IDEAL CLASSES AND MATRIX CONJUGATION OVER Z

KEITH CONRAD

1. Introduction

When R is a commutative ring, matrices A and B in $\mathrm{M}_n(R)$ are called conjugate when $UAU^{-1} = B$ for some $U \in \mathrm{GL}_n(R)$. The conjugacy problem in $\mathrm{M}_n(R)$ is: decide when two matrices in $\mathrm{M}_n(R)$ are conjugate. We want to look at the conjugacy problem in $\mathrm{M}_n(\mathbf{Z})$, where ideal theory and class groups make an interesting appearance.

The most basic invariant for conjugacy classes of matrices is the characteristic polynomial: conjugate matrices have the same characteristic polynomial. This is not a complete invariant in general: the matrices $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, both have characteristic polynomial $(T-1)^2$, but $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ are not conjugate (in $\mathrm{M}_2(R)$ for arbitrary R) since the identity matrix is conjugate only to itself. While there are refinements of the characteristic polynomial that settle the conjugacy problem in $\mathrm{M}_n(F)$ for F a field (use the rational canonical form), we don't pursue that direction. Instead our starting point is a special case where the characteristic polynomial is a complete invariant.

Theorem 1.1. *Let F be a field and $f(T) \in F[T]$ be monic irreducible of degree $n \geq 1$.*

(a) *A matrix A in $\mathrm{M}_n(F)$ has characteristic polynomial $f(T)$ if and only if $f(A) = O$.*
(b) *All matrices in $\mathrm{M}_n(F)$ with characteristic polynomial $f(T)$ are conjugate in $\mathrm{M}_n(F)$.*

The key point here is irreducibility of the characteristic polynomial. If that assumption is dropped, the theorem breaks down completely (the matrices in $\mathrm{M}_n(F)$ sharing a common reducible characteristic polynomial are not all conjugate to each other in $\mathrm{M}_n(F)$).

Proof. (a) If A has characteristic polynomial $f(T)$ then $f(A) = O$ by the Cayley-Hamilton theorem. Conversely, suppose $f(A) = O$. Let $\chi(T)$ be the characteristic polynomial of A. We want to show $\chi(T) = f(T)$. Since $f(T)$ is irreducible in $F[T]$, $f(T)$ is the minimal polynomial of A in $F[T]$, so $f(T)|\chi(T)$ in $F[T]$ because $\chi(A) = O$. Since $f(T)$ and $\chi(T)$ are monic of the same degree, the divisibility relation forces equality.

(b) Suppose $A \in \mathrm{M}_n(F)$ has characteristic polynomial $f(T)$. Make F^n an $F[T]$-module by letting multiplication by T on F^n be the action of A: $g(T)v = g(A)v$ for all $v \in F^n$. We are going to show F^n as an $F[T]$-module in this way is isomorphic to $F[T]/(f(T))$ as an $F[T]$-module. Therefore two matrices in $\mathrm{M}_n(F)$ with characteristic polynomial $f(T)$ give F^n isomorphic $F[T]$-module structures (it always looks like $F[T]/(f(T))$ as an $F[T]$-module), so the two matrices are conjugate because matrices in $\mathrm{M}_n(F)$ give rise to isomorphic $F[T]$-module structures on F^n if and only if they are conjugate.

Since $f(A) = O$, so $f(T)v = 0$ for all $v \in F^n$, the $F[T]$-module structure on F^n can be replaced with an $F[T]/(f(T))$-module structure: $\overline{g(T)}v = g(A)v$. The ring $F[T]/(f(T))$ is a field since $f(T)$ is irreducible, so F^n is an $F[T]/(f(T))$-vector space. Fixing a nonzero $v_0 \in F^n$, the multiples $F[T] \cdot v_0 = (F[T]/(f(T))) \cdot v_0$ form a subspace of F^n with F-dimension $\dim_F(F[T]/(f(T))) = n$, so it fills up all of F^n: $F^n = (F[T]/(f(T))) \cdot v_0$. Since

$v_0 \neq 0$ and $F[T]/(f(T))$ is a field, $F[T]/(f(T)) \cong F[T]/(f(T)) \cdot v_0$ as $F[T]$-modules, so $F^n \cong F[T]/(f(T))$ as $F[T]$-modules. □

Another key point in Theorem 1.1 besides irreducibility of the characteristic polynomial is that we are working over a field. If we work over $\mathbf{Z}$, irreducible characteristic polynomials don't necessarily provide a complete conjugacy invariant anymore. That is, two matrices in $\mathrm{M}_n(\mathbf{Z})$ can have a common irreducible characteristic polynomial in $\mathbf{Z}[T]$ while not being conjugate in $\mathrm{M}_n(\mathbf{Z})$ (that is, they are not conjugate by a matrix in $\mathrm{GL}_n(\mathbf{Z})$), although they must be conjugate in $\mathrm{M}_n(\mathbf{Q})$.

Example 1.2. The integral matrices $A = \left(\begin{smallmatrix} 0 & 4 \\ 2 & 0 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 0 & 8 \\ 1 & 0 \end{smallmatrix}\right)$ both have characteristic polynomial $T^2 - 8$, which is irreducible in $\mathbf{Z}[T]$, but they are not conjugate in $\mathrm{M}_2(\mathbf{Z})$. We show this by contradiction. Assume $UAU^{-1} = B$ for some $U \in \mathrm{GL}_2(\mathbf{Z})$, so $\det U = \pm 1$. Write the conjugacy relation as $UA = BU$ and let $U = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Computing UA and BU shows $a = 2d$ and $b = 4c$, so $\det U = ad - bc = 2d^2 - 4c^2$, which can't be ± 1.

Since $T^2 - 8$ is irreducible in $\mathbf{Q}[T]$, Theorem 1.1 says A and B are conjugate in $\mathrm{M}_2(\mathbf{Q})$, and indeed $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1/2 \end{smallmatrix}\right) A \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1/2 \end{smallmatrix}\right)^{-1} = B$.

If two integral matrices have the same irreducible characteristic polynomial $f(T)$, what additional data is needed to decide if the matrices are conjugate over $\mathbf{Z}$? We'll see that this task turns out to be equivalent to determining the ideal classes in an order in a number field generated by a root of $f(T)$. Our main result is stated and proved in Section 2 and we'll work out many examples of it in Section 3.

2. The bijection between conjugacy classes and ideal classes

Let's recall some terminology. For an order $\mathcal{O}$ in a number field K, a *fractional $\mathcal{O}$-ideal* is a nonzero finitely generated $\mathcal{O}$-module in K. We call two fractional $\mathcal{O}$-ideals I and J *equivalent* if $I = xJ$ for some $x \in K^\times$. The equivalence classes are called (fractional) $\mathcal{O}$-*ideal classes* and there are finitely many of them. When $\mathcal{O} = \mathcal{O}_K$, every fractional $\mathcal{O}$-ideal is invertible. When $\mathcal{O} \neq \mathcal{O}_K$ there are some noninvertible fractional $\mathcal{O}$-ideals. The label "ideal classes" here allows *all* fractional $\mathcal{O}$-ideals, invertible and noninvertible.

Theorem 2.1. *Let $f(T) \in \mathbf{Z}[T]$ be monic irreducible of degree $n \geq 1$.*

(a) *A matrix A in $\mathrm{M}_n(\mathbf{Z})$ has characteristic polynomial $f(T)$ if and only if $f(A) = O$.*
(b) *Conjugacy classes of matrices in $\mathrm{M}_n(\mathbf{Z})$ with characteristic polynomial $f(T)$ are in bijection with the $\mathbf{Z}[\alpha]$-ideal classes in $\mathbf{Q}(\alpha)$, where α is a root of $f(T)$.*

In particular, there are finitely many conjugacy classes of matrices in $\mathrm{M}_n(\mathbf{Z})$ with characteristic polynomial $f(T)$, since $\mathbf{Z}[\alpha]$ has finitely many ideal classes.

Theorem 2.1 is due to Latimer and MacDuffee [4]. See also [5, pp. 49–55], [6], and [7].

Example 2.2. Taking $f(T) = T^2 + 1$, all $A \in \mathrm{M}_2(\mathbf{Z})$ satisfying $A^2 + I_2 = O$ are conjugate in $\mathrm{M}_2(\mathbf{Z})$ since $\mathbf{Z}[i]$ has class number 1. Two such A are $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, and $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) = U\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)U^{-1}$ for $U = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbf{Z})$. (These two choices for A are *not* conjugate by a matrix in $\mathrm{SL}_2(\mathbf{Z})$.)

Example 2.3. Taking $f(T) = T^2 + 5$, all $A \in \mathrm{M}_2(\mathbf{Z})$ satisfying $A^2 + 5I_2 = O$ fall into two conjugacy classes since $\mathbf{Z}[\sqrt{-5}]$ has class number 2. We'll find representatives of these conjugacy classes in Example 3.1.

Example 2.4. The ring of integers of $\mathbf{Q}(\sqrt{10})$ is $\mathbf{Z}[\sqrt{10}]$, which has class number 2 and contains the unit $3+\sqrt{10}$. This unit is a root of T^2-6T-1. Since $\mathbf{Z}[\sqrt{10}] = \mathbf{Z}[3+\sqrt{10}]$, all $A \in \mathrm{M}_2(\mathbf{Z})$ satisfying $A^2 - 6A - I_2 = O$ fall into two conjugacy classes. We'll find representatives of these conjugacy classes in Example 3.2.

Example 2.5. Since $\mathbf{Z}[\sqrt[3]{2}]$ has class number 1, all $A \in \mathrm{M}_3(\mathbf{Z})$ satisfying $A^3 = 2I_3$ are conjugate to each other in $\mathrm{M}_3(\mathbf{Z})$.

Now we prove Theorem 2.1.

Proof. (a) The proof in Theorem 1.1(a) carries over, since a monic irreducible in $\mathbf{Z}[T]$ is irreducible in $\mathbf{Q}[T]$.

(b) For a $\mathbf{Z}[\alpha]$-fractional ideal $\mathfrak{a}$ in $\mathbf{Q}(\alpha)$, multiplication by α is a $\mathbf{Z}$-linear map $m_\alpha\colon \mathfrak{a} \to \mathfrak{a}$. Since $\mathfrak{a}$ as a $\mathbf{Z}$-module has a basis of size n, choosing a $\mathbf{Z}$-basis lets us represent m_α by a matrix $[m_\alpha] \in \mathrm{M}_n(\mathbf{Z})$. Changing the $\mathbf{Z}$-basis of $\mathfrak{a}$ changes the matrix representation of m_α to a conjugate matrix. So independent of a choice of basis we can associate to a fractional ideal $\mathfrak{a}$ the *conjugacy class* in $\mathrm{M}_n(\mathbf{Z})$ of a matrix representation for $m_\alpha\colon \mathfrak{a} \to \mathfrak{a}$. All matrices A in this conjugacy class satisfy $f(A) = O$ since $f(A) = f([m_\alpha]) = [m_{f(\alpha)}] = [m_0] = O$.

For an equivalent fractional $\mathbf{Z}[\alpha]$-ideal $\mathfrak{b} = x\mathfrak{a}$, where $x \in \mathbf{Q}(\alpha)^\times$, its conjugacy class of matrices (the matrices representing $m_\alpha\colon \mathfrak{b} \to \mathfrak{b}$ with respect to $\mathbf{Z}$-bases of $\mathfrak{b}$) is the same as that for $\mathfrak{a}$, since the matrix for m_α with respect to a $\mathbf{Z}$-basis $\{e_1, \ldots, e_n\}$ of $\mathfrak{a}$ is the same matrix as that for m_α with respect to the $\mathbf{Z}$-basis $\{xe_1, \ldots, xe_n\}$ of $\mathfrak{b}$. Thus we have a well-defined function

$$\text{(2.1)}\quad \mathbf{Z}[\alpha]\text{-ideal classes in } \mathbf{Q}(\alpha) \rightsquigarrow \text{conjugacy classes of } A \in \mathrm{M}_n(\mathbf{Z}) \text{ such that } f(A) = O$$

by the rule: pick a fractional ideal in an ideal class, pick a $\mathbf{Z}$-basis of it, write a matrix representation for m_α in terms of this basis, and use the conjugacy class of that matrix. We will show this function from fractional ideal classes to conjugacy classes of matrices is a bijection.

To show surjectivity, for every $A \in \mathrm{M}_n(\mathbf{Z})$ satisfying $f(A) = O$ we will find a $\mathbf{Z}[\alpha]$-fractional ideal $\mathfrak{a}$ in $\mathbf{Q}(\alpha)$ such that A is the matrix representation for $m_\alpha\colon \mathfrak{a} \to \mathfrak{a}$ with respect to some $\mathbf{Z}$-basis of $\mathfrak{a}$. Let $K = \mathbf{Q}(\alpha) = \mathbf{Q}[\alpha]$. Make $\mathbf{Q}^n$ into a K-vector space in the following way. For $\gamma \in K$, write $\gamma = g(\alpha)$ for $g(T) \in \mathbf{Q}[T]$. For all $v \in \mathbf{Q}^n$, set

$$\gamma v = g(\alpha)v := g(A)v \tag{2.2}$$

This is well-defined: if $\gamma = h(\alpha)$ for $h(T) \in \mathbf{Q}[T]$ then $g(\alpha) = h(\alpha)$, so $g(T) - h(T)$ is divisible by $f(T)$ (because f is the minimal polynomial of α in $\mathbf{Q}[T]$, as it is monic irreducible with root α) and therefore $g(A) = h(A)$ as matrices (since $f(A) = O$), so $g(A)v = h(A)v$ for all $v \in \mathbf{Q}^n$. If $v \in \mathbf{Z}^n$ then $\alpha v = Av$ is in $\mathbf{Z}^n$ since A has integral entries, so $\mathbf{Z}^n$ is a $\mathbf{Z}[\alpha]$-submodule of $\mathbf{Q}^n$ that is finitely generated since $\mathbf{Z}^n$ is already finitely generated as an abelian group. From the way we define $\mathbf{Q}^n$ as a K-vector space, the equation $\alpha v = Av$ tells us the matrix representation of m_α on $\mathbf{Z}^n$ with respect to the standard basis of $\mathbf{Z}^n$ is A.

Treating $\mathbf{Q}^n$ as both a $\mathbf{Q}$-vector space and as K-vector space (by (2.2)), we have

$$n = \dim_{\mathbf{Q}}(\mathbf{Q}^n) = [K : \mathbf{Q}] \dim_K(\mathbf{Q}^n) = n \dim_K(\mathbf{Q}^n),$$

so $\mathbf{Q}^n$ is 1-*dimensional* as a K-vector space. That means for a nonzero $v_0 \in \mathbf{Q}^n$, the K-linear map $\varphi_{v_0}\colon K \to \mathbf{Q}^n$ given by $\varphi_{v_0}(\gamma) = \gamma v_0$ is an isomorphism of 1-dimensional K-vector spaces. The inverse image $\varphi_{v_0}^{-1}(\mathbf{Z}^n)$ is a finitely generated $\mathbf{Z}[\alpha]$-submodule of K since $\mathbf{Z}^n$

has these properties inside the K-vector space $\mathbf{Q}^n$. So $\varphi_{v_0}^{-1}(\mathbf{Z}^n)$ is a fractional $\mathbf{Z}[\alpha]$-ideal in K. Call it $\mathfrak{a}$, so

$$\mathfrak{a} = \varphi_{v_0}^{-1}(\mathbf{Z}^n) = \{\gamma \in K : \gamma v_0 \in \mathbf{Z}^n\}.$$

(Changing v_0 may change $\mathfrak{a}$, but not its $\mathbf{Z}[\alpha]$-ideal class: for another nonzero $w_0 \in \mathbf{Q}^n$, $v_0 = xw_0$ for some $x \in K^\times$ since $\mathbf{Q}^n$ is 1-dimensional over K, and $\varphi_{w_0}^{-1}(\mathbf{Z}^n) = x\mathfrak{a}$, which is in the same $\mathbf{Z}[\alpha]$-ideal class as $\mathfrak{a}$.)

Since A is the matrix representation of m_α on $\mathbf{Z}^n$ with respect to its standard basis $\{e_1, \ldots, e_n\}$, A is also the matrix representation of m_α on $\mathfrak{a}$ with respect to the $\mathbf{Z}$-basis $\{\varphi_{v_0}^{-1}(e_1), \ldots, \varphi_{v_0}^{-1}(e_n)\}$ of $\mathfrak{a}$. We have realized A as a matrix representation for m_α on a fractional $\mathbf{Z}[\alpha]$-ideal $\mathfrak{a}$, so (2.1) is onto.

To show (2.1) is injective, suppose $\mathfrak{a}$ and $\mathfrak{b}$ are two fractional $\mathbf{Z}[\alpha]$-ideals in K such that the matrices A and B for m_α with respect to some $\mathbf{Z}$-bases of $\mathfrak{a}$ and $\mathfrak{b}$ are conjugate in $\mathrm{M}_n(\mathbf{Z})$. We want to show $\mathfrak{a}$ and $\mathfrak{b}$ are in the same ideal class: $\mathfrak{b} = x\mathfrak{a}$ for some $x \in K^\times$.

Since A represents $m_\alpha \colon \mathfrak{a} \to \mathfrak{a}$ with respect to some $\mathbf{Z}$-basis $\mathcal{A}$ of $\mathfrak{a}$, there is a commutative diagram

$$\begin{array}{ccc} \mathfrak{a} & \xrightarrow{[\cdot]_{\mathcal{A}}} & \mathbf{Z}^n \\ {\scriptstyle m_\alpha}\downarrow & & \downarrow{\scriptstyle A} \\ \mathfrak{a} & \xrightarrow[[\cdot]_{\mathcal{A}}]{} & \mathbf{Z}^n \end{array}$$

where the horitzonal arrows are the coordinate isomorphisms that identify $\mathcal{A}$ with the standard basis of $\mathbf{Z}^n$. Similarly for the basis $\mathcal{B}$ of $\mathfrak{b}$ with respect to which $m_\alpha \colon \mathfrak{b} \to \mathfrak{b}$ has matrix representation B, we have a commutative diagram

$$\begin{array}{ccc} \mathfrak{b} & \xrightarrow{[\cdot]_{\mathcal{B}}} & \mathbf{Z}^n \\ {\scriptstyle m_\alpha}\downarrow & & \downarrow{\scriptstyle B} \\ \mathfrak{b} & \xrightarrow[[\cdot]_{\mathcal{B}}]{} & \mathbf{Z}^n \end{array}.$$

Since A and B are conjugate in $\mathrm{M}_n(\mathbf{Z})$, $B = UAU^{-1}$ for some $U \in \mathrm{GL}_n(\mathbf{Z})$, so

$$\begin{array}{ccc} \mathbf{Z}^n & \xrightarrow{U} & \mathbf{Z}^n \\ {\scriptstyle A}\downarrow & & \downarrow{\scriptstyle B} \\ \mathbf{Z}^n & \xrightarrow[U]{} & \mathbf{Z}^n \end{array}$$

commutes. Let's put these diagrams together:

$$\begin{array}{ccccccc} \mathfrak{a} & \xrightarrow{[\cdot]_{\mathcal{A}}} & \mathbf{Z}^n & \xrightarrow{U} & \mathbf{Z}^n & \xrightarrow{[\cdot]_{\mathcal{B}}^{-1}} & \mathfrak{b} \\ {\scriptstyle m_\alpha}\downarrow & & {\scriptstyle A}\downarrow & & {\scriptstyle B}\downarrow & & \downarrow{\scriptstyle m_\alpha} \\ \mathfrak{a} & \xrightarrow[[\cdot]_{\mathcal{A}}]{} & \mathbf{Z}^n & \xrightarrow[U]{} & \mathbf{Z}^n & \xrightarrow[[\cdot]_{\mathcal{B}}^{-1}]{} & \mathfrak{b} \end{array}$$

Each square in the diagram commutes, so the whole diagram commutes. The top and bottom maps are $\mathbf{Z}$-linear isomorphisms, so the common composite map $\mathfrak{a} \to \mathfrak{b}$ on the top and bottom is a $\mathbf{Z}$-linear isomorphism that commutes with m_α by commutativity of the diagram around the boundary. That implies the composite map $\mathfrak{a} \to \mathfrak{b}$ is $\mathbf{Z}[\alpha]$-linear, not

just $\mathbf{Z}$-linear, so $\mathfrak{a}$ and $\mathfrak{b}$ are isomorphic as $\mathbf{Z}[\alpha]$-modules. Isomorphic fractional $\mathbf{Z}[\alpha]$-ideals are scalar multiples of each other, so $\mathfrak{b} = x\mathfrak{a}$ for some $x \in K^\times$. More specifically, each $\mathbf{Z}[\alpha]$-linear isomorphism of fractional $\mathbf{Z}[\alpha]$-ideals must be multiplication by some $x \in K^\times$, so the composite map $\mathfrak{a} \to \mathfrak{b}$ along the top and bottom of the above commutative diagram is multiplication by x. □

Remark 2.6. Here is a more conceptual version of the proof of part b.

The fractional $\mathbf{Z}[\alpha]$-ideals are the $\mathbf{Z}[\alpha]$-modules in $\mathbf{Q}(\alpha)$ that are free of rank n as $\mathbf{Z}$-modules. Moreover, an abstract $\mathbf{Z}[\alpha]$-module M whose underlying additive group is $\mathbf{Z}^n$ is isomorphic as a $\mathbf{Z}[\alpha]$-module to a fractional $\mathbf{Z}[\alpha]$-ideal. (Proof: Since $M \cong \mathbf{Z}^n$ as abelian groups, $\mathbf{Q} \otimes_{\mathbf{Z}} M \cong \mathbf{Q}^n$ as $\mathbf{Q}$-vector spaces. Since M is a $\mathbf{Z}[\alpha]$-module, $\mathbf{Q} \otimes_{\mathbf{Z}} M$ is a module over $\mathbf{Q} \otimes_{\mathbf{Z}} \mathbf{Z}[\alpha] \cong \mathbf{Q}(\alpha)$, and since $\dim_{\mathbf{Q}}(\mathbf{Q}(\alpha)) = n = \dim_{\mathbf{Q}}(\mathbf{Q} \otimes_{\mathbf{Z}} M)$, $\mathbf{Q} \otimes_{\mathbf{Z}} M$ has dimension 1 as a $\mathbf{Q}(\alpha)$-vector space. Using a vector space isomorphism of $\mathbf{Q} \otimes_{\mathbf{Z}} M$ with $\mathbf{Q}(\alpha)$ lets us identify the subset $1 \otimes M$ of $\mathbf{Q} \otimes_{\mathbf{Z}} M$ with a $\mathbf{Z}[\alpha]$-module in $\mathbf{Q}(\alpha)$ that's additively $\mathbf{Z}^n$, and this is a fractional $\mathbf{Z}[\alpha]$-ideal in $\mathbf{Q}(\alpha)$.) Fractional $\mathbf{Z}[\alpha]$-ideals are equivalent precisely when they are isomorphic as $\mathbf{Z}[\alpha]$-modules, so $\mathbf{Z}[\alpha]$-ideal classes in $\mathbf{Q}(\alpha)$ can be identified with isomorphism classes of $\mathbf{Z}[\alpha]$-modules that additively are $\mathbf{Z}^n$, or equivalently with isomorphism classes of $\mathbf{Z}[\alpha]$-module *structures* on $\mathbf{Z}^n$.

Next we show $\mathbf{Z}[\alpha]$-module structures on $\mathbf{Z}^n$ are in bijection with conjugacy classes of $A \in \mathrm{M}_n(\mathbf{Z})$ such that $f(A) = O$. To upgrade $\mathbf{Z}^n$ from a $\mathbf{Z}$-module to a $\mathbf{Z}[\alpha]$-module means making sense of multiplication by α on $\mathbf{Z}^n$, and that amounts to equipping $\mathbf{Z}^n$ with a $\mathbf{Z}$-linear map $A\colon \mathbf{Z}^n \to \mathbf{Z}^n$ such that $f(A) = O$: the action of A on $\mathbf{Z}^n$ is interpreted as multiplication by α. Two A such that $f(A) = O$ define isomorphic $\mathbf{Z}[\alpha]$-module structures on $\mathbf{Z}^n$ precisely when they are conjugate in $\mathrm{M}_n(\mathbf{Z})$[1], so $\mathbf{Z}[\alpha]$-module structures on $\mathbf{Z}^n$, up to isomorphism, can be identified with conjugacy classes of solutions to $f(A) = O$ in $\mathrm{M}_n(\mathbf{Z})$.

Remark 2.7. In Theorem 2.1 our matrix conjugations used the group $\mathrm{GL}_n(\mathbf{Z})$ of all invertible $n \times n$ integral matrices. The relation of conjugation by $\mathrm{SL}_n(\mathbf{Z})$ is, potentially, more restrictive: matrices that are conjugate by an element of $\mathrm{GL}_n(\mathbf{Z})$ might not be conjugate by an element of $\mathrm{SL}_n(\mathbf{Z})$. (This can only happen for even n, since the relation $B = UAU^{-1}$ implies $B = (-U)A(-U)^{-1}$ and for odd n, $\det(-U) = (-1)^n \det(U) = -\det(U)$. So if $U \in \mathrm{GL}_n(\mathbf{Z}) - \mathrm{SL}_n(\mathbf{Z})$ then $-U \in \mathrm{SL}_n(\mathbf{Z})$ when n is odd.) If we use $\mathrm{SL}_n(\mathbf{Z})$-conjugacy classes in Theorem 2.1b, then to maintain a bijection we need to make the equivalence relation on $\mathbf{Z}[\alpha]$-fractional ideals more restrictive: use $\mathfrak{a} \sim \mathfrak{b}$ when $\mathfrak{b} = x\mathfrak{a}$ for an $x \in K^\times$ such that $\mathrm{N}_{K/\mathbf{Q}}(x)$ is *positive*. The proof of Theorem 2.1b can be modified to show there is a bijection between $\mathrm{SL}_n(\mathbf{Z})$-conjugacy classes of $A \in \mathrm{M}_n(\mathbf{Z})$ satisfying $f(A) = O$ and equivalence classes of $\mathbf{Z}[\alpha]$-fractional ideals in $\mathbf{Q}(\alpha)$ where equivalent ideals are equal up to scaling by an element of $K^\times$ with positive norm.

Every monic polynomial of degree $n \geq 1$ is the characteristic polynomial of its companion matrix: if $f(T) = T^n + c_{n-1}T^{n-1} + \cdots + c_1T + c_0$, its companion matrix is

$$C_f = \begin{pmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{n-1} \end{pmatrix}$$

[1]This is similar to a vector space V over a field F having isomorphic $F[T]$-module structures from two linear operators A and B on V if and only if A and B are conjugate in $\mathrm{End}_F(V)$.

and $\det(TI_n - C_f) = f(T)$. When $f(T) \in \mathbf{Z}[T]$ is monic irreducible, with root α, which $\mathbf{Z}[\alpha]$-ideal class corresponds to the conjugacy class of C_f? This question has the following pretty answer.

Corollary 2.8. *Let $f(T) \in \mathbf{Z}[T]$ be monic irreducible of degree $n \geq 1$ with root α.*

(a) *The $\mathbf{Z}[\alpha]$-ideal class associated to the conjugacy class of C_f is the class of principal fractional $\mathbf{Z}[\alpha]$-ideals.*

(b) *A matrix $A \in \mathrm{M}_n(\mathbf{Z})$ with characteristic polynomial $f(T)$ is conjugate to C_f if and only if A represents multiplication by α on some principal fractional $\mathbf{Z}[\alpha]$-ideal, in which case A represents multiplication by α only on the fractional $\mathbf{Z}[\alpha]$-ideals that are principal.*

Part (b) answers a question posed to me by Lev Lerman, who was led to ask when A and C_f are conjugate from his work on dynamical systems.

Proof. (a) By the bijectivity of (2.1), to prove (a) it suffices to show there is a principal $\mathbf{Z}[\alpha]$-ideal on which multiplication by α with respect to some $\mathbf{Z}$-basis is represented by C_f. That ideal is simply $\mathbf{Z}[\alpha]$ itself: it has a $\mathbf{Z}$-basis $1, \alpha, \ldots, \alpha^{n-1}$, and the matrix for multiplication by α with respect to this basis is the companion matrix C_f.

(b) The first part of (b) (the part preceding "in which case") follows from (a) because (2.1) is a bijection and C_f represents multiplication by α on a $\mathbf{Z}$-basis of the principal $\mathbf{Z}[\alpha]$-ideal $\mathbf{Z}[\alpha]$. The second part of (b) follows from the first part because the proof of injectivity of (2.1) shows the fractional $\mathbf{Z}[\alpha]$-ideals on which A represents multiplication by α with respect to some $\mathbf{Z}$-basis are in the same $\mathbf{Z}[\alpha]$-ideal class. □

3. Examples

To illustrate Theorem 2.1 in examples, for some monic irreducible $f(T) \in \mathbf{Z}[T]$ we write down a set of ideals in $\mathbf{Z}[\alpha]$, where $f(\alpha) = 0$, representing the different $\mathbf{Z}[\alpha]$-ideal classes and compute the matrix for multiplication by α on each ideal with respect to a $\mathbf{Z}$-basis of that ideal. The resulting matrices are a complete set of representatives for the conjugacy classes of all $A \in \mathrm{M}_n(\mathbf{Z})$ satisfying $f(A) = O$, where $n = \deg f$. Most examples will have $\deg f = 2$.

Example 3.1. From Example 2.3, since $\mathbf{Z}[\sqrt{-5}]$ has class number 2 there are two conjugacy classes of matrices A in $\mathrm{M}_2(\mathbf{Z})$ satisfying $A^2 + 5I_2 = O$. To find matrices representing the two conjugacy classes, we use matrix representations for multiplication by $\alpha := \sqrt{-5}$ on ideals representing the two ideal classes of $\mathbf{Z}[\alpha]$.

The ideal classes of $\mathbf{Z}[\sqrt{-5}]$ are represented by (1) and $(2, 1 + \sqrt{-5})$ and to get $\mathbf{Z}$-bases of these ideals write $(1) = \mathbf{Z} + \sqrt{-5}\mathbf{Z}$ and $(2, 1 + \sqrt{-5}) = 2\mathbf{Z} + (1 + \sqrt{-5})\mathbf{Z}$. Multiplying the $\mathbf{Z}$-basis $\{1, \sqrt{-5}\}$ of (1) by $\sqrt{-5}$,

$$\begin{aligned} \sqrt{-5} \cdot 1 &= 0 \cdot 1 + 1 \cdot \sqrt{-5} \\ \sqrt{-5} \cdot \sqrt{-5} &= -5 \cdot 1 + 0 \cdot \sqrt{-5}, \end{aligned}$$

and multiplying the basis $\{2, 1 + \sqrt{-5}\}$ of $(2, 1 + \sqrt{-5})$ by $\sqrt{-5}$,

$$\begin{aligned} \sqrt{-5} \cdot 2 &= (-1) \cdot 2 + 2 \cdot (1 + \sqrt{-5}) \\ \sqrt{-5} \cdot (1 + \sqrt{-5}) &= (-3) \cdot 2 + 1 \cdot (1 + \sqrt{-5}). \end{aligned}$$

Therefore multiplication by $\sqrt{-5}$ on these two ideals is represented by the matrices $\left(\begin{smallmatrix} 0 & -5 \\ 1 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} -1 & -3 \\ 2 & 1 \end{smallmatrix}\right)$. That means each $A \in \mathrm{M}_2(\mathbf{Z})$ such that $A^2 + 5I_2 = O$ is conjugate to exactly one of those two matrices.

Another pair of ideals representing the two ideal classes is $(4+\sqrt{-5})$ and $(7, 2+3\sqrt{-5})$. Let's convert these into matrices. The first ideal has $\mathbf{Z}$-basis $\{4+\sqrt{-5}, -5+4\sqrt{-5}\}$, and

$$\begin{aligned} \sqrt{-5}\cdot(4+\sqrt{-5}) &= 0\cdot(4+\sqrt{-5}) + 1\cdot(-5+4\sqrt{-5}) \\ \sqrt{-5}\cdot(-5+4\sqrt{-5}) &= -5\cdot 1 + 0\cdot\sqrt{-5}, \end{aligned}$$

so multiplication by $\sqrt{-5}$ with respect to this basis is $\left(\begin{smallmatrix} 0 & -5 \\ 1 & 0 \end{smallmatrix}\right)$. Is it a coincidence we get the same matrix as we did for multiplication by $\sqrt{-5}$ on the ideal (1)? Not really. For a principal ideal $(a+b\sqrt{-5})$ the matrix for multiplication by $\sqrt{-5}$ with respect to the obvious first choice of $\mathbf{Z}$-basis – $\{a+b\sqrt{-5}, -5b+a\sqrt{-5}\}$ – is $\left(\begin{smallmatrix} 0 & -5 \\ 1 & 0 \end{smallmatrix}\right)$. Just compute it and see. So it's just a fluke of carrying out the computation on what happens to be the first basis that comes to mind. If you use a different basis of a principal ideal you would get a different matrix. Turning now to the second ideal $(7, 2+3\sqrt{-5})$, multiplying 7 and $2+3\sqrt{-5}$ by $\sqrt{-5}$ leads to the equations

$$\begin{aligned} \sqrt{-5}\cdot 7 &= -\frac{2}{3}\cdot 7 + 7\cdot(2+3\sqrt{-5}) \\ \sqrt{-5}\cdot(2+3\sqrt{-5}) &= -\frac{7}{3}\cdot 7 + \frac{2}{3}\cdot(2+3\sqrt{-5}). \end{aligned}$$

How come we got rational coefficients and not integral coefficients? Because the computation needs a $\mathbf{Z}$-basis of $(7, 2+3\sqrt{-5})$ but $\{7, 2+3\sqrt{-5}\}$ is *not* a $\mathbf{Z}$-basis (*e.g.*, $7\sqrt{-5}$ is in the ideal but it is not in the $\mathbf{Z}$-span of 7 and $2+3\sqrt{-5}$.) A $\mathbf{Z}$-basis of $(7, 2+3\sqrt{-5})$ is $\{7, 3+\sqrt{-5}\}$ (check!) and relative to this basis

$$\begin{aligned} \sqrt{-5}\cdot 7 &= -3\cdot 7 + 7\cdot(3+\sqrt{-5}) \\ \sqrt{-5}\cdot(3+\sqrt{-5}) &= -2\cdot 7 + 3\cdot(3+\sqrt{-5}), \end{aligned}$$

so the corresponding matrix is $\left(\begin{smallmatrix} -3 & -2 \\ 7 & 3 \end{smallmatrix}\right)$. Therefore $\left(\begin{smallmatrix} 0 & -5 \\ 1 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} -3 & -2 \\ 7 & 3 \end{smallmatrix}\right)$ are a pair of conjugacy class representatives of $A \in \mathrm{M}_2(\mathbf{Z})$ satisfying $A^2 + 5I_2 = O$.

Example 3.2. From Example 2.4, there are two conjugacy classes of matrices A in $\mathrm{M}_2(\mathbf{Z})$ satisfying $A^2 - 6A - I_2 = O$ since a root of $T^2 - 6T - 1$ is $\alpha = 3+\sqrt{10}$ and the ring $\mathbf{Z}[3+\sqrt{10}] = \mathbf{Z}[\sqrt{10}]$ has class number 2. Ideals representing the ideal classes are $(1) = \mathbf{Z} + \mathbf{Z}\sqrt{10}$ and $(2, \sqrt{10}) = \mathbf{Z}\cdot 2 + \mathbf{Z}\sqrt{10}$. A matrix representation for multiplication of each ideal by α can be found from

$$\begin{aligned} \alpha\cdot 1 &= 3\cdot 1 + 1\cdot\sqrt{10} \\ \alpha\cdot\sqrt{10} &= 10\cdot 1 + 3\cdot\sqrt{10}, \end{aligned}$$

$$\begin{aligned} \alpha\cdot 2 &= 3\cdot 2 + 2\cdot\sqrt{10} \\ \alpha\cdot\sqrt{10} &= 5\cdot 2 + 3\cdot\sqrt{10}, \end{aligned}$$

which tells us α has matrix representations $\left(\begin{smallmatrix} 3 & 10 \\ 1 & 3 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 3 & 5 \\ 2 & 3 \end{smallmatrix}\right)$: each $A \in \mathrm{M}_2(\mathbf{Z})$ satisfying $A^2 - 6A - I_2 = O$ is conjugate in $\mathrm{M}_2(\mathbf{Z})$ to exactly one of these two matrices.

If we use the $\mathbf{Z}$-bases $\{1, 3+\sqrt{10}\} = \{1, \alpha\}$ for the ideal (1) and $\{2, 2+\sqrt{10}\}$ for the ideal $(2, \sqrt{10})$, then we get a different pair of conjugacy class representatives. Multiplication by

α on (1) with respect to its new basis is $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 6 \end{smallmatrix}\right)$ since

$$\begin{aligned} \alpha \cdot 1 &= 0 \cdot 1 + 1 \cdot \alpha \\ \alpha \cdot \alpha &= 1 \cdot 1 + 6\alpha, \end{aligned}$$

and multiplication by α on $(2, \sqrt{10})$ with respect to its new basis is $\left(\begin{smallmatrix} 1 & 3 \\ 2 & 5 \end{smallmatrix}\right)$ since

$$\begin{aligned} \alpha \cdot 2 &= 1 \cdot 2 + 2(2 + \sqrt{10}) \\ \alpha \cdot (2 + \sqrt{10}) &= 3 \cdot 2 + 5(2 + \sqrt{10}), \end{aligned}$$

so each $A \in \mathrm{M}_2(\mathbf{Z})$ satisfying $A^2 - 6A - I_2 = O$ is conjugate in $\mathrm{M}_2(\mathbf{Z})$ to exactly one of $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 6 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} 1 & 3 \\ 2 & 5 \end{smallmatrix}\right)$.

In the integers of the number field $\mathbf{Q}(\sqrt{10}, \sqrt{2}) = \mathbf{Q}(\sqrt{2}, \sqrt{5})$, the ideal $(2, \sqrt{10})$ becomes principal, since there it equals $(\sqrt{2})(\sqrt{2}, \sqrt{5}) = (\sqrt{2})$. Corresponding to this, the matrices $\left(\begin{smallmatrix} 3 & 10 \\ 1 & 3 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 3 & 5 \\ 2 & 3 \end{smallmatrix}\right)$ become conjugate in $\mathrm{M}_2(\mathbf{Z}[\sqrt{2}])$: $U\left(\begin{smallmatrix} 3 & 10 \\ 1 & 3 \end{smallmatrix}\right)U^{-1} = \left(\begin{smallmatrix} 3 & 5 \\ 2 & 3 \end{smallmatrix}\right)$ where $U = \left(\begin{smallmatrix} \sqrt{2} & 5 \\ 1 & 2\sqrt{2} \end{smallmatrix}\right)$ has determinant -1. (More generally, $U\left(\begin{smallmatrix} 3 & 10 \\ 1 & 3 \end{smallmatrix}\right)U^{-1} = \left(\begin{smallmatrix} 3 & 5 \\ 2 & 3 \end{smallmatrix}\right)$ when $U = \left(\begin{smallmatrix} a & 5c \\ c & 2a \end{smallmatrix}\right)$, and for a suitable choice of a and c the determinant $2a^2 - 5c^2$ is -1.) For a broader viewpoint on this example, see [2].

Example 3.3. Let $f(T) = T^2 - T + 6 = 0$, with root $\alpha = \frac{1+\sqrt{-23}}{2}$. The ring $\mathbf{Z}[\alpha]$ has class number 3, with ideal classes represented by (1), $(2, \alpha)$, and $(2, \alpha)^2 = (4, 6 + \alpha)$. As $\mathbf{Z}$-modules, these ideals have respective $\mathbf{Z}$-bases $\{1, \alpha\}$, $\{2, \alpha\}$, and $\{4, 6 + \alpha\}$.

The matrices for multiplication by α on these ideals, with respect to the indicated $\mathbf{Z}$-bases of them, are found as follows:

$$\begin{aligned} \alpha \cdot 1 &= 0 \cdot 1 + 1 \cdot \alpha \\ \alpha \cdot \alpha &= -6 \cdot 1 + 1 \cdot \alpha, \end{aligned}$$

$$\begin{aligned} \alpha \cdot 2 &= 0 \cdot 1 + 2 \cdot \alpha \\ \alpha \cdot \alpha &= -3 \cdot 2 + 1 \cdot \alpha, \end{aligned}$$

$$\begin{aligned} \alpha \cdot 4 &= (-6) \cdot 4 + 4 \cdot (6 + \alpha) \\ \alpha \cdot (6 + \alpha) &= (-12) \cdot 4 + 7 \cdot (6 + \alpha), \end{aligned}$$

so the three conjugacy classes of $A \in \mathrm{M}_2(\mathbf{Z})$ satisfying $A^2 - A + 6I_2 = O$ are represented by $\left(\begin{smallmatrix} 0 & -6 \\ 1 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0 & -3 \\ 2 & 1 \end{smallmatrix}\right)$, and $\left(\begin{smallmatrix} -6 & -12 \\ 4 & 7 \end{smallmatrix}\right)$.

Example 3.4. Let d be a nonsquare in $\mathbf{Z}$ and $m \geq 2$. Then $f(T) = T^2 - m^2 d$ is irreducible in $\mathbf{Z}[T]$ with root $\alpha = m\sqrt{d}$. Set $\mathfrak{a} := \mathbf{Z}[\sqrt{d}] = \mathbf{Z} + \mathbf{Z}\sqrt{d}$ and $\mathfrak{b} := \mathbf{Z}[\alpha] = \mathbf{Z} + \mathbf{Z}m\sqrt{d}$. These are fractional $\mathbf{Z}[\alpha]$-ideals: $\mathfrak{b} = (1)$ is principal and $\mathfrak{a} = \frac{1}{m}(\mathbf{Z}m + \mathbf{Z}\alpha) = \frac{1}{m}(m, \alpha)$ is nonprincipal since (m, α) is a nonprincipal ideal in $\mathbf{Z}[\alpha]$.[2]

To find matrices A satisfying $f(A) = O$ that correspond to $\mathfrak{a}$ and $\mathfrak{b}$, we compute the matrix for multiplication by $\alpha = m\sqrt{d}$ with respect to $\mathbf{Z}$-bases of each of them. Using the

[2] If the ideal $(m, \alpha) = \mathbf{Z}m + \mathbf{Z}\alpha$ in $\mathbf{Z}[\alpha]$ were principal, say $(a + b\alpha)$ for $a, b \in \mathbf{Z}$, then $m \mid a$. As a rank 2 $\mathbf{Z}$-module this has bases $\{m, \alpha\}$ and $\{a + b\alpha, (a + b\alpha)\alpha\} = \{a + b\alpha, dm^2 b + a\alpha\}$. The transition matrix over $\mathbf{Z}$ from the first basis to the second has determinant $a^2/m - dmb^2 = a(a/m) - dmb^2$, which is a multiple of m since $m \mid a$, and thus this determinant can't be ± 1.

$\mathbf{Z}$-bases $\{1, \sqrt{d}\}$ for $\mathfrak{a}$ and $\{1, m\sqrt{d}\}$ for $\mathfrak{b}$,

$$\begin{aligned} m\sqrt{d} \cdot 1 &= 0 \cdot 1 + m \cdot \sqrt{d} \\ m\sqrt{d} \cdot \sqrt{d} &= md \cdot 1 + 0 \cdot \sqrt{d}, \end{aligned}$$

and

$$\begin{aligned} m\sqrt{d} \cdot 1 &= 0 \cdot 1 + 1 \cdot m\sqrt{d} \\ m\sqrt{d} \cdot m\sqrt{d} &= m^2 d \cdot 1 + 0 \cdot m\sqrt{d}, \end{aligned}$$

so the matrices are $\left(\begin{smallmatrix} 0 & md \\ m & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & m^2 d \\ 1 & 0 \end{smallmatrix}\right)$. These both satisfy $A^2 = m^2 d I_2$ and they are nonconjugate in $\mathrm{M}_2(\mathbf{Z})$. Taking $d = 2$ and $m = 2$ recovers Example 1.2, where the nonconjugate matrices $\left(\begin{smallmatrix} 0 & 4 \\ 2 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 8 \\ 1 & 0 \end{smallmatrix}\right)$ can now be interpreted as multiplication by $\sqrt{8} = 2\sqrt{2}$ on $\mathbf{Z}[\sqrt{2}]$ with respect to the $\mathbf{Z}$-basis $\{1, \sqrt{2}\}$ and on $\mathbf{Z}[2\sqrt{2}]$ with respect to the $\mathbf{Z}$-basis $\{1, 2\sqrt{2}\}$. The matrix $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1/2 \end{smallmatrix}\right)$ in $\mathrm{M}_2(\mathbf{Q})$ that conjugates $\left(\begin{smallmatrix} 0 & 4 \\ 2 & 0 \end{smallmatrix}\right)$ to $\left(\begin{smallmatrix} 0 & 8 \\ 1 & 0 \end{smallmatrix}\right)$ can now be explained: it is the change of basis matrix from $\{1, 2\sqrt{2}\}$ to $\{1, \sqrt{2}\}$ as $\mathbf{Q}$-bases of $\mathbf{Q}(\sqrt{2})$.

Unlike in Example 3.2, where the two matrices that are not conjugate in $\mathrm{M}_2(\mathbf{Z})$ become conjugate in $\mathrm{M}_2(\mathbf{Z}[\sqrt{2}])$, $\left(\begin{smallmatrix} 0 & 4 \\ 2 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 8 \\ 1 & 0 \end{smallmatrix}\right)$ are not conjugate in $\mathrm{M}_2(\overline{\mathbf{Z}})$, where $\overline{\mathbf{Z}}$ is the ring of all algebraic integers. We can show this by contradiction. Assume $U\left(\begin{smallmatrix} 0 & 4 \\ 2 & 0 \end{smallmatrix}\right)U^{-1} = \left(\begin{smallmatrix} 0 & 8 \\ 1 & 0 \end{smallmatrix}\right)$ for some $U \in \mathrm{GL}_2(\overline{\mathbf{Z}})$. As in the computation at the end of Example 1.2, $U = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ where $a = 2d$ and $b = 4c$, so $\det U = ad - bc = 2d^2 - 4c^2 = 2(d^2 - 2c^2)$, which is not a unit in $\overline{\mathbf{Z}}$ for all c and d in $\overline{\mathbf{Z}}$.

Example 3.5. So far we have computed matrices from fractional ideals. Let's go the other way around. The matrix $A = \left(\begin{smallmatrix} 2 & 3 \\ -3 & -2 \end{smallmatrix}\right)$ satisfies $A^2 + 5I_2 = O$. We will convert A into a $\mathbf{Z}[\sqrt{-5}]$-fractional ideal in $\mathbf{Q}(\sqrt{-5})$ using the proof of the surjectivity in Theorem 2.1(b). Make $\mathbf{Q}^2$ into a $\mathbf{Q}(\sqrt{-5})$-vector space as follows: for $a, b \in \mathbf{Q}$ and $\binom{x}{y} \in \mathbf{Q}^2$, define

$$\begin{aligned} (a + b\sqrt{-5})\begin{pmatrix} x \\ y \end{pmatrix} &:= (aI_2 + bA)\begin{pmatrix} x \\ y \end{pmatrix} \\ &= \left(aI_2 + b\begin{pmatrix} 2 & 3 \\ -3 & -2 \end{pmatrix}\right)\begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} a + 2b & 3b \\ -3b & a - 2b \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

Set $v_0 = \binom{1}{0}$, so there is an isomorphism $\mathbf{Q}(\sqrt{-5}) \to \mathbf{Q}^2$ of $\mathbf{Q}(\sqrt{-5})$-vector spaces by $\gamma \mapsto \gamma v_0$. The fractional ideal we want is the inverse image of $\mathbf{Z}^2$ under this isomorphism $\mathbf{Q}(\sqrt{-5}) \to \mathbf{Q}^2$. This is $\{\gamma \in \mathbf{Q}(\sqrt{-5}) : \gamma v_0 \in \mathbf{Z}^2\}$. Writing $\gamma = a + b\sqrt{-5}$,

$$\gamma v_0 = (aI_2 + bA)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a + 2b \\ -3b \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & -3 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix},$$

so

$$\gamma v_0 \in \mathbf{Z}^2 \iff \begin{pmatrix} a \\ b \end{pmatrix} \in \begin{pmatrix} 1 & 2 \\ 0 & -3 \end{pmatrix}^{-1}\mathbf{Z}^2 = \begin{pmatrix} 1 & 2/3 \\ 0 & -1/3 \end{pmatrix}\mathbf{Z}^2 = \left\{\begin{pmatrix} x + (2/3)y \\ -(1/3)y \end{pmatrix} : x, y \in \mathbf{Z}\right\}.$$

Therefore having $\gamma v_0 \in \mathbf{Z}^2$ is the same as saying $\gamma = x + (2/3)y - (1/3)y\sqrt{-5}$ for some integers x and y, so the $\mathbf{Z}[\sqrt{-5}]$-fractional ideal in $\mathbf{Q}(\sqrt{-5})$ corresponding to A is

$$\left\{x + \left(\frac{2}{3} - \frac{1}{3}\sqrt{-5}\right)y : x, y \in \mathbf{Z}\right\} = \mathbf{Z} + \left(\frac{2 - \sqrt{-5}}{3}\right)\mathbf{Z}. \tag{3.1}$$

If we change $v_0 = \binom{1}{0}$ to another nonzero vector in $\mathbf{Q}^2$, we get an equivalent fractional ideal. The ideal class of these fractional ideals is independent of the choice of v_0. Scaling a fractional ideal doesn't change its ideal class, so we multiply the fractional ideal by 3 and get the ideal $3\mathbf{Z} + (2-\sqrt{-5})\mathbf{Z} = (3, 2-\sqrt{-5})$ in $\mathbf{Z}[\sqrt{-5}]$.

As a check that our work is correct, let's compute the matrix for multiplication by $\sqrt{-5}$ on the ideal $(3, 2-\sqrt{-5})$ using the $\mathbf{Z}$-basis $\{3, 2-\sqrt{-5}\}$:

$$\begin{aligned} \sqrt{-5}\cdot 3 &= 2\cdot 3 - 3(2-\sqrt{-5}) \\ \sqrt{-5}\cdot(2-\sqrt{-5}) &= 3\cdot 3 - 2\cdot(2-\sqrt{-5}), \end{aligned}$$

so the matrix is $\left(\begin{smallmatrix} 2 & 3 \\ -3 & -2 \end{smallmatrix}\right) = A$.

Wait, there's more! In Example 3.1 we said each $A \in \mathrm{M}_2(\mathbf{Z})$ satisfying $A^2 + 5I_2 = O$ is conjugate to either $\left(\begin{smallmatrix} 0 & -5 \\ 1 & 0 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} -1 & -3 \\ 2 & 1 \end{smallmatrix}\right)$. Which of these is $\left(\begin{smallmatrix} 2 & 3 \\ -3 & -2 \end{smallmatrix}\right)$ conjugate to? We will answer this by turning it into a question about ideals. Table 3.5 summarizes the list of matrices and corresponding ideals and their $\mathbf{Z}$-bases with respect to which the matrix represents multiplication by $\sqrt{-5}$ on the ideal.

Matrix	Ideal	Basis
$\left(\begin{smallmatrix} 0 & -5 \\ 1 & 0 \end{smallmatrix}\right)$	(1)	$\{1, \sqrt{-5}\}$
$\left(\begin{smallmatrix} -1 & -3 \\ 2 & 1 \end{smallmatrix}\right)$	$(2, 1+\sqrt{-5})$	$\{2, 1+\sqrt{-5}\}$
$\left(\begin{smallmatrix} 2 & 3 \\ -3 & -2 \end{smallmatrix}\right)$	$(3, 2-\sqrt{-5})$	$\{3, 2-\sqrt{-5}\}$

The ideal $(3, 2-\sqrt{-5})$ is equivalent to (1) or $(2, 1+\sqrt{-5})$: it is principal or nonprincipal. We show $(3, 2-\sqrt{-5})$ is not principal by looking at the quotient ring

$$\mathbf{Z}[\sqrt{-5}]/(3, 2-\sqrt{-5}) \cong \mathbf{Z}[T]/(T^2+5, 3, 2-T) \cong \mathbf{Z}/3\mathbf{Z}.$$

If $(3, 2-\sqrt{-5}) = (\beta)$ is principal then $|\mathbf{Z}[\sqrt{-5}]/(\beta)| = 3$, so $\mathrm{N}_{\mathbf{Q}(\sqrt{-5})/\mathbf{Q}}(\beta) = 3$. But no element of $\mathbf{Z}[\sqrt{-5}]$ has norm 3. So $(3, 2-\sqrt{-5})$ is not principal, which makes it equivalent to $(2, 1+\sqrt{-5})$, so $\left(\begin{smallmatrix} 2 & 3 \\ -3 & -2 \end{smallmatrix}\right)$ is conjugate in $\mathrm{M}_2(\mathbf{Z})$ to $\left(\begin{smallmatrix} -1 & -3 \\ 2 & 1 \end{smallmatrix}\right)$.

Now it is natural to ask for an explicit conjugating matrix between $\left(\begin{smallmatrix} 2 & 3 \\ -3 & -2 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} -1 & -3 \\ 2 & 1 \end{smallmatrix}\right)$. To find one, we will find an explicit scaling factor between the ideals $(3, 2-\sqrt{-5})$ and $(2, 1+\sqrt{-5})$. There is an $x \in \mathbf{Q}(\sqrt{-5})$ such that $(3, 2-\sqrt{-5}) = x(2, 1+\sqrt{-5})$. To find x, multiply both sides by $(2, 1+\sqrt{-5})$. The right side becomes $x(2) = (2x)$ and the left side becomes

$$(3, 2-\sqrt{-5})(2, 1+\sqrt{-5}) = (6, 3+3\sqrt{-5}, 4-2\sqrt{-5}, 7+\sqrt{-5}).$$

We can eliminate the middle two generators since $3+3\sqrt{-5} = (-3)\cdot 6 + 3(7+\sqrt{-5})$ and $4-2\sqrt{-5} = 3\cdot 6 - 2(7+\sqrt{-5})$, so

$$\begin{aligned} (3, 2-\sqrt{-5})(2, 1+\sqrt{-5}) &= (6, 7+\sqrt{-5}) \\ &= ((1+\sqrt{-5})(1-\sqrt{-5}), (1+\sqrt{-5})(2-\sqrt{-5})) \\ &= (1+\sqrt{-5})(1-\sqrt{-5}, 2-\sqrt{-5}) \\ &= (1+\sqrt{-5}). \end{aligned}$$

Therefore $(1+\sqrt{-5}) = (2x)$, so we can use $x = \frac{1+\sqrt{-5}}{2}$:

$$(3, 2-\sqrt{-5}) = \frac{1+\sqrt{-5}}{2}(2, 1+\sqrt{-5}).$$

Since $(3, 2-\sqrt{-5})$ and $(2, 1+\sqrt{-5})$ are scalar multiples of each other, we can multiply the chosen $\mathbf{Z}$-basis of $(2, 1+\sqrt{-5})$ in Table 3.5 by $\frac{1+\sqrt{-5}}{2}$ to express the second matrix in the table as a representation of multiplication by $\sqrt{-5}$ on the ideal $(3, 2-\sqrt{-5})$:

$$\begin{pmatrix} -1 & -3 \\ 2 & 1 \end{pmatrix} = [m_{\sqrt{-5}}] \text{ on } (3, 2-\sqrt{-5}) \text{ with respect to } \{1+\sqrt{-5}, -2+\sqrt{-5}\}.$$

The matrices $\left(\begin{smallmatrix} -1 & -3 \\ 2 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 2 & 3 \\ -3 & -2 \end{smallmatrix}\right)$ are now seen as representations of $m_{\sqrt{-5}}$ with respect to different $\mathbf{Z}$-bases of the *same* ideal $(3, 2-\sqrt{-5})$. (That is the special feature of equivalent fractional ideals: a basis of one can be scaled to a basis of the other, so a matrix representation on one is also valid on the other.) All we have to do now is compute the change of basis matrix for the two bases $\{1+\sqrt{-5}, -2+\sqrt{-5}\}$ and $\{3, 2-\sqrt{-5}\}$ of $(3, 2-\sqrt{-5})$. Writing the second basis in terms of the first,

$$3 = 1\cdot(1+\sqrt{-5}) + (-1)\cdot(-2+\sqrt{-5}), \quad 2-\sqrt{-5} = 0\cdot(1+\sqrt{-5}) + (-1)\cdot(-2+\sqrt{-5}).$$

The change of basis matrix is $\left(\begin{smallmatrix} 1 & 0 \\ -1 & -1 \end{smallmatrix}\right)$, and $\left(\begin{smallmatrix} 1 & 0 \\ -1 & -1 \end{smallmatrix}\right)\left(\begin{smallmatrix} -1 & -3 \\ 2 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ -1 & -1 \end{smallmatrix}\right)^{-1} = \left(\begin{smallmatrix} 2 & 3 \\ -3 & -2 \end{smallmatrix}\right)$.

Example 3.6. The 4×4 matrix

$$A = \begin{pmatrix} 2 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 2 & -2 & 1 & 0 \end{pmatrix}$$

has characteristic polynomial $f(T) = T^4 - 4T^3 + T^2 - 4T + 1$, which is irreducible over $\mathbf{Q}$ since it is irreducible mod 3. An example of a root of $f(T)$ is $\alpha := 1 + \frac{1}{2}\sqrt{5} + \frac{1}{2}\sqrt{5+4\sqrt{5}}$. Let's find an ideal in the ring $\mathbf{Z}[\alpha]$ that is associated to the matrix A.

Set $K = \mathbf{Q}(\alpha)$. Make $\mathbf{Q}^4$ into a K-vector space by $g(\alpha)v = g(A)v$ for $g(T) \in \mathbf{Q}[T]$. Then $K \cong \mathbf{Q}^4$ as K-vector spaces by $g(\alpha) \mapsto g(A)v_0$, where v_0 is a fixed nonzero vector in $\mathbf{Q}^4$. A $\mathbf{Z}[\alpha]$-fractional ideal associated to A is $\mathfrak{a} = \{\gamma \in K : \gamma v_0 \in \mathbf{Z}^4\}$, or rather the $\mathbf{Z}[\alpha]$-*fractional ideal class* of $\mathfrak{a}$ is associated to the *conjugacy class* of A in $\mathrm{M}_4(\mathbf{Z})$.

To compute an example of $\mathfrak{a}$, we have to pick v_0. Let v_0 be the first standard basis vector of $\mathbf{Q}^4$. Writing $g(\alpha) = a + b\alpha + c\alpha^2 + d\alpha^3$,

$$g(\alpha)v_0 = (aI_4 + bA + cA^2 + dA^3)\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a+b+7c+29d \\ b+5c+18d \\ 4c+16d \\ 2b+2c+8d \end{pmatrix} = \begin{pmatrix} 1 & 1 & 7 & 29 \\ 0 & 1 & 5 & 18 \\ 0 & 0 & 4 & 16 \\ 0 & 2 & 2 & 8 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

Since

$$\begin{pmatrix} 1 & 1 & 7 & 29 \\ 0 & 1 & 5 & 18 \\ 0 & 0 & 4 & 16 \\ 0 & 2 & 2 & 8 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1/2 & -7/4 & -5/4 \\ 0 & 0 & -1/4 & 1/2 \\ 0 & 2 & -7/4 & -1 \\ 0 & -1/2 & 1/2 & 1/4 \end{pmatrix},$$

we have $g(\alpha)v_0 \in \mathbf{Z}^4$ if and only if

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 1 & 1/2 & -7/4 & -5/4 \\ 0 & 0 & -1/4 & 1/2 \\ 0 & 2 & -7/4 & -1 \\ 0 & -1/2 & 1/2 & 1/4 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} x + (1/2)y - (7/4)z - (5/4)w \\ -(1/4)z + (1/2)w \\ 2y - (7/4)z - w \\ -(1/2)y + (1/2)z + (1/4)w \end{pmatrix},$$

where $x, y, z, w \in \mathbf{Z}$. Feed these coefficient formulas for a, b, c, d into $g(\alpha)$ to get an explicit formula for the elements of $\mathfrak{a} = \{\gamma \in K : \gamma v_0 \in \mathbf{Z}^4\}$:

$$\left(x + \frac{1}{2}y - \frac{7}{4}z - \frac{5}{4}w\right) + \left(-\frac{1}{4}z + \frac{1}{2}w\right)\alpha + \left(2y - \frac{7}{4}z - w\right)\alpha^2 + \left(-\frac{1}{2}y + \frac{1}{2}z + \frac{1}{4}w\right)\alpha^3.$$

Rewrite this as a linear combination of x, y, z, and w:

$$x + \left(-\frac{\alpha^3}{2} + 2\alpha^2 + \frac{1}{2}\right)y + \left(\frac{1}{2}\alpha^3 - \frac{7}{4}\alpha^2 - \frac{\alpha}{4} - \frac{7}{4}\right)z + \left(\frac{1}{4}\alpha^3 - \alpha^2 + \frac{\alpha}{2} - \frac{5}{4}\right)w.$$

Since x, y, z, w run over $\mathbf{Z}$,

$$\mathfrak{a} = \mathbf{Z} + \left(-\frac{\alpha^3}{2} + 2\alpha^2 + \frac{1}{2}\right)\mathbf{Z} + \left(\frac{1}{2}\alpha^3 - \frac{7}{4}\alpha^2 - \frac{\alpha}{4} - \frac{7}{4}\right)\mathbf{Z} + \left(\frac{1}{4}\alpha^3 - \alpha^2 + \frac{\alpha}{2} - \frac{5}{4}\right)\mathbf{Z}.$$

Although $\mathfrak{a} \not\subset \mathbf{Z}[\alpha]$, we can scale $\mathfrak{a}$ to get an ideal in $\mathbf{Z}[\alpha]$ by using $\mathfrak{b} = 4\mathfrak{a}$:

$$\mathfrak{b} = 4\mathbf{Z} + \left(-2\alpha^3 + 8\alpha^2 + 2\right)\mathbf{Z} + \left(2\alpha^3 - 7\alpha^2 - \alpha - 7\right)\mathbf{Z} + \left(\alpha^3 - 4\alpha^2 + 2\alpha - 5\right)\mathbf{Z}. \tag{3.2}$$

General theory tells us $\mathfrak{b}$ is an ideal in $\mathbf{Z}[\alpha]$, but the reader can also verify this by checking $\mathfrak{b}$ is a $\mathbf{Z}[\alpha]$-module: show $\alpha\beta \subset \mathfrak{b}$ where β runs over the four generators of $\mathfrak{b}$.

Writing the four $\mathbf{Z}$-module generators of $\mathfrak{b}$ in (3.2) as β_1, β_2, β_3, and β_4, we can simplify the choice of generators. Since $\beta_2 + \beta_3 = \alpha^2 - \alpha - 5$ and $\beta_2 + 2\beta_4 = 4\alpha - 8$,

$$\mathfrak{b} = 4\mathbf{Z} + (4\alpha - 8)\mathbf{Z} + (\alpha^2 - \alpha - 5)\mathbf{Z} + (\alpha^3 - 4\alpha^2 + 2\alpha - 5)\mathbf{Z}. \tag{3.3}$$

There is the potential to lower the number of generators of $\mathfrak{b}$ by treating it as an ideal (a $\mathbf{Z}[\alpha]$-module). Using the generators in (3.3), since $4\alpha - 8 \in (4)$ we can write

$$\begin{aligned}\mathfrak{b} &= (4, \alpha^2 - \alpha - 5, \alpha^3 - 4\alpha^2 - 2\alpha - 5)\\ &= (4, \alpha^2 - \alpha - 1, \alpha^3 - 2\alpha - 1)\\ &= (4, \alpha^2 - \alpha - 1)\end{aligned}$$

since $\alpha^3 - 2\alpha - 1 = \alpha(\alpha^2 - \alpha - 1) + \alpha^2 - \alpha - 1$.

It can be shown with a computer algebra system that $\mathfrak{b}$ is not a principal ideal in $\mathbf{Z}[\alpha]$.[3] Then by Corollary 2.8(b), A is not conjugate in $\mathrm{M}_4(\mathbf{Z})$ to the companion matrix C_f of $f(T)$ even though A and C_f have the same irreducible characteristic polynomial in $\mathbf{Z}[T]$.

Example 3.7. Let's find an integral matrix $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{M}_2(\mathbf{Z})$ such that A and $A^\top$ are not conjugate to each other in $\mathrm{M}_2(\mathbf{Z})$. (In $\mathrm{M}_n(F)$ for a field F, a matrix and its transpose are conjugate.) The characteristic polynomials of A and $A^\top$ are the same, namely

$$\chi(T) = T^2 - (a + d)T + (ad - bc) \in \mathbf{Z}[T].$$

Suppose $\chi(T)$ is irreducible in $\mathbf{Z}[T]$, which is fairly typical anyway. Let α be a root of $\chi(T)$. We will produce ideals in $\mathbf{Z}[\alpha]$ corresponding to A and $A^\top$, and then carefully select a choice of α for which those two ideals are guaranteed to be in different ideal classes of $\mathbf{Z}[\alpha]$.

Let $K = \mathbf{Q}(\alpha)$, and make $\mathbf{Q}^2$ into a K-vector space by

$$(r + s\alpha)\begin{pmatrix} x \\ y \end{pmatrix} := (rI_2 + sA)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r + sa & sb \\ sc & r + sd \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

[3] It turns out $\mathfrak{b}$ is *also* an ideal in the ring of integers of $K = \mathbf{Q}(\alpha)$, and in that ring it is principal: $\mathfrak{b} = (\alpha^2 - \alpha - 1)\mathcal{O}_K$ since $(\alpha^2 - \alpha - 1) \mid 4$ in $\mathcal{O}_K$ while $(\alpha^2 - \alpha - 1) \nmid 4$ in $\mathbf{Z}[\alpha]$. And in fact $\mathcal{O}_K$ is a PID.

A fractional $\mathbf{Z}[\alpha]$-ideal corresponding to A in $\mathbf{Z}[\alpha]$ is

$$\mathfrak{a} = \left\{ m + n\alpha : m, n \in \mathbf{Z}, (m+n\alpha)\begin{pmatrix}1\\0\end{pmatrix} \in \mathbf{Z}^2 \right\} = \left\{ m + n\alpha : m, n \in \mathbf{Z}, \begin{pmatrix} m+na \\ nc \end{pmatrix} \in \mathbf{Z}^2 \right\}.$$

The condition

$$\begin{pmatrix} m+na \\ nc \end{pmatrix} \in \mathbf{Z}^2$$

is the same as

$$\begin{pmatrix} 1 & a \\ 0 & c \end{pmatrix} \begin{pmatrix} m \\ n \end{pmatrix} \in \mathbf{Z}^2,$$

so

$$\begin{pmatrix} m \\ n \end{pmatrix} = \begin{pmatrix} 1 & a \\ 0 & c \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - ay/c \\ y/c \end{pmatrix}$$

with integers x and y. Therefore we have the fractional $\mathbf{Z}[\alpha]$-ideal

$$\mathfrak{a} = \left\{ x - \frac{ay}{c} + \frac{y}{c}\alpha : x, y \in \mathbf{Z} \right\} = \mathbf{Z} + \left(\frac{-a+\alpha}{c} \right) \mathbf{Z}.$$

If we run through this argument using $A^\top$ in place of A, the roles of b and c get flipped, so the corresponding fractional $\mathbf{Z}[\alpha]$-ideal is

$$\mathfrak{a}' = \mathbf{Z} + \left(\frac{-a+\alpha}{b} \right) \mathbf{Z}.$$

Scaling fractional ideals doesn't change the correspondence, so we replace $\mathfrak{a}$ with $c\mathfrak{a}$ and $\mathfrak{a}'$ with $b\mathfrak{a}'$. That is, redefine as ideals in $\mathbf{Z}[\alpha]$

$$\mathfrak{a} = \mathbf{Z}c + \mathbf{Z}(\alpha - a) = (c, \alpha - a), \quad \mathfrak{a}' = \mathbf{Z}b + \mathbf{Z}(\alpha - a).$$[4]

The link between these two ideals in $\mathbf{Z}[\alpha]$ and the matrices A and $A^\top$ is that multiplication by α on $\mathfrak{a}$ with respect to its $\mathbf{Z}$-basis $\{c, \alpha - a\}$ is $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = A$, and multiplication by α on $\mathfrak{a}'$ with respect to its $\mathbf{Z}$-basis $\{b, \alpha - a\}$ is $\left(\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right) = A^\top$.

Putting everything together, we can formulate our task in terms of ideals rather than (non)conjugate matrices: find $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{M}_2(\mathbf{Z})$ such that its characteristic polynomial $\chi(T)$ is irreducible over $\mathbf{Q}$ and for a root α of $\chi(T)$, the ideals $\mathfrak{a} := (c, \alpha - a)$ and $\mathfrak{a}' := (b, \alpha - a)$ in $\mathbf{Z}[\alpha]$ are in different ideal classes: for no $\lambda \in \mathbf{Q}(\alpha)^\times$ does $\mathfrak{a}' = \lambda\mathfrak{a}$.

We can replace $\mathfrak{a}'$ by the conjugate ideal to $\mathfrak{a}$, $\overline{\mathfrak{a}} = (c, \overline{\alpha} - a)$, because they are in the same ideal class: scaling $\mathfrak{a}'$ by c and using the condition $bc = \alpha^2 - (a+d)\alpha + ad = (\alpha - a)(\alpha - d)$,

$$c\mathfrak{a}' = (bc, c\alpha - ca) = ((\alpha - a)(\alpha - d), (\alpha - a)c) = (\alpha - a)(\alpha - d, c),$$

and

$$\overline{\mathfrak{a}} = (c, \overline{\alpha} - a) = (c, a + d - \alpha - a) = (c, \alpha - d),$$

so $c\mathfrak{a}' = (\alpha - a)\overline{\mathfrak{a}}$. Thus we want $\mathfrak{a}$ and $\overline{\mathfrak{a}}$ to be in different ideal classes of $\mathbf{Z}[\alpha]$.

Let's find an example where $\mathbf{Z}[\alpha]$ is the *full ring of integers* of $\mathbf{Q}(\alpha)$, since that implies $\mathfrak{a}\overline{\mathfrak{a}}$ is a principal ideal[5], so $[\overline{\mathfrak{a}}]$ is the inverse ideal class to $[\mathfrak{a}]$. Asking for $\mathfrak{a}$ and $\overline{\mathfrak{a}}$ to be in

[4]It is not obvious initially that $\mathbf{Z}c + \mathbf{Z}(\alpha - a)$ and $\mathbf{Z}b + \mathbf{Z}(\alpha - a)$ are ideals in $\mathbf{Z}[\alpha]$, meaning they are stable under multiplication by α, but this could be checked directly by the fact that α is a root of $\chi(T) = T^2 - (a+d)T + (ad - bc)$.

[5]By Theorem 5.6 in https://kconrad.math.uconn.edu/blurbs/ugradnumthy/quadraticundergrad.pdf, $\mathfrak{a}\overline{\mathfrak{a}} = (c^2, c(d-a), bc)$ and an ideal with generators in $\mathbf{Z}$ is principal.

different ideal classes of $\mathbf{Z}[\alpha]$ is asking for $[\mathfrak{a}]$ to have order greater than 2 in the ideal class group of $\mathbf{Q}(\alpha)$.

Consulting tables of class numbers of quadratic fields, the first imaginary quadratic field with $h > 2$ is $\mathbf{Q}(\sqrt{-14})$, where $h = 4$. You can check that the ideal $(3, 1 - \sqrt{-14})$ in the ring of integers $\mathbf{Z}[\sqrt{-14}]$ has order 4 in the class group. (That is, the smallest power of this ideal that is principal is its 4th power.) Let's use this ideal. Write $(3, 1 - \sqrt{-14})$ as $(c, a - \alpha)$ using $c = 3$, $a = 1$, and $\alpha = \sqrt{-14}$. We want $A = \left(\begin{smallmatrix} 1 & ? \\ 3 & ? \end{smallmatrix}\right)$ to have characteristic polynomial $T^2 + 14$, so $d = -1$ and $14 = ad - bc = -1 - 3b$, so $b = -5$: our matrix is $\left(\begin{smallmatrix} 1 & -5 \\ 3 & -1 \end{smallmatrix}\right)$. This integral matrix is not conjugate in $\mathrm{M}_2(\mathbf{Z})$ to its transpose $\left(\begin{smallmatrix} 1 & 3 \\ -5 & -1 \end{smallmatrix}\right)$ because it corresponds to an ideal class in $\mathbf{Z}[\sqrt{-14}]$ having order greater than 2.

The statement that $\left(\begin{smallmatrix} 1 & -5 \\ 3 & -1 \end{smallmatrix}\right)$ and its transpose $\left(\begin{smallmatrix} 1 & 3 \\ -5 & -1 \end{smallmatrix}\right)$ are not conjugate in $\mathrm{M}_2(\mathbf{Z})$ is wholly elementary, not mentioning ideals at all, and the fact that they are not conjugate in $\mathrm{M}_2(\mathbf{Z})$ can be proved by contradiction in an elementary way. But if you follow that approach then you have absolutely no idea how the example was found or even how to find more examples. The way we went about finding the example shows a path through number theory by which many more examples can be found.

Over a field, a square matrix and its transpose are conjugate, so the matrix $\left(\begin{smallmatrix} 1 & -5 \\ 3 & -1 \end{smallmatrix}\right)$ from the preceding example is conjugate to its transpose $\left(\begin{smallmatrix} 1 & 3 \\ -5 & -1 \end{smallmatrix}\right)$ in $\mathrm{M}_2(\mathbf{Q})$. Two explicit $U \in \mathrm{GL}_2(\mathbf{Q})$ satisfying $U\left(\begin{smallmatrix} 1 & -5 \\ 3 & -1 \end{smallmatrix}\right)U^{-1} = \left(\begin{smallmatrix} 1 & 3 \\ -5 & -1 \end{smallmatrix}\right)$ are $U_1 = \left(\begin{smallmatrix} 1 & -1 \\ -1 & -1 \end{smallmatrix}\right)$ (with determinant -2) and $U_2 = \left(\begin{smallmatrix} 2 & 1 \\ 1 & -4 \end{smallmatrix}\right)$ (with determinant -9). Darij Grinberg noticed an interesting aspect of this example when we reduce mod p. Since $\mathbf{Z}/p\mathbf{Z}$ is a field, $\left(\begin{smallmatrix} 1 & -5 \\ 3 & -1 \end{smallmatrix}\right)$ and its transpose $\left(\begin{smallmatrix} 1 & 3 \\ -5 & -1 \end{smallmatrix}\right)$ are conjugate in $\mathrm{M}_2(\mathbf{Z}/p\mathbf{Z})$ for every prime p, and we can make this explicit: U_1 mod p is a conjugating matrix when $p \neq 2$ and U_2 mod p works when $p \neq 3$. Therefore it is false that if two matrices in $\mathrm{M}_2(\mathbf{Z})$ are conjugate in $\mathrm{M}_2(\mathbf{Z}/p\mathbf{Z})$ for all primes p then they are conjugate in $\mathrm{M}_2(\mathbf{Z})$.

The matrices U_1 and U_2 also show us that $\left(\begin{smallmatrix} 1 & -5 \\ 3 & -1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 3 \\ -5 & -1 \end{smallmatrix}\right)$ are conjugate in $\mathrm{M}_2(\mathbf{Z}_{(p)})$ for every prime p (use U_1 when $p \neq 2$ and U_2 when $p \neq 3$). Guralnick [3] proved that if two matrices in $\mathrm{M}_n(\mathbf{Z})$ are conjugate in $\mathrm{M}_n(\mathbf{Z}_{(p)})$ for all primes p then they are conjugate in $\mathrm{M}_n(\mathcal{O}_K)$ for some number field K. With a little work we can make this explicit for our example: $U\left(\begin{smallmatrix} 1 & -5 \\ 3 & -1 \end{smallmatrix}\right)U^{-1} = \left(\begin{smallmatrix} 1 & 3 \\ -5 & -1 \end{smallmatrix}\right)$ when $U = \left(\begin{smallmatrix} 1+2i & 2-2i \\ 2-2i & -3-2i \end{smallmatrix}\right)$, which lies in $\mathrm{SL}_2(\mathbf{Z}[i])$.

Moving beyond the 2×2 case, we can explain in terms of ideal classes when a matrix $A \in \mathrm{M}_n(\mathbf{Z})$ with an irreducible characteristic polynomial in $\mathbf{Z}[T]$ is not conjugate over $\mathbf{Z}$ to $A^\top$. If $f(\alpha) = 0$, $K = \mathbf{Q}(\alpha)$, and A is a matrix representing multiplication by α on a fractional $\mathbf{Z}[\alpha]$-ideal $\mathfrak{a}$, then $A^\top$ is a matrix for multiplication by α on the dual lattice $\mathfrak{a}^\vee = \{x \in K : \mathrm{Tr}_{K/\mathbf{Q}}(x\mathfrak{a}) \subset \mathbf{Z}\}$, where we view K as self-dual for the trace pairing.[6] That's because if $\{e_1, \ldots, e_n\}$ is the $\mathbf{Z}$-basis of $\mathfrak{a}$ such that m_α with respect to this basis has matrix A then the dual basis is a $\mathbf{Z}$-basis of $\mathfrak{a}^\vee$ and the matrix for m_α with respect to that dual basis is $A^\top$. So A and $A^\top$ are not conjugate over $\mathbf{Z}$ if and only if $\mathfrak{a}$ and $\mathfrak{a}^\vee$ are in different ideal classes of $\mathbf{Z}[\alpha]$.

In the special case that $\mathbf{Z}[\alpha] = \mathcal{O}_K$, all fractional $\mathbf{Z}[\alpha]$-ideals $\mathfrak{a}$ satisfy $\mathfrak{a}^\vee = \mathfrak{a}^{-1}\mathcal{D}_K^{-1}$, where $\mathcal{D}_K$ is the different ideal, so if $\mathcal{D}_K$ is a principal ideal then $\mathfrak{a}^\vee$ is in the same ideal class as $\mathfrak{a}^{-1}$. Then A and $A^\top$ are not conjugate in $\mathrm{M}_n(\mathbf{Z})$ if and only if $\mathfrak{a}$ and $\mathfrak{a}^{-1}$ lie in different ideals classes, which is the same as $\mathfrak{a}^2$ not being principal. That is a condition we met in Example 3.7, which is a special case of these considerations since the different ideal

[6] See Section 3 of https://kconrad.math.uconn.edu/blurbs/gradnumthy/different.pdf.

in a quadratic field is principal. It is also principal when the class group of K is 2-torsion since it is a theorem of Hecke that the ideal class of $\mathcal{D}_K$ is always a square in the class group of K.

References

[1] . T. Campbell and E. C. Trouy, *When are two elements of* $\mathrm{GL}(2,\mathbf{Z})$ *similar*?, Linear Algebra Appl. **157** (1991), 175–184.

[2] D. Clausen, http://mathoverflow.net/questions/95536/stable-conjugacy-for-integer-matrices.

[3] R. M. Guralnick, *A note on the local-global principle for similarity of matrices*, Linear Algebra Appl. **30** (1980), 241–245.

[4] C. Latimer and C. C. MacDuffee, *A correspondence between classes of ideals and classes of matrices*, Ann. of Math. **34** (1933), 313–316.

[5] M. Newman, "Integral Matrices," Academic Press, New York, 1972.

[6] O. Taussky, *On a theorem of Latimer and MacDuffee*, Canadian J. Math **1** (1949), 300–302.

[7] D. I. Wallace, *Conjugacy classes of hyperbolic matrices in* $\mathrm{SL}_n(\mathbf{Z})$ *and ideal classes in an order*, Trans. Amer. Math. Soc. **283** (1984), 177–184.