CONJUGACY PROBLEM IN $\text{GL}_2(\mathbb{Z}[\sqrt{-1}])$ AND UNITS OF QUADRATIC EXTENSIONS OF $\mathbb{Q}(\sqrt{-1})$

BY
HIRONORI ONISHI

ABSTRACT. A highly efficient procedure for deciding if two given elements of $\text{GL}_2(\mathbb{Z}[\sqrt{-1}])$ are conjugate or not will be presented. It makes use of a continued fraction algorithm in $\mathbb{Z}[\sqrt{-1}]$ and gives a fundamental unit of any given quadratic extension of $\mathbb{Q}(\sqrt{-1})$.

(1) Introduction. A solution to the conjugacy problem in the group $G = \text{GL}_2(\mathbb{Z}[\sqrt{-1}])$ is included in the result of Grunewald [3]. But for a nice group like this there ought to be a simpler solution which makes use of the special nature of $G$. On the other hand, since $G$ is not an amalgam of simpler groups, we should not expect too easy a solution. In this paper we present a straightforward procedure for deciding if two given elements of $G$ are conjugate or not. It is based on a continued fraction algorithm in the ring $\mathbb{Z}[\sqrt{-1}]$ and a module theoretic consideration. It combines the ideas used in [1 and 2]. As the examples show it is highly efficient. A similar solution can be given for the group $\text{GL}_2(\mathcal{O})$, where $\mathcal{O}$ is the ring of integers of any imaginary quadratic field, but in order to fix our attention we shall deal with the case when $\mathcal{O} = \mathbb{Z}[i]$, $i = \sqrt{-1}$.

(2) Actually what we solve is the similarity problem for the $2 \times 2$ matrices over $\mathcal{O} = \mathbb{Z}[i]$; given two such matrices $A$ and $B$ the problem is to decide if there is an $R \in \text{GL}_2(\mathcal{O})$ such that $RAR^{-1} = B$. Our solution gives an explicit $R$ if there is one. It also gives an effective characterization of the centralizer

$$Z(A) = \{R \in \text{GL}_2(\mathcal{O}) \mid RA = AR\}$$

for a given $A$, so that we can find all $R \in \text{GL}_2(\mathcal{O})$ such that $RAR^{-1} = B$. The characterization of $Z(A)$ is obtained by finding a fundamental unit of an order in a quadratic extension of $F = \mathbb{Q}(i)$; our method generates a fundamental unit.

(3) Given $2 \times 2$ matrices $A$ and $B$ over $\mathcal{O}$, call $A \sim B$ similar if $RAR^{-1} = B$ for some $R \in \text{GL}_2(\mathcal{O})$. If $A \sim B$, then $A$ and $B$ have the same characteristic polynomial $f$ over $\mathcal{O}$. Given a monic quadratic polynomial $f$ over $\mathcal{O}$, let $M(f)$ denote the set of $2 \times 2$ matrices over $\mathcal{O}$ whose characteristic polynomials are equal to $f$. In deciding if $A \sim B$, we may assume that $A$ and $B \in M(f)$ for some $f$. When $f$ is reducible over $F$, deciding if $A \sim B$ is easy and we discuss it in the Appendix.

(4) Assume that $f$ is irreducible over $F$. Put

$$f(t) = t^2 - qt + r, \quad \Delta = q^2 - 4r.$$
Given

\[ A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in M(f), \]

put \( \lambda = (q + \sqrt{\Delta})/2 \), where \( \text{Im}(\sqrt{\Delta}) > 0 \) or \( \sqrt{\Delta} > 0 \). The number \( \lambda \) is an eigenvalue of \( A \). Put

\[ \phi(A) = \alpha = (\lambda - d)/b = (a - d + \sqrt{\Delta})/2b. \]

Since \( f \) is irreducible, \( b\alpha \neq 0 \). The column vector \((\alpha, 1)^T\) is an eigenvector of \( A \) belonging to \( \lambda \) and \( A\alpha = \alpha \) (under the projective action of \( A \) on \( \mathbf{C} \)).

(5) Put \( K = F(\sqrt{\Delta}) \). Given \( \xi \in K \), let \( \xi' \) denote its conjugate over \( F \). Given \( A \in M(f) \), if \( \alpha = \phi(A) \) then

\[ A = \begin{pmatrix} \alpha & \alpha' \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & \lambda' \end{pmatrix} \begin{pmatrix} \alpha & \alpha' \\ 1 & 1 \end{pmatrix}^{-1}. \]

Thus the map \( \phi: M(f) \to K \) is injective (for a given \( f \)). For any \( R \in \text{GL}_2(\mathcal{O}) \),

\[ RAR^{-1} = \begin{pmatrix} \alpha & \alpha' \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & \lambda' \end{pmatrix} \begin{pmatrix} \alpha & \alpha' \\ 1 & 1 \end{pmatrix}^{-1}. \]

Thus by injectivity of \( \phi \) on \( M(f) \), we have \( \phi(RAR^{-1}) = R\phi(A) \). Given \( \alpha \) and \( \beta \in K - F \), call \( \alpha \sim \beta \) if \( R\alpha = \beta \) for some \( R \in \text{GL}_2(\mathcal{O}) \). From the discussion above, we see that, given \( A \) and \( B \in M(f) \),

\[ A \sim B \ \text{iff} \ \phi(A) \sim \phi(B). \]

Thus the problem is transformed to this: Given \( \alpha \) and \( \beta \in K - F \), decide if \( \alpha \sim \beta \).

(6) Given \( \alpha \) and \( \beta \in K \), let \( \langle \alpha, \beta \rangle \) denote the module over \( \mathcal{O} \) generated by \( \alpha \) and \( \beta \). In this paper, by a module we shall mean a finitely generated full module over \( \mathcal{O} \) contained in \( K \). Every module is of the form \( \langle \alpha, \beta \rangle \) and \( \langle \alpha, \beta \rangle \) is a basis of this module over \( \mathcal{O} \). For example, if \( \alpha \in K - F \), then \( \langle \alpha, 1 \rangle \) is a module. Given modules \( U \) and \( V \), call \( U \sim V \) similar if \( U = \lambda V \) for some \( \lambda \in K^\times \).

(7) Given \( \alpha \) and \( \beta \in K - F \), put \( U = \langle \alpha, 1 \rangle \) and \( V = \langle \beta, 1 \rangle \). Then

\[ \alpha \sim \beta \ \text{iff} \ \ U \sim V. \]

In fact, if \( \alpha \sim \beta \), say \( R\alpha = \beta \), \( R \in \text{GL}_2(\mathcal{O}) \), then \( R\alpha_1 = \lambda(\beta) \) for some \( \lambda \in K^\times \) and hence \( U = \lambda V \), i.e., \( U \sim V \). Going backward we get the converse. Thus the problem is now transformed to the following: Given modules \( U \) and \( V \), decide if \( U \sim V \).

(8) Let \( U = \langle \alpha, \beta \rangle \) be a module. An element \( \xi = x\alpha + y\beta \) of \( U \), where it is understood that \( x \) and \( y \in \mathcal{O} \), is called primitive if \( (x, y) = 1 \), i.e., \( x \) and \( y \) are coprime. The primitiveness of an element of \( U \) does not depend on the choice of a basis \( \langle \alpha, \beta \rangle \) of \( U \). A member of a basis is primitive. It is easy to see that if \( \rho \) is a primitive element of \( U \), then \( U = \langle \sigma, \rho \rangle \) for some \( \sigma \in U \). A module \( U \) is called normalized if 1 is a primitive element of \( U \) so that \( U = \langle \alpha, 1 \rangle \) for some \( \alpha \in K - F \). Given modules \( U \) and \( V \), call \( U \equiv V \) if \( U = cv \) for some \( c \in F^\times \).

(9) For any module \( U \), there is a unique normalized module \( V \) such that \( U \equiv V \).

**Proof.** \( U \cap \mathcal{O} \) is a nonzero fractional ideal of \( \mathcal{O} \) and hence \( U \cap \mathcal{O} = (b) \) for some \( b \in F^\times \) and \( b \) has to be a primitive element of \( U \). Thus \( U = \langle \alpha, b \rangle \) for
some $\alpha$. $V = b^{-1}U = \langle ab^{-1}, 1 \rangle$ is normalized and $U \equiv V$. To see the uniqueness, suppose that $U$ and $V$ are normalized modules such that $U \equiv V$, say $U = cV$, $c \in \mathcal{O}^\times$, $U = \langle \alpha, 1 \rangle$, and $V = \langle \beta, 1 \rangle$. Then $\langle \alpha, 1 \rangle = \langle c\alpha, c \rangle$ and hence there is an $R \in \text{GL}_2(\mathcal{O})$ such that $R^{(\alpha)} = c^{(\beta)}$. Since $c \in \mathcal{O}$ and $\alpha \not\in \mathcal{O}$, $R$ has to be of the form

$$R = \begin{pmatrix} x & y \\ 0 & c \end{pmatrix}$$

with $xc = \det R \in \mathcal{O}^\times = \{ \pm 1, \pm i \}$. Thus $c \in \mathcal{O}^\times$ and $U = V$.

(10) Given $\alpha$ and $\beta \in K - F$, call $\alpha \equiv \beta$ if

$$\begin{pmatrix} \varepsilon & c \\ 0 & 1 \end{pmatrix} \alpha = \varepsilon \alpha + c = \beta$$

for some $\varepsilon \in \mathcal{O}^\times$ and $c \in \mathcal{O}$. From the proof of (9), it is clear that, given normalized modules $U = \langle \alpha, 1 \rangle$ and $V = \langle \beta, 1 \rangle$,

$$U = V \iff \alpha \equiv \beta.$$

We assume that, given $\alpha$ and $\beta \in K - F$, recognizing if $\alpha \equiv \beta$ is instantaneous. For example, if

$$\alpha = (e_1 + \sqrt{\Delta})/2b_1 \quad \text{and} \quad \beta = (e_2 + \sqrt{\Delta})/2b_2,$$

where $e_1, b_1, e_2, b_2 \in \mathcal{O}$, then $\alpha \equiv \beta$ iff $\varepsilon b_1 = b_2$ for some $\varepsilon \in \mathcal{O}^\times$ and $e_1 \equiv e_2 \pmod{2b_1}$.

(11) Let $U$ be a module. A nonzero element $\rho$ of $U$ is called a convergent of $U$ if $0$ is the only element $\varnothing$ of $U$ such that $|\varnothing| < |\rho|$ and $|\varnothing'| < |\rho'|$.

Note that given $\alpha$ and $\beta \in K$, $|\alpha| = |\beta|$ iff $|\alpha'| = |\beta'|$. (This can be easily proved by looking at $\gamma = \alpha/\beta$ and its complex conjugate $\gamma$ and their norms.) For any $\lambda \in K^\times$, as $\rho$ ranges over the convergents of $U$, $\lambda \rho$ ranges over the convergents of $\lambda U$.

(12) Let $U = \langle \alpha, \beta \rangle$. If $\xi = x\alpha + y\beta \in U$, then $\xi' = x\alpha' + y\beta'$ and

$$x = (\xi \beta' - \xi' \beta)/(\alpha \beta' - \alpha' \beta) \quad \text{and} \quad y = (\alpha \xi' - \alpha' \xi)/(\alpha \beta' - \alpha' \beta).$$

Thus if $|\xi|$ and $|\xi'|$ are bounded, then $|x|$ and $|y|$ are bounded. Thus for any $c_1$ and $c_2 > 0$, $U$ contains only a finite number of element $\xi$ such that $|\xi| < c_1$ and $|\xi'| < c_2$. This shows that there are convergents of $U$.

(13) Let $\rho$ be a convergent of $U$. Then $\rho$ is a primitive element of $U$ and $U = \langle \sigma, \rho \rangle$ for some $\sigma$ and $\rho^{-1}U = \langle \sigma \rho^{-1}, 1 \rangle$ is normalized. The normalized module $\rho^{-1}U$ is called a derived module of $U$. Let $\mathcal{D}(U)$ denote the set of all derived modules of $U$, i.e., $\mathcal{D}(U) = \{ \rho^{-1}U \mid \rho \text{ is a convergent of } U \}$.

(14) If $V \in \mathcal{D}(U)$, then $U \sim V$. Thus if $\mathcal{D}(U) \cap \mathcal{D}(V) \neq \emptyset$, then $U \sim V$. Conversely, suppose $U \sim V$, say $\lambda U = V$, $\lambda \in K^\times$. The relation $\lambda \rho = \sigma$ establishes a one-to-one correspondence between the convergents $\rho$ of $U$ and the convergents $\sigma$ of $V$ and $\rho^{-1}U = \sigma^{-1}V$. Thus $\mathcal{D}(U) = \mathcal{D}(V)$. In particular, given modules $U$ and $V$, either $\mathcal{D}(U) = \mathcal{D}(V)$ or $\mathcal{D}(U) \cap \mathcal{D}(V) = \emptyset$ according as $U \sim V$ or not.

(15) By an argument similar to the one given in [2], we can show that $\mathcal{D}(U)$ is a finite set for any module $U$ and such an argument indicates how to find all members
In this paper we shall accomplish this by means of a continued fraction algorithm, which is more efficient.

Given a module \( U \), let \( O_U \) denote its coefficient ring; \( O_U \) consists of \( \omega \in K \) such that \( \omega \xi \in U \) for all \( \xi \in U \). \( O_U \) is a module and \( O \subset O_U \subset O_K \), where \( O_K \) is the ring of integers of \( K \). If \( U \sim U \), then \( O_U = O_V \). Given \( \lambda \in K^\times \), \( \lambda U = U \) iff \( \lambda \in O_U^\times \), then as \( \rho \) ranges over the convergents of \( U \), so does \( \lambda \rho \).

Given convergents \( \rho \) and \( \sigma \) of \( U \), call \( \rho \sim \sigma \) if \( \rho = \sigma \) for some \( \lambda \in O_U^\times \). A root of unity in \( O_U^\times \) is usually a 4th root of unity, i.e., in \( O^\times \), but it could be an 8th root or a 12th root of unity. We assume that given convergents \( \rho \) and \( \sigma \) of \( U \), recognizing if \( \rho \sim \sigma \) is instantaneous. If \( \rho \sim \sigma \), then \( |\rho| = |\sigma| \) (and hence \( |\rho'| = |\sigma'| \)). But since \( \rho \) and \( \sigma \) are not necessarily integers, it is possible that \( \rho \neq \sigma \) and \( |\rho| = |\sigma| \) (cf. Example 3 and (35)).

Our main objective in the rest of the paper is to show that \( C(U) \) is finite and to see how we can systematically obtain a complete set of representatives of the equivalence classes in \( C(U) \). We are going to develop a continued fraction algorithm for these purposes. We start with a simplest version. Such an algorithm has an independent interest of its own (cf. [4, pp. 181–188]).

Given \( \alpha \in C \), let \( [\alpha] \) denote the element \( \rho \in O \) such that \( \alpha - \rho \) is in the square

\[
-\frac{1}{2} < x < \frac{1}{2} \quad \text{and} \quad -\frac{1}{2} < y < \frac{1}{2}
\]

of the complex plane. Given \( \alpha \in C \), put \( \alpha_0 = \alpha \) and having defined \( \alpha_n \) for some \( n \geq 0 \), put \( \rho_n = [\alpha_n] \) and \( \alpha_{n+1} = 1/(\alpha_n - \rho_n) \) provided \( \alpha_n \neq \rho_n \), i.e., \( \alpha_n \notin O \). Note that \( |\alpha_n| \geq \sqrt{2} \) for \( n > 0 \). It is easily verified that \( \alpha_n \in O \) for some \( n \geq 0 \) iff \( \alpha \in F \).

Given \( \alpha \in C \), let \( p_n \) be as in (19) and put

\[
P_n = \begin{pmatrix} p_n & 1 \\ 1 & 0 \end{pmatrix}, \quad A_0 = I \quad \text{and} \quad A_n = P_0 P_1 \cdots P_{n-1}.
\]

Then we verify that

\[
A_n = \begin{pmatrix} a_n & a_{n-1} \\ b_n & b_{n-1} \end{pmatrix},
\]

where \( a_n \) and \( b_n \) are given by the recursions \( a_0 = 1, a_1 = p_0, a_{n+1} = a_n p_n + a_{n-1}, b_0 = 0, b_1 = 1, b_{n+1} = b_n p_n + b_{n-1} \). Since \( \det P_n = -1 \), \( \det A_n = (-1)^n \). In particular, \( (a_n, b_n) = 1 \).

From the definition of \( \alpha_n \) and \( p_n \) in (19), we have

\[
P_n^{-1} \alpha_n = \alpha_{n+1} \quad \text{and} \quad P_n^{-1} \begin{pmatrix} \alpha_n \\ 1 \end{pmatrix} = \alpha_n^{-1} \begin{pmatrix} \alpha_{n+1} \\ 1 \end{pmatrix}.
\]

Thus

\[
A_n^{-1} \alpha = \alpha_n \quad \text{and} \quad A_n^{-1} \begin{pmatrix} \alpha \\ 1 \end{pmatrix} = (\alpha_1 \cdots \alpha_n)^{-1} \begin{pmatrix} \alpha_n \\ 1 \end{pmatrix}.
\]

By looking at the second component of the second equality above, we get that \( a_n - b_n \alpha = (-1)^n / (\alpha_1 \cdots \alpha_n) \). Since \( |\alpha_n| \geq \sqrt{2} \) for \( n > 0 \), it follows that \( |a_n - b_n \alpha| \leq
1/\sqrt{2\pi}. In particular, for \( \alpha \notin F \),

\[
\lim_{n \to \infty} \frac{a_n}{b_n} = \alpha \quad \text{and} \quad \lim_{n \to \infty} b_n = \infty.
\]

(22) LEMMA. With \( \alpha_n \) and \( b_n \) as above for \( \alpha \notin F \), \( |b_n/(\alpha_1 \cdots \alpha_n)| < \sqrt{2} + 1 \) for all \( n > 0 \).

PROOF. By looking at the second component of the equality

\[
(\alpha_1 \cdots \alpha_n) \begin{pmatrix} \alpha \\ 1 \end{pmatrix} = A_n \begin{pmatrix} \alpha_n \\ 1 \end{pmatrix},
\]

we get that \( \alpha_1 \cdots \alpha_n = b_n \alpha_n + b_{n-1} \), and hence

\[
|b_n/(\alpha_1 \cdots \alpha_n)| = |\alpha_n + b_{n-1}/b_n|^{-1}.
\]

Call \( n > 0 \) good if \( |b_n| > |b_{n-1}| \). Since \( b_0 = 0 \) and \( b_1 = 1 \), \( n = 1 \) is good. (It can be shown that all \( n > 0 \) are good but we do not need this. In any case since \( b_n \to \infty \), there are infinitely many good \( n \)'s.) If \( n \) is good, then

\[
|\alpha_n + b_{n-1}/b_n| > |\alpha_n| - |b_{n-1}/b_n| > \sqrt{2} - 1,
\]

and hence

\[
|b_n/(\alpha_1 \cdots \alpha_n)| < 1/((\sqrt{2} - 1) = \sqrt{2} + 1.
\]

Suppose \( n \) is bad. Take the largest good \( k < n \). Then since \( n, n-1, \ldots, k+1 \) are bad, \( |b_n| \leq |b_{n-1}| \leq \cdots \leq |b_k| \), and hence

\[
\left| \frac{b_n}{\alpha_1 \cdots \alpha_n} \right| = \left| \frac{b_k}{\alpha_1 \cdots \alpha_k} \right| \left| \frac{b_n}{b_k \alpha_{k+1} \cdots \alpha_n} \right| < \left( \sqrt{2} + 1 \right) \left| \frac{b_n}{b_k} \right| \leq \sqrt{2} + 1.
\]

(23) THEOREM (PERIODICITY). Given \( \alpha \in C - F \), \( \alpha_{k+l} = \alpha_k \) for some \( k \) and \( l \) with \( l > 0 \) iff \( \alpha \) is quadratic over \( F \).

PROOF. Suppose \( \alpha_{k+l} = \alpha_k \), \( l > 0 \). Then with \( A = P_k \cdots P_{k+l-1} = A_k^{-1} A_{k+l} \), \( \alpha_k = A \alpha_{k+l} = A \alpha_k \). Thus \( \alpha_k \) is quadratic over \( F \). Since \( \alpha = A \alpha_k \), \( \alpha \) is quadratic over \( F \) also.

Conversely, suppose \( \alpha \) is quadratic over \( F \), say \( \delta^2 - e \alpha + c = 0 \), where \( d, e, c \in \mathcal{O} \) and \( dc \neq 0 \). Put \( C = \begin{pmatrix} e & -2c \\ 2d & e \end{pmatrix} \). Then \( C \alpha = \alpha \). Since \( \alpha = A \alpha_n \), \( A^{-1} CA \alpha_n = \alpha_n \).

Computing

\[
C_n = A_n^{-1} CA_n = \begin{pmatrix} e_n & -2c_n \\ 2d_n & -e_n \end{pmatrix}
\]

modulo \( \pm I \), we get that

\[
d_n = (2d_n - e) \alpha_n + c \text{ and } c_n = -d_n - 1.
\]

Put \( a_n = b_n \alpha + \delta_n \) and substitute this into the expression for \( d_n \) above. We get that \( d_n = (2d \alpha - e) \delta_n + d \delta_n^2 \). By (21) and (22), \( |b_n \delta_n| < \sqrt{2} + 1 \) and \( \delta_n \to 0 \) as \( n \to \infty \). Thus \( d_n \) are bounded by a constant (depending only on \( \alpha \)). Then so are \( c_n \). Since \( e_n^2 - 4d_n c_n = e^2 - 4dc, \) \( e_n \) are bounded also. Since \( d_n \alpha_n^2 - e_n \alpha_n + c_n = 0 \) and \( d_n, e_n, c_n \) are bounded, we conclude that there are only a finite number of distinct \( \alpha_n \). Thus \( \alpha_{k+l} = \alpha_k \) for some \( k \) and \( l > 0 \) (cf. [4, p. 185] for another proof).
(24) Let \( \alpha \) be quadratic over \( F \) and suppose \( \alpha_{k+l} = \alpha_k, l > 0 \). Then \( |\alpha_n - \alpha'_n| \geq \sqrt{2} - 1 \) for all \( n \geq k \).

**Proof.** Let the notations be as in (23) and put \( \Delta = e^2 - 4dc \). Then \( \alpha_n = (e_n + \sqrt{\Delta})/2d_n \) for all \( n \geq 0 \) (with \( d_0 = d, e_0 = e \) and \( c_0 = c \)) and hence \( \alpha_n - \alpha'_n = \sqrt{\Delta}/d_n \). Since \( 2d\alpha - e = \sqrt{\Delta}, d_n = \sqrt{\Delta}b_n\delta_n + d\delta_n^2 \) and \( \alpha_n = \alpha_{n+m} \) for all \( n \geq k \) and \( m \geq 0 \) and \( \delta_n \to 0 \) as \( n \to \infty \), we get that \( |d_n| \leq (\sqrt{2} + 1)/|\sqrt{\Delta}| \) for all \( n \geq k \) and hence

\[
|\alpha_n - \alpha'_n| = |\sqrt{\Delta}/d_n| \geq 1/(\sqrt{2} + 1) = \sqrt{2} - 1.
\]

(25) Given \( \alpha \in \mathbb{K} - F \), put \( U = \langle \alpha, 1 \rangle \). By means of the simple continued fraction algorithm developed above, we can find a unit \( \lambda \in \mathcal{O}_U^\times \) such that \( |\lambda| > 1 \). In fact, compute \( \alpha_n \) until we get \( \alpha_{k+l} \equiv \alpha_k, l > 0 \), and consider \( U_n = \langle \alpha_n, 1 \rangle \). Since \( A_n \in \text{GL}_2(\mathbb{O}) \) and

\[
A_n(\alpha_n^1) = (\alpha_1 \cdots \alpha_n)(\alpha_1^1),
\]

then put \( \lambda = \alpha_{k+1} \cdots \alpha_{k+l} \). Then \( |\lambda| > 1 \) and \( U_k = U_{k+l} = U_k \) and hence \( \lambda \in \mathcal{O}_U^\times \). (But \( \lambda \) may not be a fundamental unit of \( \mathcal{O}_U \).)

(26) Given a module \( U \), the norms (over \( F \)) of the convergents of \( U \) are bounded.

**Proof.** We may assume that \( U = \langle \alpha, 1 \rangle \). Let \( \rho \) be a convergent of \( U \). If \( |\rho| > 1 \), then take \( \lambda \in \mathcal{O}_U^\times \) such that \( |\lambda\rho| \leq 1 \) (cf. (25)). Then \( \sigma = \lambda\rho \) is a convergent of \( U \) such that \( |\sigma| \leq 1 \). Since \( |N\lambda| = 1, |N\sigma| = |N\rho| \), thus we may assume that \( |\rho| \leq 1 \).

Let \( \alpha_n \) and \( b_n \) be as in (20) for \( \alpha \) and consider the elements \( \xi_n = a_n - b_n\alpha \) of \( U \). Since \( \xi_n = (-1)^n(\alpha_1 \cdots \alpha_n)^{-1} \) (cf. (21)) and \( |\alpha_n| \geq \sqrt{2} \), \( \xi_n \) decreases to 0. Take \( n > 0 \) such that

\[
|\xi_n| < |\rho| \leq |\xi_{n-1}| = |\xi_n|/|\alpha_n|.
\]

Since \( \xi_n \in U \) and \( \rho \) is a convergent of \( U \), \( |\rho'| < |\xi'| \). Thus \( |N\rho| < |N\xi_n|/|\alpha_n| \). Since only a finite number of \( \alpha_n \) are distinct, \( |\alpha_n| \) are bounded. On the other hand,

\[
N\xi_n = (N\alpha_1 \cdots N\alpha_n)^{-1}.
\]

With the notations as in (23), \( d_n\alpha_n^2 - e_n\alpha_n + c_n = 0 \), and hence \( N\alpha_n = c_n/d_n \). Since \( c_n = -d_{n-1} \), \( N\alpha_1 \cdots N\alpha_n = (-1)^n d_n^{-1} \). Thus \( N\xi_n = (-1)^n d_n^{-1} \) and these are bounded. (A modification of the argument used in (11) of [2] gives another proof of this result via Minkowski Theorem.)

(27) Given \( c_2 > c_1 > 0 \), the number of convergents of \( U \) such that \( c_2 > |\rho| > c_1 \) is finite.

**Proof.** Choose \( c_0 > 0 \) such that \( |N\rho| < c_0 \) for all convergents \( \rho \) of \( U \) (cf. (26)). Let \( \rho \) be a convergent of \( U \) such that \( c_2 > |\rho| > c_1 \). Since \( |\rho|/|\rho'| < c_0, |\rho'| < c_0|\rho|^{-1} < c_0c_1^{-1} \). Since there are only a finite number of elements \( \xi \) of \( U \) such that \( |\xi| < c_2 \) and \( |\xi'| < c_0c_1^{-1} \) (cf. (12)), we get the result.

(28) **Theorem.** For every module \( U \), the set \( \mathcal{C}(U) \) is finite (cf. (17)).

**Proof.** Take \( \lambda \in \mathcal{O}_U^\times \) such that \( |\lambda| > 1 \). Given a convergent \( \rho \) of \( U \), take \( n \in \mathbb{Z} \) such that \( |\lambda|^{n-1} < |\rho| \leq |\lambda|^n \). Then \( |\lambda|^{-1} < |\rho\lambda^{-n}| \leq 1 \). Since \( \rho\lambda^{-n} \approx \rho \), we get the result by (27).

(29) We now turn to the problem of finding a complete set of representatives of the equivalence classes in \( \mathcal{C}(U) \), where \( U = \langle \alpha, 1 \rangle \). First of all, we have to find a convergent of \( U \) to get started. Let \( \alpha = (e + \sqrt{\Delta})/2b \in \mathbb{K} \), where \( b, e \in \mathbb{O} \). If \( \xi = y\alpha - x \in U \), then \( y = (b/\sqrt{\Delta})(\xi - \xi') \). Thus if \( |\xi| < 1 \) and \( |\xi'| < 1 \), then \( |y| < 2|b|/|\sqrt{\Delta}| \). In particular, if \( 2|b| \leq |\sqrt{\Delta}| \), then \( y = 0 \) and 1 must be a convergent of \( U \).
(30) Suppose $2|b|/|\sqrt{\Delta}| > 1$ but not too large (cf. (32)). Let $y$ range, in some convenient order, over the nonzero elements of $O$ in the first quadrant (including the real axis but not the imaginary axis) such that $|y| \leq 2|b|/|\sqrt{\Delta}|$. For each $y$, choose $x \in O$ such that $|y\alpha - x| \leq 1$ and compute $|y\alpha' - x|$. If $|y\alpha' - x| \geq 1$ for all $y$ and $x$, then 1 is a convergent of $U$.

(31) Assume that for some $y$ as in (30), there is an $x \in a$ such that $|y\alpha - x| < 1$ and $|y\alpha' - x| < 1$. For each $y$, compute $|y\alpha - x|$ and $|y\alpha' - x|$. If $|y\alpha' - x| \leq 1$ for all $y$ and $x$, then 1 is a convergent of $U$.

(32) In case $2|b|/|\sqrt{\Delta}|$ is large, the method of finding a convergent of $U$ described in (30) and (31) is tedious and unsatisfactory. This is where the result of (24) comes to the rescue. Compute $a_n$ as in (19) until we get $a_{k+l} = a_k$, $l > 0$, and consider $\beta = a_k$ and $V = \langle \beta, 1 \rangle$. With the notations as in (23) and (24),

$$2|d_k|/|\sqrt{\Delta}| = 2/|\beta - \beta'| \leq 2(\sqrt{2} + 1).$$

Thus we can find a convergent $\sigma$ of $V$ as in (29), (30) and (31) without much trouble. (It is likely that 1 is a convergent of $V$.) Let

$$A_k = \begin{pmatrix} a_k & a_{k-1} \\ b_k & b_{k-1} \end{pmatrix}$$

be as in (20) for $\alpha$ and put $\gamma = a_k - b_k\alpha$. Then $U = \langle \alpha, 1 \rangle = \langle b_{k-1}\alpha - a_{k-1}, \gamma \rangle = \gamma(\beta, 1) = \gamma V$. Thus $\gamma \sigma$ is a convergent of $U$.

(33) We now have a way to find a convergent $\rho_1 = p-q\alpha$ of $U = \langle \alpha, 1 \rangle$. If $\rho_1 = 1$, then put $Q_1 = I$. In any case, find $r$ and $s \in \mathbb{O}$ such that $ps - qr = 1$ and put $Q_1 = \begin{pmatrix} p & r \\ q & s \end{pmatrix}$. Although it does not matter how we find such $r$ and $s$, one definite way to find them is to apply the simple continued fraction algorithm to the “rational” element $\beta = p/q \in F^\times$. Compute $a_n$ and $b_n$ as in (20) for $\beta$. Then we arrive at $k \geq 0$ such that $qa_k - pb_k = \varepsilon \in O^\times$. Put $r = \varepsilon^{-1}a_k$ and $s = \varepsilon^{-1}b_k$. Put $\alpha_1 = Q_1^{-1}\alpha$ and $U_1 = \langle \alpha_1, 1 \rangle$. Since $Q_1^{-1}(\alpha_1) = \rho_1(\alpha_1)$,

$$Q_1(\begin{pmatrix} \alpha_1 \\ 1 \end{pmatrix}) = \rho_1^{-1}(\begin{pmatrix} \alpha \\ 1 \end{pmatrix})$$

and $U_1 = \rho_1^{-1}U$.

Since $\rho_1$ is a convergent of $U$, 1 is a convergent of $U_1$.

(34) LEMMA. If 1 is a convergent of $U = \langle \alpha, 1 \rangle$ and $\xi = y\alpha - x$ is a primitive element of $U$ such that $|\xi| \leq 1$ and $|y| \geq 2$, then $U$ contains a nonzero element $\beta$ such that $|\beta| < 1$ and $|\beta'| < |\xi'|$.

PROOF. Let $\xi = y\alpha - x$ be a primitive element of $U$ such that $|\xi| \leq 1$ and $|y| \geq 2$. Choose $p \in O$ such that $|\alpha - p| < 1$ and $|\alpha' - p|$ is least. Since 1 is a
convergent of \( U \), \(|a' - p| > 1\). Choose \( \varepsilon \in \mathcal{O}^\times \) such that \( \beta = \varepsilon(a - p) \) is in the first quadrant. Consider the half-planes

\[
H_1: \text{Re}(z) \leq \frac{1}{2}, \quad H_2: \text{Im}(z) \leq \frac{1}{2}, \quad H_3: \text{Re}(z) + \text{Im}(z) \leq 1.
\]

By the choice of \( p \) we have the following implications:

- If \(|\beta - 1| < 1\), then \( \beta' \in H_1 \).
- If \(|\beta - i| < 1\), then \( \beta' \in H_2 \).
- If \(|\beta - 1 - i| < 1\), then \( \beta' \in H_3 \).

Let \( A: |z| \geq 1 \). Since \( \beta \) is in the first quadrant and \(|\beta| < 1\), there are five cases:

1. \(|\beta - 1| < 1\) and \(|\beta - i| < 1\): put \( B = A \cap H_1 \cap H_2 \).
2. \(|\beta - 1| \geq 1\) and \(|\beta - 1 - i| < 1\): put \( B = A \cap H_2 \cap H_3 \).
3. \(|\beta - 1 - i| \geq 1\) and \(|\beta - 1 - i| \geq 1\): put \( B = A \cap H_2 \).
4. \(|\beta - i| \geq 1\) and \(|\beta - 1 - i| < 1\): put \( B = A \cap H_1 \cap H_3 \).
5. \(|\beta - i| \geq 1\) and \(|\beta - 1 - i| \geq 1\): put \( B = A \cap H_1 \).

Since \( \xi \) is primitive, \((x, y) = 1\). Since \(|\xi| \leq 1\), \(|\alpha - xy^{-1}| \leq |y|^{-1}\). Put \( r = |p - xy^{-1}| \).

The inequality

\[
|z - \varepsilon p| \geq |y| |z - \varepsilon xy^{-1}| = |yz - \varepsilon x|
\]

on \( z \) defines a disk \( D \) of radius \( r\sqrt{|y|/|y| - 1} \) with the center \( c \) on the line through \( \varepsilon p \) and \( \varepsilon xy^{-1} \) so that \( \varepsilon xy^{-1} \) is between \( \varepsilon p \) and \( c \) and \( |\varepsilon xy^{-1} - c| = r/(|y| - 1) \).

Since \(|y| \geq 2\), in any of the five cases above, if \( B \) is defined as indicated, then we see that \((B + \varepsilon p) \cap D = \emptyset\). Since \( \varepsilon a' \in B + \varepsilon p, \)

\[
|\beta'| = |\varepsilon a' - \varepsilon p| < |ya' - x| = |\xi'|.
\]

(35) Although it is possible for \( U \) to have two convergents \( \rho \) and \( \sigma \) such that \(|\rho| = |\sigma| \) and \( \rho \neq \sigma \), if \( \rho \), \( \sigma \) and \( \tau \) are convergents of \( U \) such that \(|\rho| = |\sigma| = |\tau|, \rho \neq \sigma \) and \( \rho \neq \tau \), then \( \sigma \equiv \tau \). In fact, by considering \( \rho^{-1} U \), we may assume that \( \rho = 1 \). Let \( \sigma \) and \( \tau \) be convergents of \( U \) such that \(|\sigma| = |\tau| = 1, \sigma \neq 1 \) and \( \tau \neq 1 \). Put \( \sigma = y\alpha - x \). If \(|y| \geq 2 \), then there is \( \beta \in U \) such that \( \beta \neq 0, |\beta| < 1 \) and \( |\beta'| < |\sigma'| = 1 \) by (34), which contradicts that 1 is a convergent of \( U \). Thus we may assume that \( \sigma = \alpha - x \) or \( \sigma = (1 + i)\alpha - x \). Similarly, we may assume that \( \tau = \alpha - y \) or \( \tau = (1 + i)\alpha - y \) for some \( y \in \mathcal{O} \). Since \(|\sigma| = |\tau| = 1, \) if \( \sigma \neq \tau \), then we get that \( \alpha \in \mathcal{O} \) or \( \alpha \in \mathcal{O} + \zeta \) for some 12th root of unity \( \zeta \). Since \( \alpha \notin \mathcal{O}, \) we get that \( \sigma \) and \( \tau \) are 12th roots of unity, and hence \( \sigma \equiv \tau \).

(36) Having chosen a convergent \( \rho_1 \) of \( U = \langle \alpha, 1 \rangle \), we are going to choose convergents \( \rho_2, \rho_3, \ldots \) of \( U \) so that \(|\rho_1| \geq |\rho_2| \geq |\rho_3| \geq \ldots, \rho_n \neq \rho_{n+1} \) for any \( n > 0 \), at most two \( \rho_n \)'s have the same modulus, and if \( \rho \) is a convergent of \( U \) such that \(|\rho_i| \geq |\rho_j| > |\rho_j| \) for some \( j > i > 0 \), then \( \rho \cong \rho_n \) for some \( n, \max\{1, i - 1\} \leq n < j \). (The possibility that \( \rho \cong \rho_{i-1} \) occurs only if \(|\rho_{i-1}| = |\rho_i| = |\rho| \) and \( i > 1 \).)

(37) Suppose we have found convergents \( \rho_1, \ldots, \rho_n \) of \( U = \langle \alpha, 1 \rangle \) satisfying the conditions stated in (36). We have done so for \( n = 1 \) (in which case the various conditions are vacuous). Moreover, assume that we have matrices \( A_1, \ldots, A_n \in \text{GL}_2(\mathcal{O}) \) of determinant 1 such that with \( A_n = \begin{pmatrix} a_n & c_n \\ b_n & d_n \end{pmatrix}, \rho_n = a_n - b_n \alpha \). For \( n = 1, A_1 = Q_1 \). (The meanings of \( a_n, b_n \) and \( A_n \) are now different from those in (20).)

Put \( \alpha_n = A_n^{-1} \alpha \) and \( U_n = \langle \alpha_n, 1 \rangle \). Since \( A_n^{-1} (\alpha) = \rho_n^{-1} \alpha \),

\[
A_n \begin{pmatrix} \alpha_n \\ 1 \end{pmatrix} = \rho_n^{-1} \begin{pmatrix} \alpha \\ 1 \end{pmatrix} \quad \text{and} \quad U_n = \rho_n^{-1} U.
\]
Since $\rho_n$ is a convergent of $U$, 1 is a convergent of $U_n$.  

(38) To find $\rho_{n+1}$, choose $p \in \mathcal{O}$ such that $|\alpha_n - p| \leq 1$ and $|\alpha'_n - p|$ is least. There are at most two choices for such $p$, and if so, choose the one of smaller modulus. If $n > 1$ and $|\rho_n^{-1}\rho_{n-1}| = 1$, then make sure that $|\alpha_n - p| < 1$. Choose $\varepsilon \in \mathcal{O}^\times$ such that 

$$|\alpha_n - p - \frac{\varepsilon}{1 + i}| \leq \frac{1}{\sqrt{2}}$$

and put 

$$\sigma_{n+1} = \begin{cases} 
  p - \alpha_n & \text{if } \sqrt{2}c \geq |\alpha'_n - p|, \\
  (1 + i)p + \varepsilon - (1 + i)\alpha_n & \text{if } \sqrt{2}c < |\alpha'_n - p|,
\end{cases}$$

and then put $\rho_{n+1} = \rho_n \sigma_{n+1}$.

(39) To see that $\rho_{n+1}$ produced in (38) is a next desired convergent of $U$, we claim that, if $\xi$ is a primitive element of $U_n$ such that $|\xi| \leq 1$ and $|\xi'| < |\sigma_{n+1}'|$, then $\xi \equiv 1$ or $\xi \equiv \rho_n^{-1}\rho_{n-1}$ (only if $n > 1$ and $|\rho_{n-1}| = |\rho_n|$). In fact, let $\xi$ be such an element, say $\xi = y_n - x$. If $|y| \geq 2$, then there is $\beta = \alpha_n - q \in U_n$ such that $|\beta| < 1$ and $|\beta'| < |\xi'|$ by (34). Since $|\xi'| < |\sigma_{n+1}'| \leq |\alpha'_n - p|$, this contradicts the choice of $p$ in (38). Thus $|y| < 2$, and we may assume that $\xi = \alpha_n - x$ or $\xi = (1 + i)\alpha_n - x$. First suppose that $\xi = \alpha_n - x$. If $|\xi| < 1$, then $|\xi'| \geq |\alpha'_n - p|$ by the choice of $p$. But since $|\alpha'_n - p| \geq |\sigma_{n+1}'|$, this is impossible. Thus $|\xi| = 1$. Suppose $\xi \not\equiv 1$. If $n = 1$ or $n > 1$ and $|\rho_{n-1}| > |\rho_n|$, then since $|\xi'| = 1$, $\xi \equiv \alpha_n - p$ by the choice of $p$ and $\sigma_{n+1} = \rho_n \alpha$. But since $|\xi'| < |\sigma_{n+1}'|$, this is impossible. Thus $n > 1$ and $|\rho_{n-1}| = |\rho_n|$. Then since $|\rho_{n-1}^{-1}\rho_{n-1}| = |\xi'| = 1$ and $\rho_n^{-1}\rho_{n-1} \not\equiv 1$, $\xi \equiv \rho_n^{-1}\rho_{n-1}$ by (35). On the other hand, if $\xi = (1 + i)\alpha_n - x$, then it contradicts the choice of $\varepsilon$ in (38). This proves the claim.

(40) Let $\xi$ be any element of $U_n$ such that $|\xi| < |\sigma_{n+1}|$ and $|\xi'| < |\sigma_{n+1}'|$. If $\xi \not\equiv 0$, then we may assume that $\xi$ is primitive. Since $|\sigma_{n+1}| \leq 1$, $|\xi| = 1$ by (39), which is absurd. Thus $\xi = 0$ and $\sigma_{n+1}$ is a convergent of $U_n$ and hence $\rho_{n+1}$ is a convergent of $U$. Clearly, $|\rho_n| \geq |\rho_{n+1}|$. Since $1 \not\equiv \sigma_{n+1}$, $\rho_n \not\equiv \rho_{n+1}$. If $n > 1$ and $|\rho_{n-1}| = |\rho_n|$, then $1 > |\sigma_{n+1}|$ and $|\rho_n| > |\rho_{n+1}|$. Let $\rho$ be a convergent of $U$ such that $|\rho_i| \geq |\rho| > |\rho_j|$, $0 < i < j \leq n + 1$. To see $\rho \equiv \rho_k$ for some $k$, max{1, $i - 1$} $\leq k < j$, we may assume that $|\rho_n| \geq |\rho| > |\rho_{n+1}|$. Then $\xi = \rho_n^{-1}\rho$ is a convergent of $U_n$ such that $1 \geq |\xi| > |\sigma_{n+1}|$. Since $\xi$ is a convergent, $|\xi'| < |\sigma_{n+1}'|$. Thus by (39), $\xi \equiv 1$ or $\xi \equiv \rho_n^{-1}\rho_{n-1}$, and hence $\rho \equiv \rho_n$ or $\rho \equiv \rho_{n-1}$. This completes the proof that $\rho_{n+1}$ is a next desired convergent of $U$.

(41) Put

$$Q_{n+1} = \begin{pmatrix} 
  p & -1 \\
  1 & 0 
\end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 
  p & r \\
  1 + i & 1 
\end{pmatrix}$$

according as $\sigma_{n+1} = p - \alpha_n$ or $p - (1 + i)\alpha_n$, where $r = (p - 1)/(1 + i)$. Note that in the second case, since $(p, 1 + i) = 1$, $1 + i$ divides $p - 1$ and $r \in \mathcal{O}$. In either case, det $Q_{n+1} = 1$. ($Q_{n+1}$ is rarely of the second type.) Put

$$A_{n+1} = A_n Q_{n+1} \quad \text{and} \quad \alpha_{n+1} = Q_{n+1}^{-1}\alpha_n = A_{n+1}^{-1}\alpha.$$ 

Since $Q_{n+1}^{-1}(\alpha_n^1) = \sigma_{n+1}(\alpha_{n+1}^1)$,

$$A_{n+1}^{-1}(\alpha^1) = Q_{n+1}^{-1} A_{n+1}^{-1}(\alpha^1) = \rho_n Q_{n+1}^{-1}(\alpha^1) = \rho_{n+1}(\alpha_{n+1}^1),$$

and hence $\rho_{n+1} = a_{n+1} - b_{n+1} \alpha$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
We now have a way to generate a sequence of convergents \( \rho_1, \rho_2, \rho_3, \ldots \) of \( U = \langle \alpha, 1 \rangle \) satisfying the conditions stated in (36). Moreover, corresponding to these convergents, we have matrices \( Q_1, Q_2, Q_3, \ldots \in \text{GL}_2(\mathcal{O}) \) of determinant 1 such that, for each \( n > 0 \), if
\[
A_n = \begin{pmatrix} a_n & c_n \\ b_n & d_n \end{pmatrix} = Q_1 \cdots Q_n,
\]
then \( \rho_n = a_n - b_n \alpha \) so that, with \( \alpha_n = A_n^{-1} \alpha \), \( A_n(\alpha_n) = \rho_n^{-1}(\alpha) \) and \( U_n = \langle \alpha_n, 1 \rangle = \rho_n^{-1}U \in \mathcal{D}(U) \).

Since \( \mathcal{D}(U) \) is finite by (28), \( U_{l+1} = U_k \) for some \( l \geq k > 0 \). Let \( l \) be the least such integer. Then we claim that \( k = 1 \). In fact, put \( \lambda = \rho_k \rho_{l+1}^{-1} \). Since \( \rho_k^{-1}U = U_k = U_{l+1} = \rho_{l+1}^{-1}U \), \( U = \lambda U \) and \( \lambda \in \mathcal{O}_U^\times \). Since \( |\rho_k| \geq |\rho_{l+1}| \), \( |\lambda| \geq 1 \). Since \( \lambda \in \mathcal{O}_U \), if \( |\lambda| = 1 \), then \( \lambda \) is a root of unity and \( \rho_k \cong \rho_{l+1} \). Thus \( |\lambda| > 1 \). Given a convergent \( \rho \) of \( U \), choose \( n \in \mathbb{Z} \) such that \( |\rho_k| |\lambda|^n > |\rho| > |\rho_k| |\lambda|^{-n} \).

Thus \( |\rho_k| \geq |\rho \lambda^{-n}| > |\rho \lambda^{-1}| = |\rho_{l+1}| \).

Theorem. If \( l \) is the period of \( U \), then \( \{\rho_1, \ldots, \rho_l\} \) is a complete set of representatives of the equivalence classes in \( \mathcal{C}(U) \) or equivalently
\[
\mathcal{D}(U) = \{U_1, \ldots, U_l\}, \quad U_n = \langle \alpha_n, 1 \rangle.
\]

Theorem. If \( l \) is the period of \( U \), then \( \lambda_0 = \rho_{l+1}^{-1} \rho_{l+1}^{-1} \) is a fundamental unit of \( \mathcal{O}_U^\times \), i.e., every \( \lambda \in \mathcal{O}_U^\times \) is uniquely of the form \( \lambda = \lambda_0^n \zeta \), where \( n \in \mathbb{Z} \) and \( \zeta \) is a root of unity.

Proof. Given \( \lambda \in \mathcal{O}_U^\times \), choose \( n \in \mathbb{Z} \) such that
\[
|\rho_1| \geq |\rho_1 \lambda \lambda_0^{-n}| > |\rho_1| |\lambda|^{-1} = |\rho_{l+1}|.
\]
Since \( \rho_1 \lambda \lambda_0^{-1} \cong \rho_1, \rho_1 \lambda \lambda_0^{-n} \cong \rho_1 \) and \( \lambda \lambda_0^{-n} = \zeta \) is a root of unity. The uniqueness is clear.
Here is a summary of the procedure for deciding if $A \sim B$ for given $A$ and $B \in M(f)$. First compute $\alpha = \phi(A)$ and $\alpha_n$ as in (33) and (42) until we get $\alpha_{l+1} = \alpha_1$ for the first time. Next compute $\beta = \phi(B)$ and $\beta_1$ for $\beta$ as in (33). Then $A \sim B$ iff $\beta_1 = \alpha_n$ for some $n, 1 \leq n \leq l$.

Given $A$ and $B \in M(f)$, suppose that $A \sim B$ so that $\beta_1 = \alpha_n$, $1 \leq n \leq l$, as in (47), say $\beta_1 = \varepsilon \alpha_n + c$, $\varepsilon \in \mathcal{O}_K^{\times}$, $c \in \mathcal{O}$. Compute $A_n$ for $\alpha = \phi(A)$ and $B_1 (= Q_1)$ for $\beta = \phi(B)$ and put

$$R_1 = B_1 \begin{pmatrix} \varepsilon & c \\ 0 & 1 \end{pmatrix} A_n^{-1}.$$ 

Then $R_1 \in GL_2(\mathcal{O})$ and $R_1 \alpha = \beta$, and hence $R_1 A R_1^{-1} = B$.

Given $A$, put $Z(A) = \{ R \in GL_2(\mathcal{O}) | RA = AR \}$, the centralizer of $A$ in $GL_2(\mathcal{O})$. $Z(A)$ is a subgroup of $GL_2(\mathcal{O})$. If $R_1 A R_1^{-1} = B$, then the coset $R_1 Z(A)$ consists of those $R \in GL_2(\mathcal{O})$ such that $RAR^{-1} = B$.

If $\alpha = \phi(A)$ and $U = \langle \alpha, 1 \rangle$, then $Z(A)$ is canonically isomorphic to $\mathcal{O}_K^{\times}$.

**Proof.** Let $R \in Z(A)$. Since $R\alpha = \alpha$, $R(\alpha_1) = \lambda(\alpha_1)$ for some $\lambda \in K^{\times}$. Then $U = \langle \alpha, 1 \rangle = \lambda \alpha \lambda$ and hence $\lambda \in \mathcal{O}_K^{\times}$. This defines a map $R \mapsto \lambda: Z(A) \rightarrow \mathcal{O}_K^{\times}$, and it is clear that it is a homomorphism. Suppose $\lambda = 1$ for the image $\lambda$ of $R \in Z(A)$. Then $R(\alpha \alpha') = (\alpha \alpha')$, and hence $R = I$. Thus the map is injective.

Let $\lambda \in \mathcal{O}_K^{\times}$. Then $U = \lambda U = \langle \lambda \alpha, \lambda \rangle$, and hence there is an $R \in GL_2(\mathcal{O})$ such that $R(\alpha_1) = \lambda(\alpha_1)$. Then $R\alpha = \alpha$ and $R \in Z(A)$. Thus the map $R \mapsto \lambda$ is onto $\mathcal{O}_K^{\times}$.

Let $l$ be the period of $\alpha = \phi(A)$, say $\alpha_{l+1} = \varepsilon \alpha_1 + c$, $\varepsilon \in \mathcal{O}_K^{\times}$, $c \in \mathcal{O}$. Put $R_0 = A_{l+1}(\begin{pmatrix} \varepsilon & c \\ 0 & 1 \end{pmatrix}) A_1^{-1}$. Then $R_0 \in GL_2(\mathcal{O})$ and

$$R_0 \begin{pmatrix} \alpha_1 \\ 1 \end{pmatrix} = \rho_1^{-1} A_{l+1} \begin{pmatrix} \varepsilon & c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ 1 \end{pmatrix} = \rho_1^{-1} A_{l+1} \begin{pmatrix} \alpha_{l+1} \\ 1 \end{pmatrix} = \rho_1^{-1} \rho_{l+1} \begin{pmatrix} \alpha \alpha' \\ 1 \end{pmatrix} = \lambda_0 \begin{pmatrix} \alpha \alpha' \\ 1 \end{pmatrix}.$$ 

Since $\lambda_0$, together with a root of unity $\zeta$, generates $\mathcal{O}_K^{\times}$ by (46), in view of (50), $R_0$ together with an element of order 4, 8 or 12 corresponding to $\zeta$, generates $Z(A)$.

(52) As a final remark, let us apply our method to find a fundamental unit of $\mathcal{O}_K$, $K = F(\sqrt{\Delta})$, where $\Delta \in \mathcal{O}$, $\Delta \neq 0, \pm 1$, $\Delta$ is square-free in $\mathcal{O}$. Put $\pi = 1 - i$ and

$$\alpha = \begin{cases} \sqrt{\Delta} & \text{if } \Delta \equiv 0 \pmod{\pi} \text{ or } \Delta \equiv \pm i \pmod{2}, \\ (1 + \sqrt{\Delta})/\pi & \text{if } \Delta \equiv \pm 1 + 2i \pmod{4}, \\ (1 + \sqrt{\Delta})/2 & \text{if } \Delta \equiv 1 \pmod{4}, \\ (1 + \sqrt{-\Delta})/2 & \text{if } \Delta \equiv -1 \pmod{4}. \end{cases}$$

Then $\mathcal{O}_K = \langle \alpha, 1 \rangle$. The proof of this is a straightforward exercise and is left to the reader. Clearly $\mathcal{O}_K$ is the coefficient ring of the module $\mathcal{O}_K$. Thus by finding the convergents of $\mathcal{O}_K$ we get a fundamental unit of $\mathcal{O}_K$ via (46).

**Appendix.** Reducible case.

(1) We shall summarize the results for the case when the characteristic polynomial $f$ is reducible over $F$. Since the proofs are straightforward, we shall omit them. Put $f(t) = (t - e_1)(t - e_2)$, where $e_1$ and $e_2 \in \mathcal{O}$.
(2) Given \( A \in M(f) \), we can find \( R \in GL_2(O) \) such that
\[
RAR^{-1} = \begin{pmatrix} e_1 & a \\ 0 & e_2 \end{pmatrix},
\]
where \( a \) is in the first quadrant (including the real axis but not the imaginary axis).

(3) Suppose \( e_1 = e_2 = e \). If \( a \) and \( b \) are in the first quadrant and \( (e_a) \sim (e_b) \), then \( a = b \).

(4) Let \( A = \begin{pmatrix} e_1 & a \\ 0 & e_2 \end{pmatrix} \). If \( a = 0 \), then \( Z(A) = GL_2(O) \). If \( a \neq 0 \), then \( Z(A) \) is generated by \( \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix} \), \( \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} \).

(5) Assume \( e_1 \neq e_2 \). Put \( e = e_1 - e_2 \). Given \( a \in O \), we can find \( R \in GL_2(O) \) such that
\[
R \begin{pmatrix} e_1 & a \\ 0 & e_2 \end{pmatrix} R^{-1} = \begin{pmatrix} e_1 & r \\ 0 & e_2 \end{pmatrix},
\]
where (i) \( r = 0 \), (ii) \( r/e = (1 + i)/2 \) or (iii) \( 0 < \text{Re}(r/e) \leq \frac{1}{2} \) and \( 0 \leq \text{Im}(r/e) < \frac{1}{2} \).

(6) If \( a/e \) and \( b/e \) are in the quarter square in the sense of (5) for \( r/e \) and \( (e_{a1} \ a2) \sim (e_{b1} \ b2) \), then \( a = b \).

(7) Let \( a/e \) be as in (6) and \( A = \begin{pmatrix} e_{a1} & a \\ 0 & e_{a2} \end{pmatrix} \). Then \( Z(A) \) is generated by
(i) \( \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} \), \( \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \) if \( a/e = 0 \),
(ii) \( \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} \), \( \begin{pmatrix} 0 & 1 \end{pmatrix} \) if \( a/e = 1/2 \),
(iii) \( \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} \), \( \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \) if \( a/e = (1 + i)/2 \),
(iv) \( \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} \) otherwise.

**Example 1.**
\[
A = \begin{pmatrix} 16 + 33i & 17 + 67i \\ 11 + 14i & 15 + 30i \end{pmatrix}, \quad B = \begin{pmatrix} 26 + 61i & 1 + 51i \\ 7 & 5 + 2i \end{pmatrix}.
\]
The characteristic polynomial of \( A \) and \( B \) is \( f(t) = t^2 - (31 + 63i)t + 1 \) and its discriminant is \( \Delta = -3012 + 3906i \).
\[
\alpha = \phi(A) = \frac{1 + 3i + \sqrt{\Delta}}{2(11 + 14i)}, \quad \beta = \phi(B) = \frac{21 + 59i + \sqrt{\Delta}}{2(7)}.
\]
Computing \( Q_n \) and \( \alpha_n = Q_{n-1}^{-1} \alpha_{n-1} \) for \( \alpha \), we get
\[
\begin{align*}
Q_1 &= I, & \alpha_1 &= \alpha, \\
Q_2 &= \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}, & \alpha_2 &= \frac{43 + 53i + \sqrt{\Delta}}{2(25 - 17i)}, \\
Q_3 &= \begin{pmatrix} 2i & -1 \\ 1 & 0 \end{pmatrix}, & \alpha_3 &= \frac{25 + 47i + \sqrt{\Delta}}{2(17 - 4i)}, \\
Q_4 &= \begin{pmatrix} 3i & -1 \\ 1 & 0 \end{pmatrix}, & \alpha_4 &= \frac{-1 + 55i + \sqrt{\Delta}}{2(13 - 56i)}.
\end{align*}
\]
\[
Q_5 = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}, \quad \alpha_5 = \frac{-25 + 57i + \sqrt{\Delta}}{2(29 - 5i)},
\]
\[
Q_6 = \begin{pmatrix} -1 + 2i & -1 \\ 1 & 0 \end{pmatrix}, \quad \alpha_6 = \frac{-13 + 69i + \sqrt{\Delta}}{2(-5 - 20i)},
\]
\[
Q_7 = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}, \quad \alpha_7 = \frac{23 + 31i + \sqrt{\Delta}}{2(11 + 14i)} \equiv \alpha_1.
\]

On the other hand, \(\beta_1 = \beta\) and this is not \(\equiv\) to any \(\alpha_n\). Thus \(A \not\sim B\). Note that computation gives
\[
Q_2 = \begin{pmatrix} 4 + 8i & -1 \\ 1 & 0 \end{pmatrix}, \quad \beta_2 = \frac{35 + 53i + \sqrt{\Delta}}{2(51 - 7i)}
\]
for \(\beta\) and the next convergent of \((\beta_2, 1)\) after 1 is \((1 + i)\beta_2 - i\) (cf. (38) and (41)).

**Example 2.**
\[
A = \begin{pmatrix} 16 + 33i & 17 + 67i \\ 11 + 14i & 15 + 30i \end{pmatrix}, \quad B = \begin{pmatrix} 72 + 85i & -5 - 29i \\ 176 - 7i & -41 - 22i \end{pmatrix}
\]
This \(A\) is the same as in Example 1 and the characteristic polynomial of \(B\) is the same as that of \(A\).
\[
\beta = \phi(B) = \frac{113 + 107i + \sqrt{\Delta}}{2(176 - 7i)}, \quad B_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \beta_1 = \frac{-113 - 107i + \sqrt{\Delta}}{2(5 + 29i)}
\]
for \(\beta\), and we recognize that \(\beta_1 \equiv \alpha_5\), in fact, \(\beta = -i\alpha_5 + (-3 + i)\). Thus \(A \sim B\).

Now compute (cf. (42))
\[
A_5 = Q_1Q_2Q_3Q_4Q_5 = \begin{pmatrix} 15 - i & 14 + 3i \\ 7 - 2i & 7 \end{pmatrix},
\]
\[
R_1 = B_1 \begin{pmatrix} -i & -3 + i \\ 0 & 1 \end{pmatrix} A_5^{-1} = \begin{pmatrix} 7 - 2i & -15 + i \\ 19 - 20i & -47 + 32i \end{pmatrix}.
\]
We have \(R_1AR_1^{-1} = B\). Noting \(\alpha_7 = \alpha_1 + 1\), we compute (cf. (51))
\[
A_7 = \begin{pmatrix} -16 - 33i & -1 - 34i \\ -11 - 14i & -4 - 16i \end{pmatrix},
\]
\[
R_0 = A_7 \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} A_1^{-1} = \begin{pmatrix} -16 - 33i & -17 - 67i \\ -11 - 14i & -15 - 30i \end{pmatrix} = -A.
\]
Thus \(Z(A)\) is generated by \(A\) and \(iI\) and we get all \(R \in \text{GL}_2(\mathcal{O})\) such that \(RAR^{-1} = B\). In view of (50), the eigenvalue of \(A\), \(\lambda = (31 + 63i + \sqrt{\Delta})/2\), is a fundamental unit of \(\mathcal{O}_U\), where \(U = (\alpha, 1)\).

**Example 3.**
\[
A = \begin{pmatrix} 1 + 4i & -5i \\ 2 + 4i & -3 - i \end{pmatrix},
\]
\[
f(t) = t^2 - (2 + 3i)t + (-19 - 3i), \quad \Delta = 71,
\]
\[ Q_1 = I, \quad \alpha_1 = \alpha = \frac{4 + 5i + \sqrt{\Delta}}{2(2 + 4i)} , \quad A_1 = I. \]

\( \sigma = \alpha_1 - 1 \) is a convergent of \( U_1 \) such that \( |\sigma| = 1 \) and \( \sigma \neq 1 \). Take \( \sigma_2 = \sigma \) (cf. (38)).

\[ Q_2 = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} , \quad \alpha_2 = \frac{3i + \sqrt{\Delta}}{2(-2 + 4i)} , \quad A_2 = Q_2. \]

\( \sigma = \alpha_2 \) is a convergent of \( U_2 \) such that \( |\sigma| = 1 \) and \( \sigma \neq 1 \). But since \( \sigma_2^{-1} = -\sigma \), \( \sigma_3 \neq \sigma \).

\[ Q_3 = \begin{pmatrix} -i & -1 \\ 1 & 0 \end{pmatrix} , \quad \alpha_3 = \frac{8 + i + \sqrt{\Delta}}{2} , \quad A_3 = \begin{pmatrix} -1 - i & -1 \\ -i & -1 \end{pmatrix} , \]

\[ Q_4 = \begin{pmatrix} 8 & -1 \\ 1 & 0 \end{pmatrix} , \quad \alpha_4 = \frac{8 - i + \sqrt{\Delta}}{2(-2 + 4i)} , \quad A_4 = \begin{pmatrix} -9 - 8i & 1 + i \\ -1 - 8i & 1 \end{pmatrix}. \]

\( \sigma = \alpha_4 - i \) is a convergent of \( U_4 \) such that \( |\sigma| = 1 \) and \( \sigma \neq 1 \). Take \( \sigma_5 = \sigma \).

\[ Q_5 = \begin{pmatrix} i & -1 \\ 1 & 0 \end{pmatrix} , \quad \alpha_5 = \frac{-3i + \sqrt{\Delta}}{2(2 - 4i)} , \quad A_5 = \begin{pmatrix} 9 - 8i & 9 + 8i \\ 8 & 1 + 8i \end{pmatrix}. \]

\( \sigma = \alpha_5 \) is a convergent of \( U_5 \) such that \( |\sigma| = 1 \) and \( \sigma \neq 1 \). But since \( \sigma_5^{-1} = -\sigma \), \( \sigma_6 \neq \sigma \),

\[ Q_6 = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} , \quad \alpha_6 = \frac{4 - 5i + \sqrt{\Delta}}{2(-5i)} , \quad A_6 = \begin{pmatrix} 18 & -9 + 8i \\ 9 + 8i & -8 \end{pmatrix} , \]

\[ Q_7 = \begin{pmatrix} i & -1 \\ 1 & 0 \end{pmatrix} , \quad \alpha_7 = \frac{6 + 5i + \sqrt{\Delta}}{2(-3 - 3i)} , \quad A_7 = \begin{pmatrix} -9 + 26i & -18 \\ -16 + 9i & -9 - 8i \end{pmatrix}. \]

\( \sigma = \alpha_7 + 1 \) is a convergent of \( U_7 \) such that \( |\sigma| = 1 \) and \( \sigma \neq 1 \). Take \( \sigma_8 = \sigma \).

\[ Q_8 = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} , \quad \alpha_8 = \frac{i + \sqrt{\Delta}}{2(3 - 3i)} , \quad A_8 = \begin{pmatrix} -9 - 26i & 9 - 26i \\ 7 - 17i & 16 - 9i \end{pmatrix}. \]

\( \sigma = \alpha_8 \) is a convergent of \( U_8 \) such that \( |\sigma| = 1 \) and \( \sigma \neq 1 \). But since \( \sigma_8^{-1} = -\sigma \), \( \sigma_9 \neq \sigma \),

\[ Q_9 = \begin{pmatrix} i & -1 \\ 1 & 0 \end{pmatrix} , \quad \alpha_9 = \frac{6 + 5i + \sqrt{\Delta}}{2(-5)} , \quad A_9 = \begin{pmatrix} 35 - 35i & 9 + 26i \\ 33 - 2i & -7 + 17i \end{pmatrix} , \]

\[ Q_{10} = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} , \quad \alpha_{10} = \frac{4 - 5i + \sqrt{\Delta}}{2(4 + 2i)} , \quad A_{10} = \begin{pmatrix} -26 + 61i & -35 + 35i \\ -40 + 19i & -33 + 2i \end{pmatrix}. \]

\( \sigma = \alpha_{10} + i \) is a convergent of \( U_{10} \) such that \( |\sigma| = 1 \) and \( \sigma \neq 1 \). Take \( \sigma_{11} = \sigma \).

\[ Q_{11} = \begin{pmatrix} -i & -1 \\ 1 & 0 \end{pmatrix} , \quad \alpha_{11} = \frac{-3i + \sqrt{\Delta}}{2(-4 + 2i)} , \quad A_{11} = \begin{pmatrix} 26 + 61i & 26 - 61i \\ -14 + 42i & 40 - 19i \end{pmatrix}. \]
We note that \( \alpha_{11} \equiv \alpha_1 \); \( \alpha_{11} = -i\alpha_1 + i \). Since \( |\sigma_{11}| = 1 \) and \( |\sigma_2| = 1 \), \( \sigma_2 \cong \sigma_{11}^{-1} \) (cf. (44)). Thus we take \( \sigma_{10} \sigma_{11} \) as \( \sigma_{10} \) and take \( Q_{10} Q_{11} \) as \( Q_{10} \):

\[
Q_{10} = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -i & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -1 + i & 1 \\ -1 & -1 \end{pmatrix}, \quad A_{10} = \begin{pmatrix} 26 + 61i & 26 - 61i \\ -14 - 42i & 40 - 19i \end{pmatrix}.
\]

Since \( \alpha_{10} = -i\alpha_1 + i \),

\[
R_0 = A_{10} \begin{pmatrix} -i & i \\ 0 & 1 \end{pmatrix} A_{10}^{-1} = \begin{pmatrix} 61 - 26i & -35 - 35i \\ 42 + 14i & -2 - 33i \end{pmatrix},
\]

and \( Z(A) \) is generated by \( R_0 \) and \( iI \).

**Example 4.** \( \Delta = 71 \), \( K = F(\sqrt{71}) \). Since \( 71 \equiv -1 \) (mod 4), with \( \alpha = (1 + \sqrt{-71})/2 \), \( \mathcal{O}_K = \langle \alpha, 1 \rangle \) (cf. (52)). Compute \( Q_n, \alpha_n \) and \( A_n \). (In this computation, we encounter convergents \( \sigma \) of some \( U_n \) such that \( |\sigma| = 1 \) and \( \sigma \not\equiv 1 \).)

\[
Q_1 = I, \quad \alpha_1 = \alpha, \quad A_1 = I,
\]

\[
Q_2 = \begin{pmatrix} 4i - 1 \\ 1 & 0 \end{pmatrix}, \quad \alpha_2 = \frac{-1 + 8i + \sqrt{-71}}{2(2 - 4i)}, \quad A_2 = Q_2,
\]

\[
Q_3 = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}, \quad \alpha_3 = \frac{-3 + \sqrt{-71}}{2(2 + 4i)}, \quad A_3 = \begin{pmatrix} -1 - 4i & -4i \\ 1 & -1 \end{pmatrix},
\]

\[
Q_4 = \begin{pmatrix} i & -1 \\ 1 & 0 \end{pmatrix}, \quad \alpha_4 = \frac{-1 - 4i + \sqrt{-71}}{2(-5i)}, \quad A_4 = \begin{pmatrix} -1 - i & 1 + i \\ -1 & 1 \end{pmatrix},
\]

\[
Q_5 = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}, \quad \alpha_5 = \frac{5 + 6i + \sqrt{-71}}{2(-3 + 3i)}, \quad A_5 = \begin{pmatrix} -3 + 9i & -4 + 5i \\ 2 + i & 1 + i \end{pmatrix},
\]

\[
Q_6 = \begin{pmatrix} -i & 1 \\ 1 & 0 \end{pmatrix}, \quad \alpha_6 = \frac{1 + \sqrt{-71}}{2(-3 - 3i)}, \quad A_6 = \begin{pmatrix} 5 + 8i & 3 - 9i \\ 2 - i & 2 - i \end{pmatrix},
\]

\[
Q_7 = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}, \quad \alpha_7 = \frac{5 + 6i + \sqrt{-71}}{2(-5)}, \quad A_7 = \begin{pmatrix} -2 - 17i & -5 - 8i \\ -4 & -2 + i \end{pmatrix},
\]

\[
Q_8 = \begin{pmatrix} -i & 1 \\ 1 & 0 \end{pmatrix}, \quad \alpha_8 = \frac{-5 + 4i + \sqrt{-71}}{2(-4 + 2i)}, \quad A_8 = \begin{pmatrix} -22 - 6i & 2 + 17i \\ -2 + 5i & 4 \end{pmatrix},
\]

\[
Q_9 = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}, \quad \alpha_9 = \frac{-3 + \sqrt{-71}}{2(-4 - 2i)}, \quad A_9 = \begin{pmatrix} -20 + 11i & 22 + 6i \\ 2 + 5i & 2 - 5i \end{pmatrix},
\]

\[
Q_{10} = \begin{pmatrix} -i & 1 \\ 1 & 0 \end{pmatrix}, \quad \alpha_{10} = \frac{-1 + 8i + \sqrt{-71}}{2(i)}, \quad A_{10} = \begin{pmatrix} 33 + 26i & 20 - 11i \\ 7 - 7i & -2 - 5i \end{pmatrix},
\]

\( \alpha_{10} \equiv \alpha_1; \alpha_{10} = -i\alpha_1 + (4 + i)\). Since \( \rho_1 = 1 \), \( \rho_{10} = (33 + 26i) - (7 - 7i)\alpha \) is a fundamental unit of \( \mathcal{O}_K \) (cf. (46)).
REFERENCES


DEPARTMENT OF MATHEMATICS, CITY COLLEGE OF NEW YORK, NEW YORK, NEW YORK 10031