt{2})^{\mathbf{Z}}$ and $\mathbf{Z}[3\sqrt{2}]^\times = \pm(17+12\sqrt{2})^{\mathbf{Z}}$.

Table 1 describes unit groups in the full ring of integers in several number fields. The unit ε in the row for $\mathbf{Q}(\sqrt[3]{2}, \zeta_3)$ is

$$\varepsilon = \frac{-1 + 2\sqrt[3]{2} + \sqrt[3]{4}}{3} + \frac{1 - \sqrt[3]{2} + \sqrt[3]{4}}{3}\zeta_3.$$

K	r_1	r_2	$r_1 + r_2 - 1$	$\mu(\mathcal{O}_K)$	$\mathcal{O}_K^\times$
$\mathbf{Q}(\sqrt{3})$	2	0	1	± 1	$\pm(2+\sqrt{3})^{\mathbf{Z}}$
$\mathbf{Q}(\sqrt{5})$	2	0	1	± 1	$\pm(\frac{1+\sqrt{5}}{2})^{\mathbf{Z}}$
$\mathbf{Q}(\zeta_5)$	0	2	1	μ_{10}	$\mu_{10}(\frac{1+\sqrt{5}}{2})^{\mathbf{Z}}$
$\mathbf{Q}(\sqrt[3]{2})$	1	1	1	± 1	$\pm(1+\sqrt[3]{2}+\sqrt[3]{4})^{\mathbf{Z}}$
$\mathbf{Q}(\sqrt[3]{6})$	1	1	1	± 1	$\pm(1-6\sqrt[3]{6}+3\sqrt[3]{36})^{\mathbf{Z}}$
$\mathbf{Q}(\sqrt[4]{2})$	2	1	2	± 1	$\pm(1+\sqrt[4]{2})^{\mathbf{Z}}(1+\sqrt{2})^{\mathbf{Z}}$
$\mathbf{Q}(\sqrt[3]{2}, \zeta_3)$	0	3	2	μ_6	$\mu_6 \cdot \varepsilon^{\mathbf{Z}}\overline{\varepsilon}^{\mathbf{Z}}$
$\mathbf{Q}(\sqrt{2}, \sqrt{3})$	4	0	3	± 1	$\pm(1+\sqrt{2})^{\mathbf{Z}}(\sqrt{2}+\sqrt{3})^{\mathbf{Z}}(\frac{\sqrt{2}+\sqrt{6}}{2})^{\mathbf{Z}}$

TABLE 1. Unit Group of $\mathcal{O}_K$

Example 2.2. The unit group of an order is finite if and only if $r_1 + r_2 - 1 = 0$. This means (r_1, r_2) is $(1, 0)$ or $(0, 1)$, so K is $\mathbf{Q}$ or an imaginary quadratic field. Moreover, the unit group of each order in an imaginary quadratic field is $\{\pm 1\}$ except for the maximal orders $\mathbf{Z}[i]$ and $\mathbf{Z}[\zeta_3]$, whose units groups have size 4 and 6, respectively. There are a number of important results in algebraic number theory that have a simpler form for $\mathbf{Q}$ and imaginary quadratic fields than for other number fields, precisely because in these (and only these) cases the unit group is finite.

Example 2.3. We have $r_1 + r_2 - 1 = 1$ if and only if $(r_1, r_2) = (2, 0)$, $(1, 1)$, or $(0, 2)$, *i.e.*, K is real quadratic (*e.g.*, $\mathbf{Q}(\sqrt{2})$), a cubic field with only one real embedding (*e.g.*, $\mathbf{Q}(\sqrt[3]{2})$), or a totally complex quartic field (*e.g.*, $\mathbf{Q}(\zeta_5)$).

Example 2.4. If K is a totally real cubic field then $r_1 + r_2 - 1 = 2$, so each order in K has unit group of the form $\pm\varepsilon_1^{\mathbf{Z}}\varepsilon_2^{\mathbf{Z}}$.

Example 2.5. We always have $r_1 + r_2 - 1 \leq n - 1$, where $n = [K : \mathbf{Q}] = r_1 + 2r_2$. Easily $r_1 + r_2 - 1 = n - 1$ if and only if $r_2 = 0$, *i.e.*, K is a totally real number field.

Example 2.6. Let's look at a unit group with rank greater than 1 and see how to find multiplicative relations between units numerically, by using logarithms to discover them as linear relations. Set $K = \mathbf{Q}(\alpha)$ where $\alpha^3 - 3\alpha - 1 = 0$. The polynomial $f(T) = T^3 - 3T - 1$ has 3 real roots, so $\mathcal{O}_K^\times$ has rank $r_1 + r_2 - 1 = 3 - 1 = 2$.

Before looking at $\mathcal{O}_K^\times$, let's show $\mathcal{O}_K = \mathbf{Z}[\alpha]$. Since $\mathrm{disc}(\mathbf{Z}[\alpha]) = -4(-3)^3 - 27(-1)^2 = 81 = 3^4$, $[\mathcal{O}_K : \mathbf{Z}[\alpha]]$ divides 9. Therefore elements of $\mathcal{O}_K$ when written in the basis $\{1, \alpha, \alpha^2\}$ have coefficients with denominator dividing 9. Since $f(T+1) = T^3 + 3T^2 - 3$ is Eisenstein at 3 with $\alpha - 1$ as a root, elements of $\mathcal{O}_K$ when written in the basis $\{1, \alpha - 1, (\alpha - 1)^2\}$ have coefficients with denominator prime to 3. This carries over to $\{1, \alpha, \alpha^2\}$, so $\mathcal{O}_K = \mathbf{Z}[\alpha]$. (The Minkowski bound is exactly 2, and there is no prime ideal with norm 2 since $T^3 - 3T - 1$ is irreducible modulo 2, so $h(K) = 1$: $\mathbf{Z}[\alpha]$ is a PID.)

We now write down several units in $\mathbf{Z}[\alpha]$. For $a, b \in \mathbf{Q}$, $\mathrm{N}_{K/\mathbf{Q}}(a\alpha + b) = -a^3 f(-b/a)$. Check with this formula that $\alpha, \alpha+1, \alpha-2$, and $2\alpha+3$ all have norm ± 1, so they are all in $\mathbf{Z}[\alpha]^\times$. The three roots of $f(T)$ are α, $2-\alpha^2$, and $\alpha^2-\alpha-2$, so $2-\alpha^2$ and $\alpha^2-\alpha-2$ are in $\mathbf{Z}[\alpha]^\times$. The product of all three roots of $f(T)$ is $-f(0) = 1$.

Since $\mathbf{Z}[\alpha]^\times$ has rank 2, the nontrivial units we just wrote down must admit some nontrivial multiplicative relations. How can we find such relations? We will use the three different embeddings $K \to \mathbf{R}$. Call them σ_1, σ_2, and σ_3. The real roots of $f(T)$ are $\sigma_1(\alpha), \sigma_2(\alpha)$, and $\sigma_3(\alpha)$. Arranging the roots in increasing order,

$$\sigma_1(\alpha) = -1.532\ldots, \quad \sigma_2(\alpha) = -.347\ldots, \quad \sigma_3(\alpha) = 1.879\ldots.$$

For $\gamma \in K$, $\mathrm{N}_{K/\mathbf{Q}}(\gamma) = \sigma_1(\gamma)\sigma_2(\gamma)\sigma_3(\gamma)$. For $u \in \mathcal{O}_K^\times$, $|\sigma_1(u)\sigma_2(u)\sigma_3(u)| = 1$. Taking logarithms,

$$u \in \mathcal{O}_K^\times \Longrightarrow \log|\sigma_1(u)| + \log|\sigma_2(u)| + \log|\sigma_3(u)| = 0. \tag{2.1}$$

Define the logarithmic mapping $L\colon K^\times \to \mathbf{R}^3$ by

$$L(\gamma) = (\log|\sigma_1(\gamma)|, \log|\sigma_2(\gamma)|, \log|\sigma_3(\gamma)|).$$

We will use such a map L in the proof of the general unit theorem; here we will see how this map is useful computationally. Easily L is a group homomorphism and by (2.1), $L(\mathcal{O}_K^\times)$ is in the hyperplane $\{(x,y,z) \in \mathbf{R}^3 : x+y+z = 0\}$. The kernel of L on $\mathcal{O}_K^\times$ is $\{\pm 1\}$ (why?). Table 2 gives numerical approximations to the images of units under the logarithmic mapping.

γ	$L(\gamma)$ (approx.)
α	$(.4266, -1.0575, .6309)$
$\alpha+1$	$(-.6309, -.4266, 1.0575)$
$\alpha-2$	$(1.2618, .8532, -2.1151)$
$2\alpha+3$	$(-2.7460, .8352, 1.9108)$
$2-\alpha^2$	$(-1.0575, .6309, .4266)$
$\alpha^2-\alpha-2$	$(.6309, .4266, -1.0575)$

TABLE 2. Log Images of Units

From the table, it appears that $L(\alpha-2) = -2L(\alpha+1) = L(1/(\alpha+1)^2)$, so $\alpha - 2 = \pm 1/(\alpha+1)^2$. You can check that the minus sign is needed. Using a computer algebra package, the 3×3 matrix $\begin{pmatrix} L(\alpha) & L(\alpha+1) & L(2\alpha+3) \end{pmatrix}$ has $(2,-3,1)$ in its kernel, so $\alpha^2(\alpha+1)^{-3}(2\alpha+3)$ has L-value 0. Therefore $2\alpha+3 = \pm\alpha^{-2}(\alpha+1)^3$. Check that the plus sign holds. Since it looks like $L(2-\alpha^2) = L(\alpha+1) - L(\alpha)$, $2-\alpha^2 = \pm(\alpha+1)/\alpha$ and the minus sign is needed. Then, since the three roots of $f(T)$ multiply to 1, $\alpha^2-\alpha-2 = 1/(\alpha(2-\alpha^2)) = (1/\alpha)(-\alpha/(\alpha+1)) = -1/(\alpha+1)$.

This evidence suggests that α and $\alpha+1$ are a system of fundamental units for $\mathbf{Z}[\alpha]^\times$.

3. Proof of the unit theorem

Our proof of the unit theorem is based on [3, Sect. 1.5] and [4, pp. 214–215] (see also [5, p. 5]), and is deduced from a compactness theorem: the unit theorem is a consequence of a certain group being compact.

We will use Minkowski's convex-body theorem in our proof. This is a standard tool for proofs of the unit theorem, although by comparison with typical applications of Minkowski's

theorem we will be able to get by with a crudely chosen convex body: a sufficiently large ball will work.

Dirichlet did not use Minkowski's theorem; he proved the unit theorem in 1846 while Minkowski's theorem appeared in 1889. Dirichlet's substitute for the convex-body theorem was the pigeonhole principle. (An account of Dirichlet's proof in German is in [2, Sect. 183] and in English is in [6, Sect. 2.8–2.10].) Dirichlet did not state the unit theorem for all orders, but only those of the form $\mathbf{Z}[\alpha]$, since at the time these were the kinds of rings that were considered. According to an oft-repeated story, the main idea for the proof of the unit theorem came to Dirichlet while he attended a concert in the Sistine Chapel.[1]

We set some notation. As in the statement of the unit theorem, K is a number field of degree n, r_1 is the number of real embeddings of K and $2r_2$ is the number of complex embeddings of K (that is, embeddings $K \to \mathbf{C}$ whose image is not in $\mathbf{R}$), so $n = r_1 + 2r_2$. Set $V = \mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$, so $\dim_{\mathbf{R}}(V) = n$. The *Euclidean embedding* $\theta_K \colon K \to V$ is defined using the real and complex embeddings of K, as follows. Let the real embeddings of K be $\sigma_1, \ldots, \sigma_{r_1}$ and let the complex embeddings of K be $\sigma_{r_1+1}, \overline{\sigma}_{r_1+1}, \ldots, \sigma_{r_1+r_2}, \overline{\sigma}_{r_1+r_2}$, where we collect the complex embeddings into conjugate pairs $\sigma_j, \overline{\sigma}_j$. For $\alpha \in K$, we set[2]

$$\theta_K(\alpha) = (\sigma_1(\alpha), \cdots, \sigma_{r_1}(\alpha), \sigma_{r_1+1}(\alpha), \cdots, \sigma_{r_1+r_2}(\alpha)) \in V.$$

Algebraically, V is a commutative ring using component wise operations. Give V its natural topology as a Euclidean space and *all subsets of V will be given the subspace topology.* A particular subset we will care about is $V^\times = (\mathbf{R}^\times)^{r_1} \times (\mathbf{C}^\times)^{r_2}$.

Let $\mathrm{N} \colon V \to \mathbf{R}$ by

$$\mathrm{N}(x_1, \ldots, x_{r_1}, z_1, \ldots, z_{r_2}) = x_1 \cdots x_{r_1} |z_1|^2 \cdots |z_{r_2}|^2 = x_1 \cdots x_{r_1} z_1 \overline{z}_1 \cdots z_{r_2} \overline{z}_{r_2}.$$

On the image of K in V, N looks like the norm: $\mathrm{N}(\theta_K(\alpha)) = \mathrm{N}_{K/\mathbf{Q}}(\alpha)$ for all $\alpha \in K$. Set

$$G = \{v \in V^\times : |\mathrm{N}(v)| = 1\}.$$

This is a subgroup of $V^\times$, and it is closed in V since G is the inverse image of $\{1\}$ under the continuous map $V \to \mathbf{R}$ given by $v \mapsto |\mathrm{N}(v)|$. Thus G is a closed subgroup of $V^\times$.

Let $\mathcal{O}$ be an order in K and set

$$U = \theta_K(\mathcal{O}^\times) = G \cap \theta_K(\mathcal{O}).$$

(Think "U = units".) We have $U \subset G$ since $\mathcal{O}^\times = \{\alpha \in \mathcal{O} : |\mathrm{N}_{K/\mathbf{Q}}(\alpha)| = 1\}$. Since we give G the subspace topology from V and the image of $\mathcal{O}$ in V under the Euclidean embedding is discrete, U is discrete in G. We will be interested in the quotient group G/U.

[1] For instance, in 1905 Minkowski [9, pp. 156–7] wrote "Es wird erzählt, da nach langjährigen vergeblichen Bemühungen um das schwierige Problem Dirichlet die Lösung in Rom in der Sixtinischen Kapelle während des Anhörens der Ostermusik ergründet hat. Inwieweit dieses Faktum für die von manchen behauptete Wahlverwandtschaft zwischen Mathematik und Musik spricht, wage ich nicht zu erörtern." (*translation*: "People say that, after many years of unsuccessful efforts in trying to solve this difficult problem, Dirichlet found the solution in Rome in the Sistine Chapel while listening to Easter music. I do not dare to discuss to what extent this fact confirms the conjectured (by some people) relationship between mathematics and music.")

[2] The Euclidean embedding of K, as defined here, depends on the ordering of the different real and complex embeddings as well as on the choice of one complex embedding from each conjugate pair. A coordinate-free way of defining the Euclidean embedding uses tensor products: the natural mapping $K \to \mathbf{R} \otimes_{\mathbf{Q}} K$ where $\alpha \mapsto 1 \otimes \alpha$ is a ring homomorphism into a finite-dimensional real vector space of dimension n, just like V.

Example 3.1. Let $K = \mathbf{Q}(\sqrt{2})$ and $\mathcal{O} = \mathcal{O}_K = \mathbf{Z}[\sqrt{2}]$. Then $V = \mathbf{R}^2$ and $\mathrm{N}\colon V \to \mathbf{R}$ by $\mathrm{N}(x, y) = xy$. The Euclidean embedding $\theta\colon \mathbf{Q}(\sqrt{2}) \to \mathbf{R}^2$ places $\mathbf{Z}[\sqrt{2}]^\times$ on the curve $G = \{(x, y) \in \mathbf{R}^2 : |xy| = 1\}$, a union of two hyperbolas. We know $\mathbf{Z}[\sqrt{2}]^\times = \pm(1+\sqrt{2})^{\mathbf{Z}}$ and $U = \theta_K(\mathbf{Z}[\sqrt{2}]^\times)$ is a discrete subset of G ("equally spaced" in a multiplicative sense). See Figure 1.

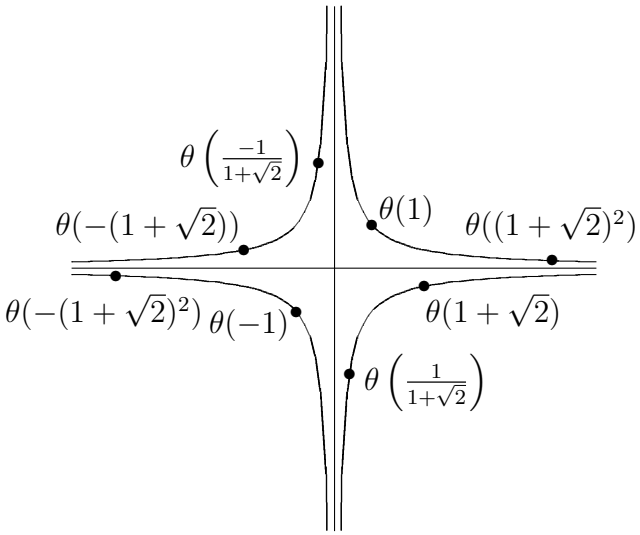


FIGURE 1. Units in $\mathbf{Z}[\sqrt{2}]$ on $G = \{(x, y) \in \mathbf{R}^2 : |xy| = 1\}$.

Let's see how the subgroup U moves G around by multiplication in Figure 1. Multiplying G by some $u \in U$ moves the arcs between consecutive points of U in Figure 1 to other arcs between consecutive point, and it exchanges the hyperbolas $y = 1/x$ and $y = -1/x$ if $\mathrm{N}(u) = -1$. Multiplication by $\theta(-1) = (-1, -1)$ on G exchanges the two branches on each hyperbola.

Modulo U each $(x, y) \in G$ is congruent to a point on the arc between $\theta(1)$ and $\theta((1+\sqrt{2})^2)$, so the map $[1, (1+\sqrt{2})^2] \to G/U$ given by $x \mapsto (x, 1/x)U$ is surjective and continuous, which implies G/U is compact.

In Example 3.1 we used knowledge of the unit group of $\mathbf{Z}[\sqrt{2}]$ to see G/U is compact. The key to proving the unit theorem is showing the compactness of G/U without knowing the structure of the unit group in advance.

Lemma 3.2. *For nonzero a in $\mathcal{O}$, $[\mathcal{O} : (a)] = |\mathrm{N}_{K/\mathbf{Q}}(a)|$.*

Proof. This follows from $\mathcal{O}$ being a free $\mathbf{Z}$-module of rank $[K : \mathbf{Q}]$. □

Lemma 3.3. *For each positive integer N, finitely many $a \in \mathcal{O}$ satisfy $|\mathrm{N}_{K/\mathbf{Q}}(a)| = N$ up to multiplication by $\mathcal{O}^\times$. That is, there are $a_1, \ldots, a_k \in \mathcal{O}$, where k depends on N, such that $|\mathrm{N}_{K/\mathbf{Q}}(a_i)| = N$ and each $a \in \mathcal{O}$ satisfying $|\mathrm{N}_{K/\mathbf{Q}}(a)| = N$ is a unit multiple of an a_i.*

Proof. If $|\mathrm{N}_{K/\mathbf{Q}}(a)| = N$ then $[\mathcal{O} : (a)] = N$ by Lemma 3.2, so $N\mathcal{O} \subset (a) \subset \mathcal{O}$. Since $\mathcal{O}/N\mathcal{O}$ is finite, there are only finitely many principal ideals between $N\mathcal{O}$ and $\mathcal{O}$. Let $(a_1), \ldots, (a_k)$ be those ideals. Then $(a) = (a_i)$ for some i, so a and a_i are unit multiples. □

Theorem 3.4. *The group G/U is compact in the quotient topology.*

Proof. We will find a compact subset S of G that represents all cosets in G/U. The continuous map $S \to G/U$ is onto and thus G/U is compact. (Usually G itself is not compact. See Figure 1.)

We begin with a remark about volumes. For $v \in V^\times$, multiplication of $V = \mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$ by v is an $\mathbf{R}$-linear map (hence continuous) given by a matrix with determinant $\mathrm{N}(v)$, so for a region $R \subset V$ with finite volume, the volume of vR is $|\mathrm{N}(v)|$ times the volume of R. In particular, if $v \in G$ then $\mathrm{vol}(vR) = \mathrm{vol}(R)$ because $|\mathrm{N}(v)| = 1$. When R is compact so is vR, by continuity of multiplication.

Pick a compact, convex, centrally symmetric region $C \subset V$ with $\mathrm{vol}(C) > 2^n \,\mathrm{vol}(\theta_K(\mathcal{O}))$, where the "volume" of the lattice $\theta_K(\mathcal{O})$ means the volume of a fundamental domain for this lattice as a subset of V. For instance, C could be a large ball in V centered at the origin. For *each* $g \in G$, gC is also compact and centrally symmetric. It is convex too, since multiplication by g on V is an invertible linear transformation, and invertible linear transformations send convex sets to convex sets. Using g^{-1} instead of g, Minkowski's convex body theorem applies to $g^{-1}C$ and the lattice $\mathcal{O} \subset V$ (we identify $\mathcal{O}$ with $\theta_K(\mathcal{O})$):

$$g^{-1}C \cap (\mathcal{O} - \{0\}) \neq \emptyset.$$

Let a be a nonzero element of $\mathcal{O}$ lying in $g^{-1}C$. Then $|\mathrm{N}_{K/\mathbf{Q}}(a)| = |\mathrm{N}(a)| \in |\mathrm{N}(g^{-1}C)| = |\mathrm{N}(C)|$, which is a bounded subset of $\mathbf{R}$ since C is compact. Note $|\mathrm{N}(C)|$ is independent of g. The number $|\mathrm{N}_{K/\mathbf{Q}}(a)|$ is also an integer, so $|\mathrm{N}_{K/\mathbf{Q}}(a)|$ lies in a *finite set* (a bounded set of integers is finite). Combining that with Lemma 3.3, there is a finite set $\{a_1, \dots, a_m\}$ of nonzero elements of $\mathcal{O}$ such that every $g^{-1}C$ meets some $a_i\mathcal{O}^\times = a_iU$, which implies every gU meets some $a_i^{-1}C$.

We have shown the quotient group G/U is represented by $G \cap \bigcup_{i=1}^m a_i^{-1}C$. The union $\bigcup_{i=1}^m a_i^{-1}C$ is a compact subset of V, since each $a_i^{-1}C$ is compact, and since G is closed in V the intersection $G \cap \bigcup_{i=1}^m a_i^{-1}C$ is compact in G. Hence G/U has a compact set of representatives in G, so G/U is compact in the quotient topology. □

Now we prove the unit theorem. Recall that, by definition, $G = \{v \in V : |\mathrm{N}(v)| = 1\}$. Each element of $V = \mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$ can be written in the form $(x_1, \dots, x_{r_1}, z_{r_1+1}, \dots, z_{r_1+r_2})$. Define the logarithmic mapping $L\colon V^\times \to \mathbf{R}^{r_1+r_2}$ by

$$L(x_1, \dots, z_{r_1+r_2}) := (\dots, \log|x_i|, \dots, 2\log|z_j|, \dots),$$

where the coefficients 2 in this formula are related to the exponents 2 in the definition of N. The function L is a continuous group homomorphism and, for each $g \in G$, $L(g)$ lies in the hyperplane

$$H = \{(y_1, \dots, y_{r_1+r_2}) \in \mathbf{R}^{r_1+r_2} : \sum_i y_i = 0\}.$$

It is easy to see that $L(G) = H$, so $L(G)$ has dimension $r_1 + r_2 - 1$ over $\mathbf{R}$. What we really care about is $L(U)$, which provides a linearized geometric picture for U (once we determine the kernel of $L|_U$). The basic plan is to show $L(U)$ is a "full" lattice in the hyperplane $L(G)$ and the kernel of L restricted to U is finite cyclic (coming from roots of unity in U).

First we treat the **kernel** of $L|_U$. As a map out on $V^\times$, L has compact kernel:

$$\ker L = \{\pm 1\}^{r_1} \times (S^1)^{r_2}.$$

Every root of unity in U gets sent to 0 by L. Let's check these are the only elements of $U = \mathcal{O}^\times$ in $\ker L$. Since U is closed in $V^\times$ (all discrete subsets are closed), the kernel of

$L|_U$ is closed and thus (as a subset of $\{\pm 1\}^{r_1} \times (S^1)^{r_2}$) is compact. Since $\mathcal{O}$ is discrete in V (it's a lattice), U is discrete in $V^\times$, so the kernel of $L|_U$ is also discrete (a subset of a discrete set is discrete), so $\ker(L|_U)$ is compact and discrete: it is finite! A subgroup of U with finite order can only contain roots of unity. Therefore the elements of $\ker(L|_U)$ are the roots of unity in $U = \mathcal{O}^\times$, which form a finite cyclic group since every finite subgroup of $K^\times$ is a cyclic group. (Warning: it is *false* that the kernel of L as a map out of $K^\times$ is only the roots of unity in K. An element of $K^\times$ that has all of its $\mathbf{Q}$-conjugates lying on the unit circle must be in the kernel of L. An example is $3/5 + (4/5)i$, or more generally $a/c + (b/c)i$ where (a, b, c) is a Pythagorean triple. But these are not algebraic *integers*, so they don't belong to U.)

Now we look at the **image** $L(U)$ in the hyperplane $L(G) \subset \mathbf{R}^{r_1+r_2}$. We have already seen (and used) that the group U is discrete in $V^\times$, so also in G. The image of a discrete set under a continuous map need not be discrete (consider $\mathbf{Z}^2 \to \mathbf{R}$ by $(m, n) \mapsto m + n\sqrt{2}$), but $L(U)$ is discrete in $L(G)$ since there are only finitely many elements in $L(U)$ that lie in a bounded region of $\mathbf{R}^{r_1+r_2}$. Indeed, consider the box

$$\{(y_1, \ldots, y_{r_1+r_2}) \in \mathbf{R}^{r_1+r_2} : |y_i| \leq b\}.$$

Suppose $L(u)$ is in this box for some $u \in U$. The real embeddings[3] of u have absolute value at most e^b and the complex embeddings of u have absolute value at most $e^{b/2}$. That puts an upper bound in terms of b (and $n = [K : \mathbf{Q}]$) on the coefficients of the polynomial $\prod_\sigma (T - \sigma(u)) \in \mathbf{Z}[T]$. The coefficients have only finitely many possibilities, since there are finitely many integers with absolute value below a given bound, so there are finitely many such polynomials. As u is a root of such a polynomial, there are finitely many choices for u. This shows $L(U)$ is discrete.

Since $L(U)$ is a discrete subgroup of $L(G) \cong \mathbf{R}^{r_1+r_2-1}$, $L(U) \cong \mathbf{Z}^{r'}$ where $r' \leq r_1 + r_2 - 1$. Since $L \colon G \to L(G)$ is a continuous and surjective group homomorphism, the induced map $G/U \to L(G)/L(U)$ is also continuous and surjective where both quotient groups get the quotient topology. From Theorem 3.4, G/U is compact so $L(G)/L(U)$ is compact. Since $L(G)$ is (r_1+r_2-1)-dimensional over $\mathbf{R}$ and $L(U)$ has $\mathbf{Z}$-rank $r' \leq r_1+r_2-1$, compactness of $L(G)/L(U)$ forces $r' = r_1 + r_2 - 1$: Euclidean space modulo a discrete subgroup is compact *only* when the subgroup has rank equal to the dimension of the space (*e.g.*, $\mathbf{R}^2/(\mathbf{Z} \times \{0\})$ is a non-compact infinite cylinder). That proves $L(U) \cong \mathbf{Z}^{r_1+r_2-1}$ and $L(U)$ is a lattice in the hyperplane H.

We're now basically done. Let $\varepsilon_1, \ldots, \varepsilon_r$ ($r = r_1 + r_2 - 1$) be elements of $\mathcal{O}^\times$ whose Euclidean embeddings in U provide a $\mathbf{Z}$-basis of $L(U)$. The ε_i's are multiplicatively independent, since their L-images are $\mathbf{Z}$-linearly independent. For $\varepsilon \in \mathcal{O}^\times$, $L(\varepsilon) = m_1 L(\varepsilon_1) + \cdots + m_r L(\varepsilon_r)$ for some integers m_i, so $L(\varepsilon) = L(\varepsilon_1^{m_1} \cdots \varepsilon_r^{m_r})$. Since $\ker(L|_U)$ is the Euclidean image of the roots of unity in $\mathcal{O}^\times$, $\varepsilon = \zeta \varepsilon_1^{m_1} \cdots \varepsilon_r^{m_r}$ for some $\zeta \in \mu(\mathcal{O})$. This concludes the proof of the unit theorem.

The most difficult part of the proof of the unit theorem is showing there are $r_1 + r_2 - 1$ independent units of infinite order. For instance, using the logarithmic map it was not hard for us to show $L(U)$ is a discrete subgroup of $L(G) \cong \mathbf{R}^{r_1+r_2-1}$, so $\mathcal{O}^\times \cong U \cong W \times \mathbf{Z}^{r'}$ where $r' \leq r_1 + r_2 - 1$ and W is the group of roots of unity in $\mathcal{O}^\times$. Thus $\mathcal{O}^\times$ has *at most* $r_1 + r_2 - 1$ independent units of infinite order, but this alone doesn't tell us there are units

[3]We are identifying $U = \theta_K(\mathcal{O}^\times)$ with $\mathcal{O}^\times$ when we speak of real embeddings of u. If we did not make that identification, and wrote $u = \theta_K(\alpha)$, then we would speak instead of real embeddings of α, which are the initial coordinates of u.

of infinite order in $\mathcal{O}^\times$. The place in the proof where saw there are units of infinite order (if $r_1+r_2-1>0$) is when we went from $r' \leq r_1+r_2-1$ to $r' = r_1+r_2-1$. This happened two paragraphs up and relied on G/U being compact.

4. Fundamental Unit in the Rank 1 Case

As noted already in Example 2.3, an order $\mathcal{O}$ in number field K has a rank 1 unit group precisely when K is real quadratic, cubic with 1 real embedding (that is, a cubic field that is not totally real), or a totally complex quartic field. In the first two cases, the only roots of unity in K are ± 1, which are always in $\mathcal{O}$, so $\mathcal{O}^\times = \pm\varepsilon^{\mathbf{Z}}$.[4] Viewing K in $\mathbf{R}$, the choice of $\varepsilon > 1$ is called *the* fundamental unit of $\mathcal{O}$.

Example 4.1. Since $\mathbf{Z}[\sqrt{2}]^\times = \pm(1+\sqrt{2})^{\mathbf{Z}}$, the fundamental unit of $\mathbf{Z}[\sqrt{2}]$ is $1+\sqrt{2}$.

Example 4.2. Since $\mathbf{Z}[3\sqrt{2}]^\times = \pm(17+12\sqrt{2})^{\mathbf{Z}}$, $\mathbf{Z}[3\sqrt{2}]$ has fundamental unit $17+12\sqrt{2}$.

Example 4.3. In Example 4.9 we will show $\mathbf{Z}[\sqrt[3]{6}]^\times = \pm(109+60\sqrt[3]{6}+33\sqrt[3]{36})^{\mathbf{Z}}$, so $109+60\sqrt[3]{6}+33\sqrt[3]{36} \approx 326.990$ is the fundamental unit of $\mathbf{Z}[\sqrt[3]{6}]$.

In a real quadratic field, one way to find the fundamental unit in an order is by brute force: if we write a unit greater than 1 as $a+b\sqrt{d}$ or $a+b(1+\sqrt{d})/2$ with $a, b \in \mathbf{Z}$, necessarily $a \geq 0$ and $b \geq 1$ (check!). This allows one to systematically search for the smallest unit greater than 1 by sifting through pairs of integers in the first quadrant by increasing values of a and b. (There is a more efficient method, using continued fractions.)

To give examples of fundamental unit computations in the cubic case, we will use an inequality due to Artin [1, pp. 169–170]. Mordell [10], near the end of his review of [1] in 1962, described Artin's inequality as a "surprise" since "one would have thought that there was not much opportunity for new results on cubic units".

Theorem 4.4 (Artin)**.** *Let $\mathcal{O}$ be an order in a cubic field K with $r_1 = 1$. Viewing K in $\mathbf{R}$, if $v > 1$ is a unit of $\mathcal{O}^\times$ then $|\operatorname{disc}(\mathcal{O})| < 4v^3 + 24$.*

Proof. This argument is similar to Artin's in [1] and may look like a messy calculation. Consider reading the corollary and its applications first, and then return to this proof.

Since v is a unit and is not ± 1, $v \notin \mathbf{Q}$. Thus $\mathbf{Q}(v) = K$, so $\mathbf{Z}[v]$ is an order inside $\mathcal{O}$. From $\mathbf{Z}[v] \subset \mathcal{O}$, $|\operatorname{disc}(\mathcal{O})| \leq |\operatorname{disc}(\mathbf{Z}[v])|$. We will show $|\operatorname{disc}(\mathbf{Z}[v])| < 4v^3+24$.

Let $\sigma\colon K \to \mathbf{C}$ be one of the non-real embeddings of K. Then $\mathrm{N}_{K/\mathbf{Q}}(v) = v\sigma(v)\overline{\sigma}(v) = v|\sigma(v)|^2 > 0$, so v has norm 1. Let $x = \sqrt{v}$ (as a positive real number), so $1 = x^2|\sigma(v)|^2$. Therefore $|\sigma(v)| = 1/x$, so in polar form $\sigma(v) = x^{-1}e^{it}$ for some real number t. Then

$$\begin{aligned}
\operatorname{disc}(\mathbf{Z}[v]) &= ((\sigma(v)-v)(\overline{\sigma}(v)-v)(\sigma(v)-\overline{\sigma}(v)))^2 \\
&= ((x^{-1}e^{it}-x^2)(x^{-1}e^{-it}-x^2)(x^{-1}e^{it}-x^{-1}e^{-it}))^2 \\
&= ((x^{-2}+x^4-2x\cos t)(-2ix^{-1}\sin t))^2 \\
&= -4(\sin^2 t)(x^3+x^{-3}-2\cos t)^2.
\end{aligned}$$

Since $x > 1$, $x^3 + x^{-3} > 2$. Set $a = (x^3+x^{-3})/2$, so $a > 1$ and by taking absolute values,

$$\begin{aligned}
|\operatorname{disc}(\mathbf{Z}[v])| &= 4(\sin^2 t)(2a-2\cos t)^2 \\
&= 16(1-\cos^2 t)(a-\cos t)^2. \qquad (4.1)
\end{aligned}$$

[4] Don't confuse $\pm\varepsilon^{\mathbf{Z}}$ with $\varepsilon^{\pm\mathbf{Z}}$; the latter is just $\varepsilon^{\mathbf{Z}}$.

Set $y = \cos t$, so $y \in [-1,1]$. Then (4.1) is

$$f(y) = 16(1-y^2)(a-y)^2$$

and we want to maximize this on $[-1,1]$. Let a maximum occur at y_0. Since $f(y) \geq 0$ on $[-1,1]$ with $f(1) = f(-1) = 0$ and $f(0) = 16a^2 > 0$, we have $y_0 \in (-1,1)$ and $f'(y_0) = 0$.

By the product rule, $f'(y) = 32(a-y)(2y^2 - ay - 1)$ and the root of the linear factor is $a > 1$, so $2y_0^2 - ay_0 - 1 = 0$. Rewrite this as

$$ay_0 = 2y_0^2 - 1. \tag{4.2}$$

Thus

$$|\operatorname{disc}(\mathbf{Z}[v])| = f(\cos t) \leq f(y_0) = 16(1-y_0^2)(a-y_0)^2. \tag{4.3}$$

Expanding $(a-y_0)^2$ and using the relation (4.2) a couple of times, we get

$$16(1-y_0^2)(a-y_0)^2 = 16(a^2 + 1 - y_0^4 - y_0^2).$$

Substituting $a = (x^3 + x^{-3})/2$ into this,

$$f(y_0) = 16\left(\frac{x^6}{4} + \frac{3}{2} + \left(\frac{x^{-6}}{4} - y_0^4 - y_0^2\right)\right).$$

We will show $x^{-6}/4 < y_0^2$, so the right side is less than $16(x^6/4 + 3/2) = 4x^6 + 24 = 4v^3 + 24$. Then by (4.3), $|\operatorname{disc}(\mathbf{Z}[v])| < 4v^3 + 24$, as desired.

Let $h(y) = 2y^2 - ay - 1$, the quadratic factor of $f'(y)$, so $h'(y_0) = 0$. The roots of $h(y)$ have product $-1/2$, so they have opposite sign. Since $h(1) = 1 - a < 0$ and $h(y) > 0$ for large y, $h(y)$ has a root in $(1,\infty)$. Thus $-1 < y_0 < 0$, and recall $x > 1$, so the inequality $x^{-6}/4 < y_0^2$ is the same as $y_0 < -1/(2x^3)$. The graph of $h(y)$ is a concave up parabola and y_0 is the smaller root of $h(y)$, so to prove $y_0 < -1/(2x^3)$ it is enough to show $h(-1/(2x^3)) < 0$:

$$h\left(\frac{-1}{2x^3}\right) = \frac{2}{4x^6} + \frac{a}{2x^3} - 1 = \frac{1}{2x^6} + \frac{1}{2x^3}\frac{x^3 + x^{-3}}{2} - 1 = \frac{3}{4x^6} - \frac{3}{4} = \frac{3}{4}\left(\frac{1}{x^6} - 1\right) < 0$$

since $x > 1$. □

Remark 4.5. The condition on v in Theorem 4.4 is $v > 1$, *not* $v > 0$. If we could use $0 < v < 1$ in Artin's inequality, then replacing v with a high power of itself would imply $|\operatorname{disc}(\mathcal{O})| < 24$, which is false for all cubic orders with one exception. See Footnote 6 below.

Corollary 4.6. *Let $\mathcal{O}$ be an order in a cubic field K with $r_1 = 1$. Viewing K inside $\mathbf{R}$, let $\varepsilon > 1$ be the fundamental unit of $\mathcal{O}$. For each unit $u > 1$ in $\mathcal{O}^\times$, if $4u^{3/m} + 24 \leq |\operatorname{disc}(\mathcal{O})|$ for an integer $m \geq 2$ then $u = \varepsilon^k$ where $1 \leq k < m$. In particular, if $4u^{3/2} + 24 \leq |\operatorname{disc}(\mathcal{O})|$ then $u = \varepsilon$.*

Proof. The group $\mathcal{O}^\times$ is $\pm\varepsilon^{\mathbf{Z}}$, so $u = \varepsilon^k$ for some positive integer k. Artin's inequality using $v = \varepsilon$ says

$$|\operatorname{disc}(\mathcal{O})| < 4\varepsilon^3 + 24 = 4u^{3/k} + 24.$$

If $k \geq m$ then $|\operatorname{disc}(\mathcal{O})| < 4u^{3/k} + 24 \leq 4u^{3/m} + 24 \leq |\operatorname{disc}(\mathcal{O})|$, so we have a contradiction. Thus $k < m$. □

Example 4.7. Let $K = \mathbf{Q}(\sqrt[3]{2})$, so $\mathcal{O}_K = \mathbf{Z}[\sqrt[3]{2}]$ and $\operatorname{disc}(\mathcal{O}_K) = \operatorname{disc}(T^3 - 2) = -108$. Since

$$1 = \sqrt[3]{2}^3 - 1 = (\sqrt[3]{2} - 1)(\sqrt[3]{4} + \sqrt[3]{2} + 1),$$

we have a unit $u = 1 + \sqrt[3]{2} + \sqrt[3]{4} \approx 3.847$. Since $4u^{3/2} + 24 \approx 54.185 < 108$, u is the fundamental unit of $\mathcal{O}_K$.

Example 4.8. Let $K = \mathbf{Q}(\alpha)$, where $\alpha^3+2\alpha+1 = 0$. The polynomial is irreducible modulo 3, so $K/\mathbf{Q}$ is cubic. It has one real root, approximately $-.45$. Since $\mathrm{disc}(T^3+2T+1) = -59$, $\mathcal{O}_K = \mathbf{Z}[\alpha]$. Clearly α is a unit. We view K in $\mathbf{R}$. Since $\alpha \approx -.45$, we get a unit greater than 1 using

$$u = -\frac{1}{\alpha} \approx 2.205.$$

Since $4u^{3/2} + 24 \approx 37.10 < 59$, u is the fundamental unit of $\mathcal{O}_K$.

Example 4.9. Let $K = \mathbf{Q}(\sqrt[3]{6})$. This will be an example where the unit u we find will satisfy $4u^{3/2} + 24 > |\mathrm{disc}(\mathcal{O}_K)|$, so we will have to be more creative to prove u is the fundamental unit.

First we show $\mathcal{O}_K = \mathbf{Z}[\sqrt[3]{6}]$. Since

$$\mathrm{disc}(\mathbf{Z}[\sqrt[3]{6}]) = \mathrm{disc}(T^3 - 6) = -2^2 3^5 = -972 = [\mathcal{O}_K : \mathbf{Z}[\sqrt[3]{6}]]^2 \,\mathrm{disc}(\mathcal{O}_K),$$

the index $[\mathcal{O}_K : \mathbf{Z}[\sqrt[3]{6}]]$ is a factor of $2 \cdot 3^2$. At the same time, since $T^3 - 6$ is Eisenstein at 2 and 3 the index $[\mathcal{O}_K : \mathbf{Z}[\sqrt[3]{6}]]$ is not divisible by 2 or 3. Therefore the index is 1, so $\mathcal{O}_K = \mathbf{Z}[\sqrt[3]{6}]$.

To find units in $\mathcal{O}_K$, we seek two different descriptions of a principal ideal in $\mathcal{O}_K$: if $(\alpha) = (\beta)$ then $\alpha = \beta u$ where u is a unit. Here is a table of how the first few primes p decompose in $\mathcal{O}_K$, based on how $T^3 - 6 \bmod p$ decomposes.

p	$T^3 - 6 \bmod p$	(p)
2	T^3	$\mathfrak{p}_2^3$
3	T^3	$\mathfrak{p}_3^3$
5	$(T-1)(T^2+T+1)$	$\mathfrak{p}_5\mathfrak{p}_{25}$
7	$(T-3)(T-5)(T-6)$	$\mathfrak{p}_7\mathfrak{p}_7'\mathfrak{p}_7''$

The only ideal of norm 2 is $\mathfrak{p}_2$. We will prove $\mathfrak{p}_2$ is principal by finding an element of absolute norm 2. For $c \in \mathbf{Z}$, $\mathrm{N}_{K/\mathbf{Q}}(\sqrt[3]{6} + c) = c^3 + 6$. Therefore $\mathrm{N}_{K/\mathbf{Q}}(\sqrt[3]{6} - 2) = -2$, so the ideal $(\sqrt[3]{6} - 2)$ has norm 2 and must be $\mathfrak{p}_2$. We have the equality of principal ideals

$$(2) = \mathfrak{p}_2^3 = (\sqrt[3]{6} - 2)^3 = ((\sqrt[3]{6} - 2)^3),$$

so the numbers 2 and $(\sqrt[3]{6} - 2)^3$ are equal up to a unit multiple in $\mathcal{O}_K$. Since $(\sqrt[3]{6} - 2)^3 \approx -.0061$, to get a unit greater than 1 we use the ratio[5]

$$u = -\frac{2}{(\sqrt[3]{6} - 2)^3} \approx 326.9908.$$

Let $\varepsilon > 1$ be the fundamental unit of $\mathbf{Z}[\sqrt[3]{6}]$. Does $u = \varepsilon$? By the unit theorem $u = \varepsilon^k$ for some $k \geq 1$ and we want to show $k = 1$. Artin's inequality with $v = \varepsilon$ says that in $\mathbf{R}$,

$$|\mathrm{disc}(\mathcal{O}_K)| < 4\varepsilon^3 + 24 \Longrightarrow 972 < 4u^{3/k} + 24.$$

For large k this inequality must fail, since the right side tends to $4 + 24 = 28$ as $k \to \infty$. In fact $4u^{3/4} + 24 \approx 331.5 < 972$, so k is either 1, 2, or 3. How do we rule out $u = \varepsilon^2$ and $u = \varepsilon^3$?

Here's a great idea: to prove an algebraic integer is not a square or cube, prove it is not a square or cube modulo $\mathfrak{p}$ for some prime ideal $\mathfrak{p}$. Looking at the above table of prime ideal factorizations, we will use the ideals $\mathfrak{p}_5$ and $\mathfrak{p}_7$.

[5]Explicitly, $u = 109 + 60\sqrt[3]{6} + 33\sqrt[3]{36}$, but this representation will not be needed.

In $\mathcal{O}_K/\mathfrak{p}_5 \cong \mathbf{Z}/(5)$ we have $\sqrt[3]{6} \equiv 1 \bmod \mathfrak{p}_5$, so $u \equiv 2/(2-1)^3 \equiv 2 \bmod \mathfrak{p}_5$. The nonzero squares in $\mathbf{Z}/(5)$ are 1 and 4, so u is not a square in $\mathcal{O}_K/\mathfrak{p}_5$ and thus is not a square in $\mathcal{O}_K$.

In $\mathcal{O}_K/\mathfrak{p}_7 \cong \mathbf{Z}/(7)$ we have $\sqrt[3]{6} \equiv 3 \bmod \mathfrak{p}_7$, so $u \equiv 2/(2-3)^3 \equiv -2 \equiv 5 \bmod \mathfrak{p}_7$. The nonzero cubes in $\mathbf{Z}/(7)$ are 1 and 6, so u is not a cube in $\mathcal{O}_K$. (Using $\mathfrak{p}_7'$ or $\mathfrak{p}_7''$ would have led to the same conclusion.)

We have shown $k = 1$, so $u = \varepsilon$ is the fundamental unit of $\mathbf{Z}[\sqrt[3]{6}]$.

Example 4.10. Let $K = \mathbf{Q}(\alpha)$ where $\alpha^3 - \alpha - 1 = 0$. The polynomial $T^3 - T - 1$ is irreducible mod 5, so $K/\mathbf{Q}$ is cubic. The polynomial has one real root $\alpha \approx 1.324$, so $r_1 = 1$. Since $\operatorname{disc}(T^3 - T - 1) = -23$ is squarefree, $\mathcal{O}_K = \mathbf{Z}[\alpha]$. Clearly α is a unit in $\mathcal{O}_K$. It is natural to wonder if α is the fundamental unit of $\mathcal{O}_K$ since it is so close to 1 in the real embedding. We can't use Artin's inequality because $|\operatorname{disc}(\mathcal{O}_K)| < 24$,[6] so for every unit $u > 1$ and positive integer m, $|\operatorname{disc}(\mathcal{O}_K)| < 4u^{3/m} + 24$.

Since we know the unit group of $\mathcal{O}_K$ (modulo ± 1) is infinite cyclic, to show α is the fundamental unit we show α is the smallest unit greater than 1: no unit $u \in \mathbf{Z}[\alpha]^\times$ satisfies $1 < u < \alpha$. Let $\sigma\colon K \to \mathbf{C}$ be one of the complex embeddings of K, so $\mathrm{N}_{K/\mathbf{Q}}(u) = u\sigma(u)\overline{\sigma}(u) = u|\sigma(u)|^2 > 0$. Therefore $\mathrm{N}_{K/\mathbf{Q}}(u) = 1$. Since $u \notin \mathbf{Q}$, the minimal polynomial of u over $\mathbf{Q}$ is $T^3 + aT^2 + bT - 1$ for some integers a and b. The roots are $u, \sigma(u)$, and $\overline{\sigma}(u)$, so

$$a = -(u + \sigma(u) + \overline{\sigma}(u)), \quad b = u\sigma(u) + u\overline{\sigma}(u) + \sigma(u)\overline{\sigma}(u).$$

Then

$$|a| \leq u + 2|\sigma(u)|, \quad |b| \leq 2u|\sigma(u)| + |\sigma(u)|^2.$$

Since $1 = u|\sigma(u)|^2$, the bound $1 \leq u$ implies $|\sigma(u)| \leq 1$, so from $1 < u < \alpha$ we get

$$|a| < \alpha + 2 \approx 3.3, \quad |b| \leq 2\alpha + 1 \approx 3.6.$$

Thus a and b both lie in $\{0, \pm 1, \pm 2, \pm 3\}$. Among the 49 polynomials $T^3 + aT^2 + bT - 1$ with a and b in this set, every such polynomial that has a unit $u > 1$ of $\mathcal{O}_K$ as a root must have discriminant equal to a nonzero square multiple of $\operatorname{disc}(\mathcal{O}_K) = -23$ (because $\operatorname{disc}(\mathbf{Z}[u]) = [\mathcal{O}_K : \mathbf{Z}[u]]^2 \operatorname{disc}(\mathcal{O}_K)$). There are four such polynomials (see table below), including $T^3 - T - 1$ itself, and the real root of each polynomial other than $T^3 - T - 1$ is larger than α.

Polynomial	Discriminant	Real Root
$T^3 - T - 1$	-23	$\alpha \approx 1.324$
$T^3 - 2T^2 + T - 1$	-23	$\alpha^2 \approx 1.754$
$T^3 - 3T^2 + 2T - 1$	-23	$\alpha^3 \approx 2.324$
$T^3 - 2T^2 - 3T - 1$	-23	$\alpha^4 \approx 3.079$

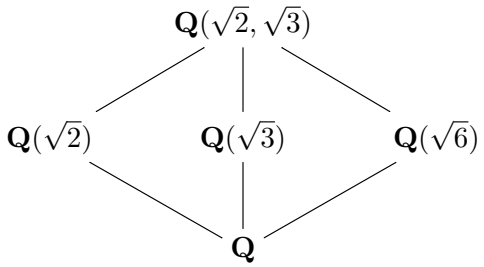
Thus α is the fundamental unit of $\mathcal{O}_K$.

5. Units in a multiquadratic field

A real quadratic field has unit rank 1. A biquadratic field $\mathbf{Q}(\sqrt{m}, \sqrt{n})$, where m, n, and mn are positive integers and not squares, has unit rank $4 - 1 = 3$. There are three quadratic subfields, $\mathbf{Q}(\sqrt{m})$, $\mathbf{Q}(\sqrt{n})$, and $\mathbf{Q}(\sqrt{mn})$, and each has a fundamental unit. A choice of one unit from each quadratic subfield need not be a set of 3 fundamental units for the biquadratic field.

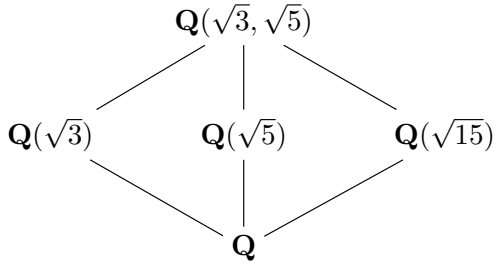
[6] This field K is the only cubic field, up to isomorphism, with absolute discriminant less than 24, so $\mathcal{O}_K$ is the only cubic order with absolute discriminant less than 24.

Example 5.1. In the field $\mathbf{Q}(\sqrt{2},\sqrt{3})$, a system of fundamental units is $1+\sqrt{2}$, $\sqrt{2}+\sqrt{3}$, and $\frac{\sqrt{2}+\sqrt{6}}{2}$ (see Table 1).



Fundamental units for the three quadratic subfields are $u_1 = 1+\sqrt{2}$, $u_2 = 2+\sqrt{3}$, and $u_3 = 5+2\sqrt{6}$. In terms of the fundamental units of $\mathbf{Q}(\sqrt{2},\sqrt{3})$, $u_2 = \left(\frac{\sqrt{2}+\sqrt{6}}{2}\right)^2$ and $u_3 = (\sqrt{2}+\sqrt{3})^2$. In terms of the fundamental units of the quadratic subfields, the fundamental units we listed for $\mathbf{Q}(\sqrt{2},\sqrt{3})$ are $u_1, \sqrt{u_3}, \sqrt{u_2}$.

Example 5.2. In the field $\mathbf{Q}(\sqrt{3},\sqrt{5})$, a system of fundamental units is $\frac{1+\sqrt{5}}{2}$, $4+\sqrt{15}$, and $\frac{3+\sqrt{3}+\sqrt{5}+\sqrt{15}}{2}$.



Fundamental units of the quadratic subfields are $u_1 = 2+\sqrt{3}$, $u_2 = \frac{1+\sqrt{5}}{2}$, and $u_3 = 4+\sqrt{15}$. In terms of the fundamental units of $\mathbf{Q}(\sqrt{3},\sqrt{5})$, $u_1 = \frac{1}{4+\sqrt{15}}\left(\frac{3+\sqrt{3}+\sqrt{5}+\sqrt{15}}{2}\right)^2$ by PARI. In terms of the fundamental units of the quadratic subfields, the fundamental units we listed for $\mathbf{Q}(\sqrt{3},\sqrt{5})$ are $u_2, u_3, \sqrt{u_1u_3}$.

Kuroda [7] showed that for every real biquadratic field $\mathbf{Q}(\sqrt{m},\sqrt{n})$, with fundamental units u_1, u_2, u_3 for its 3 quadratic subfields, a set of fundamental units for the biquadratic field is one of the following 7 lists up to relabeling the u_i's:

$$\{u_1,u_2,u_3\}, \{\sqrt{u_1},u_2,u_3\}, \{\sqrt{u_1u_2},u_2,u_3\}, \{\sqrt{u_1u_2u_3},u_2,u_3\},$$
$$\{\sqrt{u_1},\sqrt{u_2},u_3\}, \{\sqrt{u_1u_2},\sqrt{u_3},u_2\}, \{\sqrt{u_1u_2},\sqrt{u_2u_3},\sqrt{u_3u_1}\}.$$

Examples 5.1 and 5.2 illustrate two of these possibilities and the list shows $\langle -1, u_1, u_2, u_3\rangle$ has index 1, 2, 4, or 8 in the unit group of the biquadratic field.

Consider now a general multiquadratic field

$$K = \mathbf{Q}(\sqrt{d_1},\ldots,\sqrt{d_k}),$$

where the d_i's are nonsquare *positive* integers that are multiplicatively independent modulo squares (that is, they are independent in $\mathbf{Q}^\times/(\mathbf{Q}^\times)^2$). By Galois theory and induction, $[K:\mathbf{Q}] = 2^k$ and $\mathrm{Gal}(K/\mathbf{Q}) \cong (\mathbf{Z}/2\mathbf{Z})^k$ by making sign changes on every $\sqrt{d_i}$. The unit rank of K is $r_1 - 1 = 2^k - 1$, and this is also the number of quadratic subfields: such subfields

are of the form $\mathbf{Q}(\sqrt{d_I})$, where $I = \{i_1, \ldots, i_m\}$ is a nonempty subset of $\{1, 2, \ldots, k\}$ and $d_I = d_{i_1} \cdots d_{i_m}$. Since each $\mathbf{Q}(\sqrt{d_I})$ has unit rank 1, it is natural to suspect that choosing one unit (besides ± 1) from each quadratic subfield of K should give us a multiplicatively independent set of units in K.

Theorem 5.3. *With notation as above, let u_I be a unit in $\mathbf{Q}(\sqrt{d_I})$ other than ± 1. These units are multiplicatively independent: if $\prod_I u_I^{a_I} = 1$, where the exponents a_I are in $\mathbf{Z}$, then each a_I is 0.*

Proof. Our argument is taken from [8, Lemma 2] (which includes some extraneous hypotheses on the d_i's). The special feature of a unit in a real quadratic field is that its $\mathbf{Q}$-conjugate is, up to sign, its inverse: $u' = \pm u^{-1}$. This fact will interact well with multiplication relations.

One $\mathbf{Q}$-basis of K is all the square roots $\sqrt{d_I}$ together with 1 (we could set $d_\emptyset = 1$ and $1 = \sqrt{d_\emptyset}$). For each nonempty subset J of $\{1, 2, \ldots, k\}$, there is a $\sigma \in \mathrm{Gal}(K/\mathbf{Q})$ such that $\sigma(\sqrt{d_J}) = -\sqrt{d_J}$ and $\sigma(\sqrt{d_I}) = \sqrt{d_I}$ for all $I \neq J$. Since σ is the identity on $\mathbf{Q}(\sqrt{d_I})$ and is nontrivial on $\mathbf{Q}(\sqrt{d_J})$, $\sigma(u_I) = u_I$ while $\sigma(u_J) = \pm u_J^{-1}$.

Applying σ to $\prod_I u_I^{a_I} = 1$ turns it into $\prod_{I \neq J} u_I^{a_I} \cdot (\pm u_J^{-1})^{a_J} = 1$. Dividing one multiplicative relation by the other, $(\pm u_J^2)^{a_J} = 1$. Since u_J has infinite order, $a_J = 0$. □

Corollary 5.4. *The units u_I generate a subgroup of $\mathcal{O}_K^\times$ with finite index.*

Proof. By their multiplicative independence, the u_I's generate a group of rank $2^k - 1$, which is the rank of $\mathcal{O}_K^\times$. □

References

[1] E. Artin, *Theory of Algebraic Numbers*, George Striker, Schildweg 12, Göttingen, 1959.

[2] P. G. L. Dirichlet, *Vorlesungen über Zahlentheorie*, 4th ed., Vieweg und Sohn, Braunschweig, 1894. Online at https://archive.org/details/vorlesungenberz02dirigoog.

[3] R. Godement, *Introduction à la Théorie des Groups de Lie*, Springer–Verlag, Berlin, 2003.

[4] E. Kleinert, "Units of Classical Orders: A Survey," L'Enseignement Math. **40** (1994), 205–248.

[5] E. Kleinert, *Units in Skew Fields*, Birkhaüser, Basel, 2000.

[6] H. Koch, *Number Theory: Algebraic Numbers and Functions*, Amer. Math. Society, Providence, 2000.

[7] S. Kuroda, "Über den Dirichletschen Körper," J. Fac. Sci. Imp. Univ. Tokyo Sect. I **4** (1943), 383–406.

[8] F. Luca and I. E. Shparlinski, "On the Square-free Parts of $\lfloor en! \rfloor$," Glasgow Math. J. **49** (2007), 391–403.

[9] H. Minkowski, "Peter Gustav Lejeune Dirichlet und seine Bedeutung für die heutige Mathematik," Jahresbericht der Deutschen Mathematiker-Vereinigung **14** (1905), 149–163.

[10] L. J. Mordell, Review of E. Artin's "Theory of Algebraic Numbers", Bull. Amer. Math. Soc. **3** (1962), 162–166.