On the Conjugates of 2×2 Matrices

R. C. Daileda

October 11, 2017

Let F be a field and let $A \in \mathrm{M}_n(F)$. The group $\mathrm{GL}_n(F)$ acts on $\mathrm{M}_n(F)$ by conjugation. Slighty abusing terminology and symbolism, we will call the orbit of A under this action its *conjugacy class* and will denote it by $\mathrm{Cl}(A)$:

$$\mathrm{Cl}(A) = \{XAX^{-1} \,|\, X \in \mathrm{GL}_n(F)\}.$$

The goal of this note is to study $|\mathrm{Cl}\, A|$ when $n = 2$, i.e. the number of $\mathrm{GL}_2(F)$ conjugates of $A \in \mathrm{M}_2(F)$. Our main result is the following.

Theorem 1. *If F is an infinite field and $A \in \mathrm{M}_2(F)$ is not a scalar matrix, then $\mathrm{Cl}(A)$ is infinite.*

Theorems 2 through 5 allow one to compute $|\,\mathrm{Cl}(A)|$ explicitly when F is finite.

According to the Orbit-Stabilizer Theorem, there is a bijective correspondence between $\mathrm{Cl}(A)$ and the coset space of the stablizer

$$C(A) = \{X \in \mathrm{GL}_n(F) \,|\, XAX^{-1} = A\} = \{X \in \mathrm{GL}_n(F) \,|\, AX = XA\},$$

(the *centralizer* of A) given by $X \cdot C(A) \mapsto XAX^{-1}$. In particular, $|\,\mathrm{Cl}(A)| = [\mathrm{GL}_n(F) : C(A)]$, so a natural place to start our work is with the study of $C(A)$.

Because it will allow us to bring the techniques of linear algebra to bear on the problem, we consider the larger set

$$C'(A) = \{X \in \mathrm{M}_n(F) \,|\, AX = XA\}.$$

Clearly $C(A) = C'(A) \cap \mathrm{GL}_n(F)$. We define the commutator of two $n \times n$ matrices to be $[A, B] = AB - BA$. The map $X \mapsto [A, X]$ is an F-vector space endomorphism of $\mathrm{M}_n(F)$ whose kernel is precisely $C'(A)$. It can be studied by analyzing the null space of the matrix for $X \mapsto [A, X]$ relative to some ordered basis. We choose to use the basis $\{E_{ij}\}$, $i, j \in \{1, 2, \ldots, n\}$, where E_{ij}'s only nonzero entry is a 1 in the (i, j) position, declaring that $E_{ij} < E_{\ell m}$ if $i = \ell$ and $j < m$ or $i < \ell$. Relative to this basis, it is not difficult to show that the matrix of $X \mapsto [A, X]$ is

$$T = A \otimes I - I \otimes A, \tag{1}$$

$\otimes$ denoting the Kronecker product of matrices.

We now set $n = 2$ and $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Relative to the ordered basis

$$E_{1,1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; E_{1,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \; E_{2,1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \; E_{2,2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \tag{2}$$

by either using (1) or computing directly, it is not difficult to see that the matrix for $X \mapsto [A, X]$ in this case is

$$T = \begin{pmatrix} 0 & -c & b & 0 \\ -b & a-d & 0 & b \\ c & 0 & d-a & -c \\ 0 & c & -b & 0 \end{pmatrix}.$$

$C'(A)$ consists precisely of those matrices whose coordinate vectors relative to (2) lie in the null space of T, and the dimension of that space is $\dim \operatorname{null} T = 4 - \operatorname{rank} T$.

Because $(1, 0, 0, 1)^T \in \operatorname{null} T$, we immediately conclude that all *scalar matrices* $\left(\begin{smallmatrix} \lambda & \\ & \lambda \end{smallmatrix}\right)$ commute with A. Of course, this observation can be made directly, but it is interesting to see it arise naturally from our considerations. The real question is what else, if anything, commutes with A? Notice that any polynomial in A will commute with A so that $F[A]$ is a subspace of $C'(A)$.

If $(b,c) \neq (0,0)$, then A and I are linearly independent, which means that the minimal polynomial of A must have degree 2, and hence $\dim F[A] = 2$. Similarly, if $b = c = 0$ and $a \neq d$, then A has two distinct eigenvalues and its minimal polynomial again has degree 2, implying that $\dim F[A] = 2$. But in both cases at least two of the first three columns of T are linearly independent, so that $\operatorname{rank} T \geq 2$ and $\dim C'(A) = \dim A \leq 2$. We conclude that $\dim C'(A) = 2$ and $C'(A) = F[A] = \{\lambda I + \mu A \,|\, \lambda, \mu \in F\}$.

At this point a remark is in order. In the latter case, the elements of $F[A]$ take the form

$$\begin{pmatrix} \lambda + \mu a & \\ & \lambda + \mu d \end{pmatrix}.$$

It is not difficult to show that since $a \neq d$, the diagonal entries can take on any values in F, so that we have the alternate, and more explicit, description

$$C'(A) = F[A] = \left\{ \begin{pmatrix} \lambda & \\ & \mu \end{pmatrix} \,\middle|\, \lambda, \mu \in F \right\}.$$

The remaining case occurs when $b = c = 0$ and $a = d$, i.e. when A is a scalar matrix. Since scalar matrices commute with all others (or since $T = 0$ in this case), we conclude that $C'(A) = \mathrm{M}_2(F)$. Table 1 summarizes our findings.

Conditions	$\boldsymbol{C'(A)}$
$(b,c) = (0,0)$, $a = d$	$\mathrm{M}_2(F)$
$(b,c) = (0,0)$, $a \neq d$	$F[A] = \left\{ \begin{pmatrix} \lambda & \\ & \mu \end{pmatrix} \,\middle|\, \lambda, \mu \in F \right\}$
$(b,c) \neq (0,0)$	$F[A] = \{\lambda I + \mu A \,\|\, \lambda, \mu \in F\}$

Table 1: The centralizer of $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{M}_2(F)$

We now return to our original question, that of determining $|\operatorname{Cl}(A)| = [\mathrm{GL}_2(F) : C(A)]$. If A is scalar we clearly have $|\operatorname{Cl}(A)| = 1$. If A is diagonal but nonscalar, then according to the results above

$$C(A) = C'(A) \cap \mathrm{GL}_2(F) = \left\{ \begin{pmatrix} \lambda & \\ & \mu \end{pmatrix} \,\middle|\, \lambda, \mu \in F^\times \right\}$$

and we need to determine the order of the coset space $\mathrm{GL}_2(F)/\{\left(\begin{smallmatrix} \lambda & \\ & \mu \end{smallmatrix}\right)\}$.

Theorem 2. *There is a bijection between* $\mathrm{GL}_2(F)/\{\left(\begin{smallmatrix} \lambda & \\ & \mu \end{smallmatrix}\right)\}$ *and* $N = \{(P,Q) \in \mathbb{P}^1(F) \times \mathbb{P}^1(F) \,|\, P \neq Q\}$.

Proof. We define

$$\begin{aligned} \Phi : \mathrm{GL}_2(F)/\{\left(\begin{smallmatrix} \lambda & \\ & \mu \end{smallmatrix}\right)\} &\to N \\ (\mathbf{a}\ \mathbf{b}) \cdot \{\left(\begin{smallmatrix} \lambda & \\ & \mu \end{smallmatrix}\right)\} &\mapsto ([\mathbf{a}], [\mathbf{b}]), \end{aligned}$$

where $\mathbf{a}$ and $\mathbf{b}$ are column vectors and $[\,\cdot\,]$ denotes the projective equivalence class of a vector. To verify that Φ is well-defined we need to check two things: that it does not depend on the coset representative chosen and that its image actually lies in N. The latter is trivial. If $(\mathbf{a}\ \mathbf{b})$ is invertible, $\mathbf{a}$ and $\mathbf{b}$ must be linearly independent, i.e. nonmultiples. This is equivalent to $[\mathbf{a}] \neq [\mathbf{b}]$. Now suppose that $(\mathbf{a}\ \mathbf{b})$ and $(\mathbf{c}\ \mathbf{d})$ belong to the same coset of $\{\left(\begin{smallmatrix} \lambda & \\ & \mu \end{smallmatrix}\right)\}$. Then

$$(\mathbf{c}\ \mathbf{d}) = (\mathbf{a}\ \mathbf{b}) \begin{pmatrix} \lambda & \\ & \mu \end{pmatrix} = (\lambda\mathbf{a}\ \mu\mathbf{b})$$

so that $[\mathbf{c}] = [\mathbf{a}]$ and $[\mathbf{d}] = [\mathbf{b}]$, as needed.

If the cosets of $(\mathbf{a}\ \mathbf{b})$ and $(\mathbf{c}\ \mathbf{d})$ have the same image under Φ, then $([\mathbf{a}], [\mathbf{b}]) = ([\mathbf{c}], [\mathbf{d}])$, and so there exist $\lambda, \mu \in F^\times$ so that $\lambda\mathbf{a} = \mathbf{c}$ and $\mu\mathbf{b} = \mathbf{d}$. That is, $(\mathbf{a}\ \mathbf{b})\left(\begin{smallmatrix}\lambda & \\ & \mu\end{smallmatrix}\right) = (\mathbf{c}\ \mathbf{d})$, and the cosets of $(\mathbf{a}\ \mathbf{b})$ and $(\mathbf{c}\ \mathbf{d})$ are the same. Hence Φ is injective.

Finally, given any pair of projective points (P, Q) with $P \neq Q$, choose representative vectors so that $P = [\mathbf{a}]$ and $Q = [\mathbf{b}]$. Because $P \neq Q$, $\mathbf{a}$ and $\mathbf{b}$ are not multiples of one another. Hence $(\mathbf{a}\ \mathbf{b})$ is invertible and its coset furnishes the preimage of (P, Q) under Φ.

□

Corollary 1. *If A is a 2×2 nonscalar diagonal matrix over an infinite field, its $\mathrm{GL}_2(F)$ conjugacy class is infinite.*

Remark. If F is finite and A is a nonscalar diagonal matrix, it clearly would not be difficult to use Theorem 2 to determine the size of the conjugacy class of A explicitly.

Now suppose that A is not diagonal. An element $\lambda I + \mu A \in F[A] = C'(A)$ is invertible precisely when

$$\det(\lambda I + \mu A) = \lambda^2 + \mu\lambda \operatorname{Tr} A + \mu^2 \det A \neq 0.$$

If $\mu = 0$, this is equivalent to $\lambda \neq 0$, but if $\mu \neq 0$, this can be rewritten as

$$\mu^2 f_A(-\lambda/\mu) \neq 0 \quad\Leftrightarrow\quad f_A(-\lambda/\mu) \neq 0,$$

where f_A is the characteristic polynomial of A. This occurs precisely when $-\lambda/\mu$ is *not* an eigenvalue of A. If the eigenvalues of A do not belong to F (they simultaneously do or do not), this condition is automatic and so

$$C(A) = \{\lambda I + \mu A \,|\, (\lambda, \mu) \in F^2 \setminus \{(0, 0)\}\}.$$

Otherwise, denoting the eigenvalues by $r_1, r_2 \in F$ (which may be the same), this means

$$(\lambda, \mu) \notin \underbrace{F \cdot (-r_1, 1)}_{L_1} \cup \underbrace{F \cdot (-r_2, 1)}_{L_2}.$$

Let

$$P(r_1, r_2) = \left\{(\lambda, \mu) \,|\, (\lambda, \mu) \in F^2 \setminus L_1 \cup L_2\right\}.$$

We have just proven that

$$C(A) = \{\lambda I + \mu A \,|\, (\lambda, \mu) \in P(r_1, r_2)\}$$

in this case. Table 2 summarizes our computations of the centralizer of A.

We now explicitly compute the coset $X \cdot C(A)$ for an arbitrary $X = \left(\begin{smallmatrix} x & y \\ z & w\end{smallmatrix}\right) \in \mathrm{GL}_2(F)$. Without loss of generality, assume $c \neq 0$. Given $(\lambda, \mu) \in F^2$, we have

$$\begin{aligned} X(\lambda I + \mu A) &= \left(X\begin{pmatrix} \lambda + \mu a \\ \mu c \end{pmatrix} \ X\begin{pmatrix} \mu b \\ \lambda + \mu d \end{pmatrix} \right) \\ &= \left(X\begin{pmatrix} 1 & a \\ 0 & c \end{pmatrix}\begin{pmatrix} \lambda \\ \mu \end{pmatrix} \ X\begin{pmatrix} 0 & b \\ 1 & d \end{pmatrix}\begin{pmatrix} \lambda \\ \mu \end{pmatrix} \right). \end{aligned} \tag{3}$$

Let $\left(\begin{smallmatrix} r \\ s \end{smallmatrix}\right) = X\left(\begin{smallmatrix} 1 & a \\ 0 & c \end{smallmatrix}\right)\left(\begin{smallmatrix} \lambda \\ \mu \end{smallmatrix}\right)$ so that $\left(\begin{smallmatrix} \lambda \\ \mu \end{smallmatrix}\right) = \frac{1}{c}\left(\begin{smallmatrix} c & -a \\ 0 & 1 \end{smallmatrix}\right) X^{-1}\left(\begin{smallmatrix} r \\ s \end{smallmatrix}\right)$ and the right-hand side of (3) becomes

$$\begin{aligned} \left(\begin{matrix} r \\ s \end{matrix} \ \ \frac{1}{c} X\begin{pmatrix} 0 & b \\ c & d - a \end{pmatrix} X^{-1}\begin{pmatrix} r \\ s \end{pmatrix}\right) &= \left(\begin{matrix} r \\ s \end{matrix} \ \ \frac{1}{c} X(A - aI)X^{-1}\begin{pmatrix} r \\ s \end{pmatrix}\right) \\ &= \left(\begin{matrix} r \\ s \end{matrix} \ \ \frac{1}{c}\,(XAX^{-1} - aI)\begin{pmatrix} r \\ s \end{pmatrix}\right). \end{aligned} \tag{4}$$

When the eigenvalues of A are outside of F, (λ, μ) is free to take on any nonzero value. Since $X\left(\begin{smallmatrix} 1 & a \\ 0 & c \end{smallmatrix}\right)$ is invertible, the same is true of (r, s). Consequently we obtain the following result.

Conditions	$\boldsymbol{C(A)}$
$(b,c)=(0,0)$ $a=d$	$\mathrm{GL}_2(F)$
$(b,c)=(0,0)$ $a\neq d$	$\left\{ \begin{pmatrix} \lambda & \\ & \mu \end{pmatrix} \,\middle|\, \lambda,\mu\in F^\times \right\}$
$(b,c)\neq(0,0)$ $r_1,r_2\notin F$	$\{\lambda I+\mu A \,\vert\, (\lambda,\mu)\in F^2\setminus\{(0,0)\}\}$
$(b,c)\neq(0,0)$ $r_1,r_2\in F$	$\{\lambda I+\mu A \,\vert\, (\lambda,\mu)\in P(r_1,r_2)\}$

Table 2: The centralizer of $A=\left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right)$ in $\mathrm{GL}_2(F)$

Lemma 1. *If $A=\left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right)\in\mathrm{M}_2(F)$, $(b,c)\neq(0,0)$ and the eigenvalues of A do not belong to F, then for any $X\in\mathrm{GL}_2(F)$*

$$X\cdot C(A)=\left\{ \left(\begin{array}{cc} r & \frac{1}{c}\,(XAX^{-1}-aI)\begin{pmatrix} r\\ s\end{pmatrix} \\ s & \end{array}\right) \,\middle|\, (r,s)\in F^2\setminus\{(0,0)\} \right\}.$$

We can now parametrize the coset space $\mathrm{GL}_2(F)/C(A)$ in this case.

Theorem 3. *If $A=\left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right)\in\mathrm{M}_2(F)$, $(b,c)\neq(0,0)$ and the eigenvalues of A do not belong to F, then there is a bijection $\mathrm{GL}_2(F)/C(A)\to F\times F^\times$.*

Proof. According to Lemma 1, each coset $X\cdot C(A)$ contains a unique upper triangular element of the form $\left(\begin{smallmatrix} 1 & u\\ 0 & v\end{smallmatrix}\right)$ with $(u,v)\in F\times F^\times$. We map $X\cdot C(A)$ to (u,v). This is injective since distinct cosets are disjoint. It is surjective since given $(u,v)\in F\times F^\times$, the matrix $Y=\left(\begin{smallmatrix} 1 & u\\ 0 & v\end{smallmatrix}\right)$ belongs to the coset $Y\cdot C(A)$, which therefore maps to (u,v). $\square$

Corollary 2. *If $A=\left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right)\in\mathrm{M}_2(F)$, $(b,c)\neq(0,0)$, the eigenvalues of A do not belong to F, and F is infinite, then the conjugacy class $\mathrm{Cl}(A)$ is infinite.*

In the case that the eigenvalues of A do belong to F, for $i=1,2$ let

$$\mathbf{s}_i(X)=X\begin{pmatrix} 1 & a\\ 0 & c\end{pmatrix}\begin{pmatrix} -r_i\\ 1\end{pmatrix}=X\begin{pmatrix} a-r_i\\ c\end{pmatrix}. \tag{5}$$

Then $(r,s)^T$ of (4) must belong to $P(\mathbf{s}_1(X),\mathbf{s}_2(X))$. This yields the next result.

Lemma 2. *If $A=\left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right)\in\mathrm{M}_2(F)$, $(b,c)\neq(0,0)$ and the eigenvalues of A belong to F, then for any $X\in\mathrm{GL}_2(F)$*

$$X\cdot C(A)=\left\{ \left(\begin{array}{cc} r & \frac{1}{c}\,(XAX^{-1}-aI)\begin{pmatrix} r\\ s\end{pmatrix} \\ s & \end{array}\right) \,\middle|\, \begin{pmatrix} r\\ s\end{pmatrix}\in P(\mathbf{s}_1(X),\mathbf{s}_2(X)) \right\}.$$

We can now begin to parametrize the coset space $\mathrm{GL}_2(F)/C(A)$ when the eigenvalues of A belong to F.

Theorem 4. *Suppose $A=\left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right)\in\mathrm{M}_2(F)$, $(b,c)\neq(0,0)$ and the eigenvalues of A are distinct and belong to F. Then there is a bijection $\mathrm{GL}_2(F)/C(A)\to\{(P,Q)\in\mathbb{P}^1(F)\times\mathbb{P}^1(F)\,\vert\, P\neq Q\}$.*

Proof. Again we assume $c \neq 0$, without loss of generality. We map $X \cdot C(A)$ to $([\mathbf{s}_1(X)], [\mathbf{s}_2(X)])$. Because we have assumed $r_1 \neq r_2$, the definition of $\mathbf{s}_i$ shows that $[\mathbf{s}_1(X)] \neq [\mathbf{s}_2(X)]$, so this is well-defined. For the same reason, it is always possible to choose $X \in \mathrm{GL}_2(F)$ so that the $\mathbf{s}_i(X)$ take on any specified values. Hence this map is surjective.

Proving injectivity requires the most work. Suppose that $([\mathbf{s}_1(X)], [\mathbf{s}_2(X)]) = ([\mathbf{s}_1(Y)], [\mathbf{s}_2(Y)])$ for some $X, Y \in \mathrm{GL}_2(F)$. Using the definition (5), we find that this means the vectors $(a - r_i, c)^T$, $i = 1, 2$, are eigenvectors of $Y^{-1}X$. However, since $\operatorname{rank} A - r_i I = 1$ and $c \neq 0$, these vectors span the r_i-eigenspaces of A as well. It follows that the matrix

$$\begin{pmatrix} a - r_1 & a - r_2 \\ c & c \end{pmatrix}$$

diagonalizes both A and $Y^{-1}X$. As diagonal matrices commute, this can easily be shown to imply that A and $Y^{-1}X$ commute, i.e. $Y^{-1}X \in C(A)$. Thus $X \cdot C(A) = Y \cdot C(A)$, which is what we needed to show. $\square$

Corollary 3. *If $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{M}_2(F)$, $(b, c) \neq (0, 0)$, the eigenvalues of A are distinct and belong to F, and F is infinite, then the conjugacy class $\mathrm{Cl}(A)$ is infinite.*

The final case, when A satisfies $(b, c) \neq (0, 0)$ and has only a single eigenvalue r in F, is the most subtle. The Theory of Jordan Canonical Forms, there is a matrix $Z \in \mathrm{GL}_2(F)$ so that

$$ZAZ^{-1} = \begin{pmatrix} r & 1 \\ 0 & r \end{pmatrix}.$$

Because the map $X \to ZXZ^{-1}$ is an automorphism of $\mathrm{GL}_2(F)$ we may therefore assume $A = \left(\begin{smallmatrix} r & 1 \\ 0 & r \end{smallmatrix}\right)$. In this case, elements of $C(A)$ have the form

$$\lambda I + \mu A = \lambda I + \mu r I + \mu \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = (\lambda + \mu r) I + \mu \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

The restriction $(\lambda, \mu) \in P(r, r)$ means that μ is free to take on any value while $\lambda + \mu r \neq 0$. We conclude that

$$C(A) = \left\{ \begin{pmatrix} \alpha & \beta \\ 0 & \alpha \end{pmatrix} \middle| \, \alpha, \beta \in F, \alpha \neq 0 \right\}. \tag{6}$$

We now make an important observation.

Lemma 3. *If $X \in \mathrm{GL}_2(F)$ is upper triangular, so is every matrix in $X \cdot C(A)$. Otherwise, $X \cdot C(A)$ contains no upper triangular matrices.*

Proof. Since the set of upper triangular matrices in $\mathrm{GL}_2(F)$ is a subgroup, the first statement follows in light of (6). As for the second, if $X = \left(\begin{smallmatrix} * & * \\ x & * \end{smallmatrix}\right)$ with $x \neq 0$, then

$$X \begin{pmatrix} \alpha & \beta \\ 0 & \alpha \end{pmatrix} = \begin{pmatrix} * & * \\ \alpha x & * \end{pmatrix},$$

and $\alpha x \neq 0$ since $\alpha \neq 0$. $\square$

Now let $X = \left(\begin{smallmatrix} x & y \\ z & w \end{smallmatrix}\right) \in \mathrm{GL}_2(F)$ and suppose first that $z \neq 0$. Then there is an $\alpha \in F^\times$ so that $\alpha z = 1$. Let $\beta = -\alpha^2 w$. Then

$$X \begin{pmatrix} \alpha & \beta \\ 0 & \alpha \end{pmatrix} = \begin{pmatrix} u & v \\ 1 & 0 \end{pmatrix} \tag{7}$$

for some $u, v \in F$, $v \in F^\times$. We claim that $\left(\begin{smallmatrix} u & v \\ 1 & 0 \end{smallmatrix}\right)$ is the only element of the coset $X \cdot C(A)$ of this form. Indeed, $\left(\begin{smallmatrix} p & q \\ 1 & 0 \end{smallmatrix}\right)$ is in the same coset if and only if

$$\begin{pmatrix} p & q \\ 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} u & v \\ 1 & 0 \end{pmatrix} = -\frac{1}{q} \begin{pmatrix} 0 & -q \\ -1 & p \end{pmatrix} \begin{pmatrix} u & v \\ 1 & 0 \end{pmatrix} = -\frac{1}{q} \begin{pmatrix} -q & 0 \\ p - u & -v \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ (u - p)/q & v/q \end{pmatrix}$$

belongs to $C(A)$. But by (6) this can only happen when $u = p$ and $v = q$. This proves the next statement.

Lemma 4. *The non-upper-triangular cosets of $C(A)$ are parametrized by the matrices $\left(\begin{smallmatrix} u & v \\ 1 & 0 \end{smallmatrix}\right) \in \mathrm{GL}_2(F)$.*

At last, consider the case when $X = \left(\begin{smallmatrix} x & 0 \\ z & w \end{smallmatrix}\right) \in \mathrm{GL}_2(F)$. Let $\alpha = 1/x$ and $\beta = -\alpha^2 z$. Then $X \left(\begin{smallmatrix} \alpha & \beta \\ 0 & \alpha \end{smallmatrix}\right)$ has the form

$$\begin{pmatrix} 1 & 0 \\ 0 & u \end{pmatrix}.$$

A computation similar to that above shows that no two matrices of this form can belong to the same coset of $C(A)$, proving our last lemma.

Lemma 5. *The upper-triangular cosets of $C(A)$ are parametrized by the matrices $\left(\begin{smallmatrix} 1 & 0 \\ 0 & u \end{smallmatrix}\right) \in \mathrm{GL}_2(F)$.*

Lemmas 4 and 5 immediately imply our penultimate theorem.

Theorem 5. *Suppose $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{M}_2(F)$, $(b, c) \neq (0, 0)$ and that A has a single repeated eigenvalue belonging to F. Then there is a bijection $\mathrm{GL}_2(F)/C(A) \to F^\times \cup F \times F^\times$.*

Corollary 4. *If $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{M}_2(F)$, $(b, c) \neq (0, 0)$, A has a single repeated eigenvalue belonging to F, and F is infinite, then the conjugacy class $\mathrm{Cl}(A)$ is infinite.*

Corollaries 1, 2, 3 and 4 together now finally prove Theorem 1.

As closing commentary, we mention that we used the theory of the Jordan Canonical Form in the final case only out of desperation. A more elementary approach, along the lines of the other cases, would have been preferable, if for no other reason than aesthetics. Or we could have worked with canonical forms from the outset. In the case that A has no eigenvalues in F, there's no real savings. But when A's eigenvalues are distinct and belong to F we are immediately reduced to the case of non-scalar diagonal matrices, which were handled earlier. One disadvantage to this approach, however, is that we lose the explicit descriptions of the cosets $X \cdot C(A)$, since these are only obtained for the canonical forms, not the matrices themselves (as in the final case above).