# Math 129: Algebraic Number Theory
# Lecture 10: Dirichlet's Unit Theorem

William Stein

Tuesday, March 9, 2004

Announcement: Unfortunately, I'll miss my office hours 2–3pm today, since I'll be flying to Baltimore to give a talk at Johns Hopkins University. However, you can talk to me from 11:30–12:00 today, and 1:00–1:45 (I'll be in my office then).

In this lecture we will prove the main structure theorem for the group of units of the ring of integers of a number field. The answer is remarkably simple: if $K$ has $r$ real and $s$ complex embeddings, then

$$\mathcal{O}_K^* \approx \mathbf{Z}^{r+s-1} \oplus W,$$

where $W$ is the finite cyclic group of roots of unity in $K$. Examples will follow on Thursday (application: the solutions to Pell's equation $x^2 - dy^2 = 1$, for $d > 1$ squarefree, form a free abelian group of rank 1).

## 1 The Group of Units

**Definition 1.1 (Unit Group).** The *group of units* $U_K$ associated to a number field $K$ is the group of elements of $\mathcal{O}_K$ that have an inverse in $\mathcal{O}_K$.

**Theorem 1.2 (Dirichlet).** *The group $U_K$ is the product of a finite cyclic group of roots of unity with a free abelian group of rank $r + s - 1$, where $r$ is the number of real embeddings of $K$ and $s$ is the number of complex conjugate pairs of embeddings.*

We prove the theorem by defining a map $\varphi : U_K \to \mathbf{R}^{r+s}$, and showing that the kernel of $\varphi$ is finite and the image of $\varphi$ is a lattice in a hyperplane in $\mathbf{R}^{r+s}$. The trickiest part of the proof is showing that the image of $\varphi$ spans a hyperplane, and we do this by a clever application of Blichfeldt's lemma (that if $S$ is closed, bounded, symmetric, etc., and has volume at least $2^n \cdot \mathrm{Vol}(V/L)$, then $S \cap L$ contains a nonzero element).

*Remark* 1.3. Theorem 1.2 is due to Dirichlet who lived 1805–1859. Thomas Hirst described Dirichlet as follows:

> He is a rather tall, lanky-looking man, with moustache and beard about to turn grey with a somewhat harsh voice and rather deaf. He was unwashed, with his cup of coffee and cigar. One of his failings is forgetting time, he pulls his watch out, finds it past three, and runs out without even finishing the sentence.

Koch wrote that:

> ... important parts of mathematics were influenced by Dirichlet. His proofs characteristically started with surprisingly simple observations, followed by extremely sharp analysis of the remaining problem.

I think Koch's observation nicely describes the proof we will give of Theorem 1.2.

The following proposition explains how to think about units in terms of the norm.

**Proposition 1.4.** *An element $a \in \mathcal{O}_K$ is a unit if and only if* $\mathrm{Norm}_{K/\mathbf{Q}}(a) = \pm 1$.

*Proof.* Write $\mathrm{Norm} = \mathrm{Norm}_{K/\mathbf{Q}}$. If $a$ is a unit, then $a^{-1}$ is also a unit, and $1 = \mathrm{Norm}(a)\,\mathrm{Norm}(a^{-1})$. Since both $\mathrm{Norm}(a)$ and $\mathrm{Norm}(a^{-1})$ are integers, it follows that $\mathrm{Norm}(a) = \pm 1$. Conversely, if $a \in \mathcal{O}_K$ and $\mathrm{Norm}(a) = \pm 1$, then the equation $aa^{-1} = 1 = \pm\mathrm{Norm}(a)$ implies that $a^{-1} = \pm\mathrm{Norm}(a)/a$. But $\mathrm{Norm}(a)$ is the product of the images of $a$ in $\mathbf{C}$ by all embeddings of $K$ into $\mathbf{C}$, so $\mathrm{Norm}(a)/a$ is also a product of images of $a$ in $\mathbf{C}$, hence a product of algebraic integers, hence an algebraic integer. Thus $a^{-1} \in \mathcal{O}_K$, which proves that $a$ is a unit. □

Let $r$ be the number of real and $s$ the number of complex conjugate embeddings of $K$ into $\mathbf{C}$, so $n = [K : \mathbf{Q}] = r + 2s$. Define a map

$$\varphi : U_K \to \mathbf{R}^{r+s}$$

by

$$\varphi(a) = (\log|\sigma_1(a)|, \ldots, \log|\sigma_{r+s}(a)|).$$

**Lemma 1.5.** *The image of $\varphi$ lies in the hyperplane*

$$H = \{(x_1, \ldots, x_{r+s}) \in \mathbf{R}^{r+s} : x_1 + \cdots + x_r + 2x_{r+1} + \cdots + 2x_{r+s} = 0\}. \tag{1.1}$$

*Proof.* If $a \in U_K$, then by Proposition 1.4,

$$\left(\prod_{i=1}^{r} |\sigma_i(a)|\right) \cdot \left(\prod_{i=r+1}^{s} |\sigma_i(a)|^2\right) = 1.$$

Taking logs of both sides proves the lemma. □

**Lemma 1.6.** *The kernel of $\varphi$ is finite.*

*Proof.* We have

$$\begin{aligned}\mathrm{Ker}(\varphi) &\subset \{a \in \mathcal{O}_K : |\sigma_i(a)| = 1 \text{ for all } i = 1, \ldots, r+2s\} \\ &\subset \sigma(\mathcal{O}_K) \cap X,\end{aligned}$$

where $X$ is the bounded subset of $\mathbf{R}^{r+2s}$ of elements all of whose coordinates have absolute value at most 1. Since $\sigma(\mathcal{O}_K)$ is a lattice [ref?], the intersection $\sigma(\mathcal{O}_K) \cap X$ is finite, so $\mathrm{Ker}(\varphi)$ is finite. □

**Lemma 1.7.** *The kernel of $\varphi$ is a finite cyclic group.*

*Proof.* It is a general fact that any finite subgroup of the multiplicative group of a field is cyclic. [Homework.] □

To prove Theorem 1.2, it suffices to proove that $\mathrm{Im}(\varphi)$ is a lattice in the hyperplane $H$ from (1.1), which we view as a vector space of dimension $r + s - 1$.

Define an embedding

$$\sigma : K \hookrightarrow \mathbf{R}^n \tag{1.2}$$

given by $\sigma(x) = (\sigma_1(x), \ldots, \sigma_{r+s}(x))$, where we view $\mathbf{C} \cong \mathbf{R} \times \mathbf{R}$ via $a + bi \mapsto (a, b)$. Note that this is exactly the same as the embedding

$$\begin{aligned}x \mapsto \big(&\sigma_1(x), \sigma_2(x), \ldots, \sigma_r(x), \\ &\mathrm{Re}(\sigma_{r+1}(x)), \ldots, \mathrm{Re}(\sigma_{r+s}(x)), \mathrm{Im}(\sigma_{r+1}(x)), \ldots, \mathrm{Im}(\sigma_{r+s}(x))\big),\end{aligned}$$

from before, except that we have re-ordered the last $s$ imaginary components to be next to their corresponding real parts.

**Lemma 1.8.** *The image of $\varphi$ is discrete in $\mathbf{R}^{r+s}$.*

*Proof.* Suppose $X$ is any bounded subset of $\mathbf{R}^{r+s}$. Then for any $u \in Y = \varphi^{-1}(X)$ the coordinates of $\sigma(u)$ are bounded in terms of $X$ (since log is an increasing function). Thus $\sigma(Y)$ is a bounded subset of $\mathbf{R}^n$. Since $\sigma(Y) \subset \sigma(\mathcal{O}_K)$, and $\sigma(\mathcal{O}_K)$ is a lattice in $\mathbf{R}^n$, it follows that $\sigma(Y)$ is finite. Since $\sigma$ is injective, $Y$ is finite, and $\varphi$ has finite kernel, so $\varphi(U_K) \cap X$ is finite, which implies that $\varphi(U_K)$ is discrete. □

To finish the proof of Theorem 1.2, we will show that the image of $\varphi$ spans $H$. Let $W$ be the $\mathbf{R}$-span of the image $\varphi(U_K)$, and note that $W$ is a subspace of $H$. We will show that $W = H$ indirectly by showing that if $v \notin H^\perp$, where $\perp$ is with respect to the dot product on $\mathbf{R}^{r+s}$, then $v \notin W^\perp$. This will show that $W^\perp \subset H^\perp$, hence that $H \subset W$, as required.

Thus suppose $z = (z_1, \ldots, z_{r+s}) \notin H^\perp$. Define a function $f : K^* \to \mathbf{R}$ by

$$f(x) = z_1 \log|\sigma_1(x)| + \cdots z_{r+s} \log|\sigma_{r+s}(x)|. \tag{1.3}$$

To show that $z \notin W^\perp$ we show that there exists some $u \in U_K$ with $f(u) \neq 0$.

Let

$$A = \sqrt{|d_K|} \cdot \left(\frac{2}{\pi}\right)^s \in \mathbf{R}_{>0}.$$

Choose any positive real numbers $c_1, \ldots, c_{r+s} \in \mathbf{R}_{>0}$ such that

$$c_1 \cdots c_r \cdot (c_{r+1} \cdots c_{r+s})^2 = A.$$

Let

$$\begin{aligned} S = \{(x_1, \ldots, x_n) \in \mathbf{R}^n : & \\ & |x_i| \le c_i \text{ for } 1 \le i \le r, \\ & |x_i^2 + x_{i+s}^2| \le c_i^2 \text{ for } r < i \le r+s\} \subset \mathbf{R}^n. \end{aligned}$$

Then $S$ is closed, bounded, convex, symmetric with respect to the origin, and of dimension $r + 2s$, since $S$ is a product of $r$ intervals and $s$ discs, each of which has these properties. Viewing $S$ as a product of intervals and discs, we see that the volume of $S$ is

$$\mathrm{Vol}(S) = \prod_{i=1}^{r} (2c_i) \cdot \prod_{i=1}^{s} (\pi c_i^2) = 2^r \cdot \pi^s \cdot A.$$

Recall *Blichfeldt's lemma* that if $L$ is a lattice and $S$ is closed, bounded, etc., and has volume at least $2^n \cdot \mathrm{Vol}(V/L)$, then $S \cap L$ contains a nonzero element. To apply this lemma, we take $L = \sigma(\mathcal{O}_K) \subset \mathbf{R}^n$, where $\sigma$ is as in (1.2). We showed, when proving finiteness of the class group, that $\mathrm{Vol}(\mathbf{R}^n/L) = 2^{-s}\sqrt{|d_K|}$. To check the hypothesis to Blichfeld's lemma, note that

$$\mathrm{Vol}(S) = 2^{r+s}\sqrt{|d_K|} = 2^n 2^{-s}\sqrt{|d_K|} = 2^n \,\mathrm{Vol}(\mathbf{R}^n/L).$$

Thus there exists a nonzero element $a \in S \cap \sigma(\mathcal{O}_K)$, i.e., a nonzero $a \in \mathcal{O}_K$ such that $|\sigma_i(a)| \le c_i$ for $1 \le i \le r+s$. We then have

$$\begin{aligned} |\operatorname{Norm}_{K/\mathbf{Q}}(a)| &= \left| \prod_{i=1}^{r+2s} \sigma_i(a) \right| \\ &= \prod_{i=1}^{r} |\sigma_i(a)| \cdot \prod_{i=r+1}^{s} |\sigma_i(a)|^2 \\ &\le c_1 \cdots c_r \cdot (c_{r+1} \cdots c_{r+s})^2 = A. \end{aligned}$$

Since $a \in \mathcal{O}_K$ is nonzero, we also have

$$|\operatorname{Norm}_{K/\mathbf{Q}}(a)| \ge 1.$$

Moreover, if for any $i \le r$, we have $|\sigma_i(a)| < \frac{c_i}{A}$, then

$$1 \le |\operatorname{Norm}_{K/\mathbf{Q}}(a)| < c_1 \cdots \frac{c_i}{A} \cdots c_r \cdot (c_{r+1} \cdots c_{r+s})^2 = \frac{A}{A} = 1,$$

a contradiction, so $|\sigma_i(a)| \geq \frac{c_i}{A}$ for $i = 1, \ldots, r$. Likewise, $|\sigma_i(a)|^2 \geq \frac{c_i^2}{A}$, for $i = r+1, \ldots, r+s$. Rewriting this we have

$$\frac{c_i}{|\sigma_i(a)|} \leq A \quad \text{for } i \leq r \quad \text{and} \quad \left(\frac{c_i}{|\sigma_i(a)|}\right)^2 \leq A \quad \text{for } i = r+1, \ldots, r+s.$$

Our strategy is to use an appropriately chosen $a$ to construct a unit $u \in U_K$ such $f(u) \neq 0$. First, let $b_1, \ldots, b_m$ be representative generators for the finitely many nonzero principal ideals of $\mathcal{O}_K$ of norm at most $A$. Since $|\operatorname{Norm}_{K/\mathbf{Q}}(a)| \leq A$, we have $(a) = (b_j)$, for some $j$, so there is a unit $u \in \mathcal{O}_K$ such that $a = ub_j$.

Let

$$s = s(c_1, \ldots, c_{r+s}) = z_1 \log(c_1) + \cdots + z_{r+s} \log(c_{r+s}),$$

and recall $f : K^* \to \mathbf{R}$ defined in (1.3) above. We first show that

$$|f(u) - s| \leq B = |f(b_j)| + \log(A) \cdot \left(\sum_{i=1}^{r} |z_i| + \frac{1}{2} \cdot \sum_{i=r+1}^{s} |z_i|\right). \tag{1.4}$$

We have

$$\begin{aligned}
|f(u) - s| &= |f(a) - f(b_j) - s| \\
&\leq |f(b_j)| + |s - f(a)| \\
&= |f(b_j)| + |z_1(\log(c_1) - \log(|\sigma_1(a)|)) + \cdots + z_{r+s}(\log(c_{r+s}) - \log(|\sigma_{r+s}(a)|))| \\
&= |f(b_j)| + |z_1 \cdot \log(c_1/|\sigma_1(a)|) + \cdots + \frac{z_{r+s}}{2} \cdot \log((c_{r+s}/|\sigma_{r+s}(a)|)^2)| \\
&\leq |f(b_j)| + \log(A) \cdot \left(\sum_{i=1}^{r} |z_i| + \frac{1}{2} \cdot \sum_{i=r+1}^{s} |z_i|\right).
\end{aligned}$$

The amazing thing about (1.4) is that the bound $B$ on the right hand side does not depend on the $c_i$. Suppose we can choose positive real numbers $c_i$ such that

$$c_1 \cdots c_r \cdot (c_{r+1} \cdots c_{r+s})^2 = A$$

and $s = s(c_1, \ldots, c_{r+s})$ is such that $|s| > B$. Then $|f(u) - s| \leq B$ would imply that $|f(u)| > 0$, which is exactly what we aimed to prove. It is possible to choose such $c_i$, by proceeding as follows. If $r + s = 1$, then we are trying to prove that $\varphi(U_K)$ is a lattice in $\mathbf{R}^0 = \mathbf{R}^{r+s-1}$, which is automatically true, so assume $r+s > 1$. Then there are at least two distinct $c_i$. Let $j$ be such that $z_j \neq 0$ (which exists since $z \neq 0$). Then $|z_j \log(c_j)| \to \infty$ as $c_j \to \infty$, so we choose $c_j$ very large and the other $c_i$, for $i \neq j$, in any way we want subject to the condition

$$\prod_{i=1, i\neq j}^{r} c_i \cdot \prod_{i=r+1}^{s} c_i^2 = \frac{A}{c_j}.$$

Since it is possible to choose the $c_i$ as needed, it is possible to find a unit $u$ such that $f(u) > 0$. We conclude that $z \notin W^\perp$, so $W^\perp \subset Z^\perp$, whence $Z \subset W$, which finishes the proof Theorem 1.2.