Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 478-493, September 2015

ELA

CONJUGACY CLASSES OF TORSION IN $GL_N(\mathbb{Z})$*

QINGJIE YANG†

Abstract. The problem of integral similarity of block-triangular matrices over the ring of integers is connected to that of finding representatives of the classes of an equivalence relation on general integer matrices. A complete list of representatives of conjugacy classes of torsion in the 4×4 general linear group over ring of integers is given. There are 45 distinct such classes and each torsion element has order of $1, 2, 3, 4, 5, 6, 8, 10$ or 12.

Key words. General linear group, Ring of integers, Integral similarity, Direct sum, Torsion, Cyclotomic polynomial.

AMS subject classifications. 53D30, 15A36.

1. Introduction. The problem that we consider in this paper is the determination of the conjugacy classes of torsion matrices in the $n \times n$ general linear group over $\mathbb{Z}$, the ring of integers.

Let $M_{n\times m}(\mathbb{Z})$ be the set of $n\times m$ matrices over $\mathbb{Z}$. For a matrix $A \in M_{n\times m}(\mathbb{Z})$, the transpose of A is denoted by A^{T}. When $n = m$ we simply denote $M_{n\times m}(\mathbb{Z})$ by $M_n(\mathbb{Z})$. Let I_n be the identity matrix in $M_n(\mathbb{Z})$.

A unimodular matrix of size n is an $n\times n$ integer matrix having determinant $+1$ or -1. The general linear group of size n over $\mathbb{Z}$, denoted by $GL_n(\mathbb{Z})$, is the set of unimodular matrices in $M_n(\mathbb{Z})$ together with the operation of ordinary matrix multiplication. That is,

$$GL_n(\mathbb{Z}) = \{A \in M_n(\mathbb{Z}) \big| |A| = \pm 1\},$$

where $|A|$ is the determinant of A. An element T of $GL_n(\mathbb{Z})$ is a torsion element if it has finite order, i.e., if there is a positive integer m such that $A^m = I$. A d-torsion element is a torsion element that has order d.

Two matrices A, B of $M_n(\mathbb{Z})$ are conjugates or integrally similar, denoted by $A \sim B$, if there is a matrix $Q \in GL_n(\mathbb{Z})$ such that $B = Q^{-1}AQ$.

Finding finite groups or torsion of integral matrices up to conjugation has a long history, see [5].

*Received by the editors on April 2, 2013. Accepted for publication on June 28, 2015. Handling Editor: Bryan L. Shader.
†School of Information, Renmin University of China, Beijing, 100872, China (yangqj@ruc.edu.cn).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 478-493, September 2015

ELA

Given a matrix $A \in M_n(\mathbb{Z})$, we denote the characteristic polynomial of A by

$$f_{_A}(x) = |xI - A|.$$

If $A \in GL_n(\mathbb{Z})$, then $f_{_A}(x)$ is a monic polynomial with constant term $f(0) = \pm 1$.

Let $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{Z}[x]$, the polynomial ring over $\mathbb{Z}$, with $f(0) = \pm 1$. The set of all integral matrices with characteristic polynomial $f(x)$ is denoted by M_f. That is,

$$M_f = \{A \in GL_n(\mathbb{Z}) \mid f_{_A}(x) = f(x)\}.$$

Let $\mathcal{M}_f$ be the set of all conjugacy classes of matrices in M_f. The size of $\mathcal{M}_f$ is denoted by $|\mathcal{M}_f|$.

The matrix C_f given by

$$C_f = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix}$$

is known as the companion matrix of $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$. It is known that $C_f \in M_f$, and thus, $\mathcal{M}_f \neq \emptyset$.

For any $A \in GL_n(\mathbb{Z})$, we use $C(A)$ to denote its centralizer in $GL_n(\mathbb{Z})$. If A is similar to the companion matrix of a polynomial over $\mathbb{Z}$, then its centralizer is

$$C(A) = \{g(A) \in GL_n(\mathbb{Z}) \mid g(x) \in \mathbb{Z}[x] \text{ is of degree less than } n\}.$$

For $A \in M_{n\times m}(\mathbb{Z})$ and $B \in M_{s\times t}(\mathbb{Z})$, the direct sum of A and B is

$$A \oplus B = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \in M_{(n+s)\times(m+t)}(\mathbb{Z}).$$

Obviously, $A \oplus B$ is a unimodular matrix if and only if both A and B are unimodular matrices.

A matrix $A \in GL_n(\mathbb{Z})$ is decomposable if it is conjugate to a direct sum of two matrices which have smaller sizes; otherwise, A is said to be indecomposable.

The characteristic polynomial of a decomposable matrix is reducible over $\mathbb{Z}$, but the converse is not true.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 478-493, September 2015

ELA

In this paper, we mainly consider the integral similarity problem for upper block-triangular matrices of the form

$$\begin{bmatrix} A & X \\ 0 & B \end{bmatrix}, \tag{1.1}$$

where A, B are unimodular matrices with coprime minimal polynomials. Our results are based on the following lemmas. We state them without proof.

Lemma 1.1. *Each A in $M_n(\mathbb{Z})$ is integrally similar to a block-triangular matrix*

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ 0 & A_{22} & \cdots & A_{2r} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_{rr} \end{bmatrix},$$

where the characteristic polynomial of A_{ii} is irreducible, $1 \le i \le r$. The block-triangularization can be attained with the diagonal blocks in any prescribed order.

See [6, 9] for a proof.

Lemma 1.2. *Let $A \in GL_n(\mathbb{Z})$ have irreducible minimal polynomial $p(x)$ with $|\mathcal{M}_p| = 1$. Then A is integrally similar to*

$$C_p \oplus C_p \oplus \cdots \oplus C_p,$$

where C_p is the companion matrix of $p(x)$. That is $|\mathcal{M}_{p^k}| = 1$.

See also [6].

Consider two monic polynomials

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0 \quad \text{and} \quad g(x) = x^m + b_{m-1}x^{m-1} + \cdots + b_0$$

in $\mathbb{Z}[x]$. Recall that the resultant of $f(x)$ and $g(x)$ is the determinant

$$R(f,g) = \begin{vmatrix} 1 & a_{n-1} & \cdot & \cdot & \cdot & a_0 & 0 & \cdot & \cdots & \cdot \\ 0 & 1 & a_{n-1} & \cdot & \cdot & \cdot & a_0 & 0 & \cdots & \cdot \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdot & \cdot & 0 & 1 & a_{n-1} & \cdot & \cdot & \cdot & \cdots & a_0 \\ 1 & b_{m-1} & \cdot & \cdot & b_0 & 0 & \cdot & \cdot & \cdots & \cdot \\ 0 & 1 & b_{m-1} & \cdot & \cdot & b_0 & 0 & \cdot & \cdots & \cdot \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdot & \cdot & \cdot & 0 & 1 & b_{m-1} & \cdot & \cdot & \cdots & b_0 \end{vmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix}1\\1\\1\\1\end{matrix}}\right\} m \text{ rows} \\ \left.\vphantom{\begin{matrix}1\\1\\1\\1\end{matrix}}\right\} n \text{ rows} \end{array}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 478-493, September 2015

ELA

It is known that $f(x)$ and $g(x)$ are coprime if and only if $R(f, g) \neq 0$.

The following theorem, which is a corollary of Lemma 3.1, gives a sufficient condition for decomposability.

Theorem 1.3. *Let $A \in M_n(\mathbb{Z})$ with its characteristic polynomial a product of two coprime polynomials whose resultant is ± 1. Then A is decomposable.*

To explain our results, we need to develop some notation. For any $A \in GL_n(\mathbb{Z})$, we use A^+, A^- to denote the block matrices $\begin{bmatrix} A & e \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} A & e \\ 0 & -1 \end{bmatrix}$ respectively, where $e = (1, 0, \ldots, 0)^{\mathrm{T}} \in M_{n\times 1}(\mathbb{Z})$. Clearly, A^+, $A^- \in GL_{n+1}(\mathbb{Z})$. Also, we let C_n denote the companion matrix of $\Phi_n(x)$, the nth cyclotomic polynomial of degree $\varphi(n)$, where φ is the Euler's totient function.

Our results are given in following theorems. We will prove them in Section 3.

Theorem 1.4. *Let $n > 1$ and $A = C_n \oplus C_n \oplus \cdots \oplus C_n$, the direct sum of s-copies of C_n. Let*

$$M = \begin{bmatrix} A & X \\ 0 & I_m \end{bmatrix}, \quad \textit{where} \quad X \in M_{s\varphi(n)\times m}(\mathbb{Z}).$$

1. *If $n = p^k$, where p is a prime number and $k \geq 1$, then*

$$M \sim \underbrace{C_n^+ \oplus \cdots \oplus C_n^+}_{t} \oplus \underbrace{C_n \oplus \cdots \oplus C_n}_{s-t} \oplus I_{m-t}, \tag{1.2}$$

where the number t of C_n^+ satisfies $0 \leq t \leq \min(s, m)$ and is uniquely determined by M.

2. *If n is not a power of a prime, then $M \sim A \oplus I_m$.*

The special case $n = 2$ was established by Hua and Reiner, [4]. Similarly, we have the following.

Theorem 1.5. *Let $n > 2$ and $A = C_n \oplus C_n \oplus \cdots \oplus C_n$, the direct sum of s-copies of C_n. Let*

$$M = \begin{bmatrix} A & X \\ 0 & -I_m \end{bmatrix}, \quad \textit{where} \quad X \in M_{s\varphi(n)\times m}(\mathbb{Z}).$$

1. *If $n = 2p^k$, where p is a prime and $k \geq 1$, then*

$$M \sim \underbrace{C_n^- \oplus \cdots \oplus C_n^-}_{t} \oplus \underbrace{C_n \oplus \cdots \oplus C_n}_{s-t} \oplus (-I_{m-t}),$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 478-493, September 2015

where the number t of C_n^- satisfies $0 \le t \le \min(s, m)$ and is uniquely determined by M.

2. *If $n \ne 2p^k$, then $M \sim A \oplus (-I_m)$.*

A complete conjugacy list of torsion in $GL_2(\mathbb{Z})$ is already known.

LEMMA 1.6. *All torsion in $GL_2(\mathbb{Z})$ up to conjugation are given in the following table together with the centralizers and minimal polynomials of the conjugacy class representatives*

order	A	$C(A)$	$m_A(x)$
1	I	$GL_2(\mathbb{Z})$	$\Phi_1(x) = (x-1)$
	$-I$	$GL_2(\mathbb{Z})$	$\Phi_2(x) = (x+1)$
2	K	$\pm I, \pm K$	$(x-1)(x+1)$
	U	$\pm I, \pm U$	$(x-1)(x+1)$
3	W	$\pm I, \pm W, \pm(I+W)$	$\Phi_3(x) = x^2 + x + 1$
4	J	$\pm I, \pm J$	$\Phi_4(x) = x^2 + 1$
6	$-W$	$\pm I, \pm W, \pm(I+W)$	$\Phi_6(x) = x^2 - x + 1$

where $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $K = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $U = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$, $W = C_3 = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$, $J = C_4 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

For a proof, see [8].

Although the maximal finite subgroups of $GL_4(\mathbb{Z})$ up to conjugation have been determined by Dade [1], a complete set of non-conjugate classes of torsion in $GL_4(\mathbb{Z})$ is of value. We have solved the closely related problem of classifying the conjugacy classes of elements of finite order in the 4×4 symplectic group over $\mathbb{Z}$, see [12].

If A is a d-torsion element in $GL_4(\mathbb{Z})$, then its minimal polynomial $m_A(x)$ is a factor of $x^d - 1$, i.e., $m_A(x)$ is a product of cyclotomic polynomials. It is easy to check that any torsion element A in $GL_4(\mathbb{Z})$ has order 1, 2, 3, 4, 5, 6, 8, 10 or 12. Note that $\varphi(5) = \varphi(8) = \varphi(10) = \varphi(12) = 4$ and then $\Phi_n(x)$ is a quartic polynomial for $n = 5, 8, 10$ or 12. According to Latimer, MacDuffee and Taussky [11], if the characteristic polynomial of A is $\Phi_5(x)$, $\Phi_8(x)$, $\Phi_{10}(x)$ or $\Phi_{12}(x)$, then A is conjugate to C_5, C_8, C_{10} or C_{12} respectively since $\mathbb{Q}(\zeta_m)$, where ζ_m is a primitive mth root of unity, has class number one for all positive integers m less than 12. We reduce the problem to the case that the characteristic polynomial of A is reducible. The cases where $m_A(x)$ is an irreducible quadratic polynomial, that is $m_A(x) = \Phi_n(x)$, $n = 3, 4$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 478-493, September 2015

ELA

or 6, can be solved by Lemma 1.2 and Lemma 1.6. Furthermore, the case where $A^2 = I$ was solved in complete generality by Hua and Reiner [4]. We only need to consider the cases where $m_A(x)$ is one of the following: $(x^2+1)(x-1)$, $(x^2+1)(x+1)$, $(x^2 \pm x+1)(x-1)$, $(x^2 \pm x+1)(x+1)$, $(x^2 \pm x+1)(x^2+1)$, $(x^2+x+1)(x^2-x+1)$, $(x^2-1)(x^2+1)$ and $(x^2-1)(x^2 \pm x+1)$. As a consequence, by Lemma 1.1, A is integrally similar to a block upper triangular matrix with different diagonal 2×2 blocks chosen from Lemma 1.6. By applying Theorem 1.4 or Theorem 1.5, we can solve the problem for the first four cases where one and only one of ± 1 is an eigenvalue. The case where $m_A(x) = (x^2 \pm x+1)(x^2+1)$ can be solved by Theorem 1.3. For the remaining three cases, we have the following result.

THEOREM 1.7. *All elements in $GL_4(\mathbb{Z})$ with some given reducible characteristic polynomials $f(x)$ up to conjugation are listed below:*

1. *When $f(x) = (x^2-1)(x^2+\lambda x+1)$, where $\lambda = \pm 1$,*

$$\mathcal{M}_f = \left\{ \lambda \begin{bmatrix} K & 0 \\ 0 & W \end{bmatrix},\ \lambda \begin{bmatrix} K & E \\ 0 & W \end{bmatrix},\ \lambda \begin{bmatrix} U & 0 \\ 0 & W \end{bmatrix},\ \lambda \begin{bmatrix} U & E \\ 0 & W \end{bmatrix} \right\};$$

2. *When $f(x) = (x^2-1)(x^2+1)$,*

$$\mathcal{M}_f = \left\{ \begin{bmatrix} K & 0 \\ 0 & J \end{bmatrix}, \begin{bmatrix} K & E \\ 0 & J \end{bmatrix}, \begin{bmatrix} K & I \\ 0 & J \end{bmatrix}, \begin{bmatrix} K & I-E \\ 0 & J \end{bmatrix}, \begin{bmatrix} U & 0 \\ 0 & J \end{bmatrix}, \begin{bmatrix} U & E \\ 0 & J \end{bmatrix}, \begin{bmatrix} U & I \\ 0 & J \end{bmatrix} \right\};$$

3. *When $f(x) = (x^2+x+1)(x^2-x+1)$,*

$$\mathcal{M}_f = \left\{ \begin{bmatrix} W & 0 \\ 0 & -W \end{bmatrix},\ \begin{bmatrix} W & E \\ 0 & -W \end{bmatrix} \right\},$$

where $E = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

To prove our results we need to develop some new tools. The idea we use in this paper comes from Roth's Theorem [10]. For a general form of Roth's Theorem, see [2] or [3]. We shall generalize Roth's Theorem to the integral similarity problem for upper block-triangular matrices of the form (1.1) with the diagonal blocks have coprime characteristic polynomials. In Section 2 we shall study (A, B)-equivalence in $M_{n\times m}(\mathbb{Z})$. Then we transform our similarity problem to the problem finding (A, B)-equivalent classes and prove our results in Section 3. We use the program Mathematica to calculate some of results in this paper.

2. (A, B)-equivalence. Let $A \in GL_n(\mathbb{Z})$, $B \in GL_m(\mathbb{Z})$ and suppose that their respective characteristic polynomials $f(x)$ and $g(x)$ are coprime. We define a linear

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 478-493, September 2015

transformation ψ on $M_{n\times m}(\mathbb{Z})$ by

$$\psi : M_{n\times m}(\mathbb{Z}) \to M_{n\times m}(\mathbb{Z}), \quad T \mapsto AT - TB.$$

Since $f(x)$, $g(x)$ are coprime, ψ is injective. Let $\langle A, B\rangle$ be the image of ψ, that is

$$\langle A, B\rangle = \{AT - TB \mid T \in M_{n\times m}(\mathbb{Z})\}.$$

By choosing a suitable basis, the matrix of ψ is $A \otimes I_m - I_n \otimes B^{\mathrm{T}}$, where $\otimes$ is the Kronecker product of matrices. Then the determinant of ψ is equal to $R(f, g)$, the resultant of $f(x)$ and $g(x)$. Let $r = |R(f, g)|$, the absolute value of $R(f, g)$. The quotient module $M_{n\times m}(\mathbb{Z})/\langle A, B\rangle$, called the cokernel of ψ and denoted by $\operatorname{coker}\psi$, is of order r. Let $X \in M_{n\times m}(\mathbb{Z})$. Then an equivalent condition for $X \in \langle A, B\rangle$ is that the Sylvester equation

$$AT - TB = X \tag{2.1}$$

has a unique integral solution for matrix T. Clearly, if $X \equiv 0 \pmod r$, then $X \in \langle A, B\rangle$.

LEMMA 2.1. *Let C_f be the companion matrix of $f(x)$ of degree n and $\alpha \in M_{n\times 1}(\mathbb{Z})$ be an integral column vector. Then $\alpha \in \langle C_f, I_1\rangle$ if and only if the integer number $f(1)$ divides $\ell(\alpha)$, the sum of components of α.*

Proof. Let $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ and $\alpha = (c_1, c_2, \ldots, c_n)^{\mathrm{T}}$. In this case,

$$\langle C_f, I_1\rangle = \{(C_f - I)X \mid X = (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in M_{n\times 1}(\mathbb{Z})\}.$$

So, $\alpha \in \langle C_f, I_1\rangle$ if and only if the system of linear equations

$$\begin{cases} -x_1 & & & -a_0x_n = c_1 \\ x_1 - x_2 & & & -a_1x_n = c_2 \\ & x_2 - x_3 & & -a_2x_n = c_3 \\ & & & \vdots \\ & & x_{n-1} & -(1 + a_{n-1})x_n = c_n \end{cases}$$

has an integral solution. This system is equivalent to the following system,

$$\begin{cases} -x_1 & -a_0x_n = c_1 \\ -x_2 & -(a_0 + a_1)x_n = c_1 + c_2 \\ & \vdots \\ -x_{n-1} & -(a_0 + a_1 + \cdots + a_{n-1})x_n = c_1 + c_2 + \cdots + c_{n-1} \\ & -(1 + a_0 + a_1 + \cdots + a_{n-1} + a_n)x_n = c_1 + c_2 + \cdots + c_{n-1} + c_n \end{cases}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 478-493, September 2015

ELA

which has an integer solution for $x_1, x_2, \ldots, x_n$ if and only if

$$f(1) = (1 + a_0 + a_1 + \cdots + a_{n-1} + a_n) | (c_1 + c_2 + \cdots + c_{n-1} + c_n) = \ell(\alpha).$$

Thus, $\alpha \in \langle C_f, I_1 \rangle$ if and only if $f(1)$ divides $\ell(\alpha)$. □

Similarly, $\alpha \in \langle C_f, -I_1 \rangle$ if and only if $f(-1)$ divides $c_1 - c_2 + \cdots + (-1)^{n-1} c_n$, the alternating sum of components of α.

We now define an equivalence relation on $M_{n\times m}(\mathbb{Z})$.

Definition 2.2. Let $X, Y \in M_{n\times m}(\mathbb{Z})$ be any two matrices. X and Y are said to be (A, B)-*equivalent*, denoted by $X \cong Y \pmod{A, B}$ or $X \cong Y$ for short, if there exist $P \in C(A)$ and $Q \in C(B)$ such that $XQ - PY \in \langle A, B \rangle$. The set of (A, B)-equivalent classes is denoted by $\mathcal{S}(A, B)$.

It is obvious that if $X - Y \in \langle A, B \rangle$, then $X \cong Y \pmod{A, B}$. But the converse is not necessarily true.

Lemma 2.3. *Let $X, Y \in M_{n\times m}(\mathbb{Z})$. Then*

1. $X \cong Y \pmod{A, B}$ *if and only if* $X - PYQ^{-1} \in \langle A, B \rangle$ *for some* $P \in C(A)$, $Q \in C(B)$*;*
2. $X \cong PXQ \pmod{A, B}$*, where* $P \in C(A)$ *and* $Q \in C(B)$*. In particular,* $X \cong -X$*;*
3. *If* $X \equiv Y \pmod{r}$*, then* $X \cong Y \pmod{A, B}$*, where* $r = |R(f, g)|$.

Proof. By definition, $X \cong Y$ if and only if there exist $P \in C(A)$ and $Q \in C(B)$ such that

$$XQ - PY = AT - TB \tag{2.2}$$

for some $T \in M_{n\times m}(\mathbb{Z})$. Since Q commutes B, so does Q^{-1}, and then (2.2) is equivalent to

$$X - PYQ^{-1} = A(TQ^{-1}) - (TQ^{-1})B.$$

Therefore, Part 1 is true.

Part 2 is obtained by $PXQ - (PXQ) = 0$, and $(-I_n) \in C(A)$.

Part 3 is true since $X - Y \in \langle A, B \rangle$, whenever $X \equiv Y \pmod{r}$. □

Suppose that the cokernel of ψ has a set of representative

$$\operatorname{coker} \psi = \{\overline{S}_1, \overline{S}_2, \ldots, \overline{S}_r \mid S_i \in M_{n\times m}(\mathbb{Z})\},$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 478-493, September 2015

where $\overline{S}_i = S_i + \langle A, B\rangle$, $i = 1, \dots, r$, are all cosets of $\langle A, B\rangle$ in $M_{n\times m}(\mathbb{Z})$. We define a group action of $C(A) \times C(B)$ on $\operatorname{coker}\psi$ given by

$$P\overline{S}_iQ = \overline{PS_iQ}$$

for any $(P, Q) \in C(A) \times C(B)$. The action is well defined since $P\langle A, B\rangle = \langle A, B\rangle = \langle A, B\rangle Q$. The set of orbits is denoted by $\operatorname{coker}\psi/C(A) \times C(B)$.

LEMMA 2.4. *Let $X, Y \in M_{n\times m}(\mathbb{Z})$. Then a necessary and sufficient condition for $X \cong Y$ is that $\overline{X}$ and $\overline{Y}$ are in the same orbit of the action. That is*

$$\mathcal{S}(A, B) = \operatorname{coker}\psi/C(A) \times C(B).$$

Proof. Note that $X \cong Y$ if and only if $X - PYQ \in \langle A, B\rangle$ for some $(P, Q) \in C(A) \times C(B)$, which is equivalent to $\overline{X} = P\overline{Y}Q$. Therefore, $X \cong Y$ if and only if $\overline{X}$ and $\overline{Y}$ are in the same orbit. $\square$

From Lemma 2.4, when $r = 1$, $|\mathcal{S}(A, B)|$, the class number of (A, B)-equivalence, is equal to 1. If $r > 1$, then $1 < |\mathcal{S}(A, B)| \le r$, because $\langle A, B\rangle$ is a fixed point for all elements in $C(A) \times C(B)$. In particular, if $r = 2$, $|\mathcal{S}(A, B)| = 2$.

Let $M = A\oplus A\oplus\cdots\oplus A$ be the direct sum of s-copies of A, and $N = B\oplus B\oplus\cdots\oplus B$ be the direct sum of t-copies of B. Let

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1t} \\ X_{21} & X_{22} & \cdots & X_{2t} \\ \vdots & \vdots & & \vdots \\ X_{s1} & X_{s2} & \cdots & X_{st} \end{bmatrix} \in M_{sn\times tm}(\mathbb{Z}),$$

where $X_{ij} \in M_{n\times m}$. Then we have the following.

LEMMA 2.5. *$X \in \langle M, N\rangle$ if and only if $X_{ij} \in \langle A, B\rangle$, for $i = 1, \dots, s$; $j = 1, \dots, t$.*

It is also easy to prove that the following block row/column elementary operations on X preserve the equivalence:

1. Row/column switching
   $R_i \leftrightarrow R_j$: switch the i-th row blocks and the j-th row blocks;
   $C_i \leftrightarrow C_j$: switch the i-th column blocks and the j-th column blocks.
2. Row/column multiplication
   $R_i \to PR_i$: left-multiply the i-th row blocks by P, where $P \in C(A)$;
   $C_i \to C_iQ$: right-multiply the i-th column blocks by Q, where $Q \in C(B)$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 478-493, September 2015

ELA

3. Row/column addition
$R_i \to R_i + PR_j$: add the j-th row blocks left-multiplied by P to the i-th row;
$C_i \to C_i + C_jQ$: add the j-th column blocks right-multiplied by Q to the i-th column, where $P \in M_n(\mathbb{Z})$ commutes A, and $Q \in M_m(\mathbb{Z})$ commutes B.

3. Proofs. Before the proof, we need to give a connection of (A, B)-equivalence with integral similarity of matrices of the form (1.1).

LEMMA 3.1. *Let $A \in M_n(\mathbb{Z})$, $B \in M_m(\mathbb{Z})$ and suppose that they have co-prime characteristic polynomials. Then* $\begin{bmatrix} A & X \\ 0 & B \end{bmatrix} \sim \begin{bmatrix} A & Y \\ 0 & B \end{bmatrix}$ *if and only if* $X \cong Y$ $(\mathrm{mod}\ A, B)$.

Proof. First suppose that $\begin{bmatrix} A & X \\ 0 & B \end{bmatrix} \sim \begin{bmatrix} A & Y \\ 0 & B \end{bmatrix}$. There is $Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \in GL_{n+m}(\mathbb{Z})$ such that

$$\begin{bmatrix} A & X \\ 0 & B \end{bmatrix} \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} A & Y \\ 0 & B \end{bmatrix}.$$

We get that

$$AQ_{11} + XQ_{21} = Q_{11}A, \tag{3.1}$$
$$AQ_{12} + XQ_{22} = Q_{11}Y + Q_{12}B, \tag{3.2}$$
$$BQ_{21} = Q_{21}A, \tag{3.3}$$
$$BQ_{22} = Q_{21}Y + Q_{22}B. \tag{3.4}$$

By hypothesis, (3.3) implies $Q_{21} = 0$. Then $Q = \begin{bmatrix} Q_{11} & Q_{12} \\ 0 & Q_{22} \end{bmatrix}$, and (3.1), (3.4) say that $Q_{11} \in C(A)$ and $Q_{22} \in C(B)$. Thus, (3.2) means $X \cong Y$.

Conversely, if $X \cong Y$ by some $(P, Q) \in C(A) \times C(B)$ and $T \in M_{n\times m}(\mathbb{Z})$, then $\begin{bmatrix} A & X \\ 0 & B \end{bmatrix} \sim \begin{bmatrix} A & Y \\ 0 & B \end{bmatrix}$ via the similarity $\begin{bmatrix} P & -T \\ 0 & Q \end{bmatrix}$. □

According to Lemma 3.1, integral similarity problem for block-triangular matrices of the form (1.1) can be transformed to the problem of finding (A, B)-equivalent classes.

Now we can prove our theorems.

Proof of Theorem 1.4. For any matrix $X \in M_{s\varphi(n)\times m}(\mathbb{Z})$, we write X as a block

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 478-493, September 2015

ELA

matrix

$$X = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2m} \\ \vdots & \vdots & & \vdots \\ \alpha_{s1} & \alpha_{s2} & \cdots & \alpha_{sm} \end{bmatrix},$$

where $\alpha_{ij} \in M_{\varphi(n)\times 1}(\mathbb{Z})$. Then by Lemma 2.5, $X \in \langle A, I_m\rangle$ if and only if $\alpha_{ij} \in \langle C_n, I_1\rangle$, which is equivalent to that $\Phi_n(1)$ is a factor of $\ell(\alpha_{ij})$ by Lemma 2.1, for all α_{ij}. Note that for $n > 2$, see [7],

$$\Phi_n(1) = \begin{cases} p, & n = p^k,\ p \text{ prime},\ k \geq 1 \\ 1, & \text{otherwise.} \end{cases} \tag{3.5}$$

Case 1. $n = p^k$ is a power of prime p, then $\Phi_n(1) = p$.

We first show that $\mathcal{S}(C_n, I_1) = \{\overline{0}, \overline{e}\}$. For any $\alpha \in M_{\varphi(n)\times 1}(\mathbb{Z})$, let $b = \ell(\alpha)$.

$$\ell(\alpha - be) = \ell(\alpha) - b\ell(e) = b - b \cdot 1 = 0.$$

So, $\alpha \cong be \pmod{C_n, I_1}$. We only need to show $be \cong e \pmod{C_n, I_1}$ provided $p \nmid b$. Without loss of generality, we assume $0 < b \leq p-1$. Let $P = I + C_n + C_n^2 + \cdots + C_n^{b-1}$. Then $P(I - C_n) = I - C_n^b$, and thus, the determinant of P satisfies

$$|P||I - C_n| = |I - C_n^b|.$$

Since $(b, p) = 1$, C_n and C_n^b have the same characteristic polynomial $\Phi_n(x)$. So,

$$|I - C_n| = |I - C_n^b| = \Phi_n(1) = p.$$

Therefore, $|P| = 1$ and $P \in GL_{\varphi(n)}(\mathbb{Z})$. Also, P commutes with C_n. It is easy to verify that $\ell(Pe) = b$, and then $\ell(be - Pe) = 0$. So $be \cong e \pmod{C_n, I_1}$, and hence, $\mathcal{S}(C_n, I_1) = \{\overline{0}, \overline{e}\}$.

Now suppose that there is $\alpha_{ij} \notin \langle C_n, I_1\rangle$. We can use row or column switchings move it to the left-top position. So, we may assume that $\alpha_{11} \notin \langle C_n, I_1\rangle$. There is $P \in C_n$ such that $P\alpha_{11} - e \in \langle C_n, I_1\rangle$. Then by a row multiplication and Lemma 2.5,

$$X \xrightarrow{R_1 \to PR_1} \begin{bmatrix} P\alpha_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\ \vdots & \vdots & & \vdots \\ \beta_{s1} & \beta_{s2} & \cdots & \beta_{sm} \end{bmatrix} \cong \begin{bmatrix} e & \beta_{12} & \cdots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\ \vdots & \vdots & & \vdots \\ \beta_{s1} & \beta_{s2} & \cdots & \beta_{sm} \end{bmatrix}.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 478-493, September 2015

ELA

By row additions, $R_i \to R_i - \ell(\beta_{i1})R_1$, and column additions, $C_j \to C_j - \ell(\beta_{1j})C_1$, we get

$$\begin{bmatrix} e & \beta_{12} & \cdots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\ \vdots & \vdots & & \vdots \\ \beta_{s1} & \beta_{s2} & \cdots & \beta_{sm} \end{bmatrix} \longrightarrow \begin{bmatrix} e & \gamma_{12} & \cdots & \gamma_{1m} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2m} \\ \vdots & \vdots & & \vdots \\ \gamma_{s1} & \gamma_{s2} & \cdots & \gamma_{sm} \end{bmatrix} \cong \begin{bmatrix} e & 0 & \cdots & 0 \\ 0 & \gamma_{22} & \cdots & \gamma_{2m} \\ \vdots & \vdots & & \vdots \\ 0 & \gamma_{s2} & \cdots & \gamma_{sm} \end{bmatrix},$$

where $\gamma_{i1} = \beta_{i1} - \ell(\beta_{i1})e$, $\gamma_{1j} = \beta_{1j} - \ell(\beta_{1j})e \in \langle C_n, I_1 \rangle$. Continue this process to the submatrix obtained by deleting first row block and first column, and so on, we obtain

$$X \cong \begin{bmatrix} Y & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{where} \quad Y = \underbrace{e \oplus e \oplus \cdots \oplus e}_{t}$$

for some $1 \le t \le \min(s, m)$. Therefore, by Lemma 3.1,

$$\begin{bmatrix} A & X \\ 0 & I_m \end{bmatrix} \sim \begin{bmatrix} C_n & & & & & & e & & & \\ & \ddots & & & & & & \ddots & & \\ & & C_n & & & & & & e & \\ & & & C_n & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & C_n & & & & \\ & & & & & & 1 & & & \\ & & & & & & & \ddots & & \\ & & & & & & & & 1 & \\ & & & & & & & & & I_{m-t} \end{bmatrix},$$

where the number of C_n is s and the number of e is t. By some pairs of row and column switchings, we get the matrix on the right is conjugate to the matrix (1.2).

For the uniqueness, let $X_i = \underbrace{e \oplus e \oplus \cdots \oplus e}_{t_i}$, $i = 1, 2$, with $t_1 > t_2$ and suppose that

$$\begin{bmatrix} X_1 & 0 \\ 0 & 0 \end{bmatrix} \cong \begin{bmatrix} X_2 & 0 \\ 0 & 0 \end{bmatrix}.$$

There are $P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \in C(A)$, $Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \in C(B)$, where P_{11} is $t_1 \times t_2$ matrix, Q_{11} is $t_1 \times t_2$ matrix, such that

$$\begin{bmatrix} X_1 & 0 \\ 0 & 0 \end{bmatrix} Q - P \begin{bmatrix} X_2 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} X_1Q_{11} - P_{11}X_2 & X_1Q_{12} \\ -P_{21}X_2 & 0 \end{bmatrix} \in \langle A, I_m \rangle.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 478-493, September 2015

ELA

We get $X_1Q_{12} \equiv 0 \pmod p$, so $Q_{12} \equiv 0 \pmod p$. Note that the block Q_{12} is a $t_1 \times (m - t_2)$ matrix and $t_1 + (m - t_2) > m$, the size of Q. Therefore, the determinant of Q satisfies $|Q| \equiv 0 \pmod p$. This is impossible since Q is an unimodular matrix. This completes the proof of uniqueness.

Case 2. n is not a power of prime. From (3.5), $\Phi_n(1) = 1$, and then $M_{\varphi(n)\times 1}(\mathbb{Z}) = \langle C_n, I_1\rangle$. There is only one (A, I_m)-equivalent class. Therefore, $M \sim A \oplus I_m$. □

The proof of Theorem 1.5 is similar. In this case, we use the fact that

$$\Phi_n(-1) = \begin{cases} p, & n = 2p^k,\ p \text{ prime}, k \geq 1 \\ 1, & \text{otherwise}, \end{cases} \tag{3.6}$$

see [7].

Proof of Theorem 1.7. By Lemma 1.1 and Lemma 3.1, we only need to calculate (A, B)-equivalent classes for some special pairs of 2×2 matrices in Lemma 1.6.

When $A = K$ and $B = W$. Clearly, $r = 3$. The linear transformation ψ is given by $\psi(T) = KT - TW$. Then the Sylvester equation (2.1) becomes

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} T - T \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

It is equivalent to the system of linear equations

$$\begin{cases} t_{11} - t_{12} = a \\ t_{11} + 2t_{12} = b \\ -t_{21} - t_{22} = c \\ t_{21} \phantom{- 2t_{22}} = d, \end{cases}$$

which has integral solutions if and only if $3|a - b$. Thus, the submodule $\langle K, W\rangle$, the image of ψ, is

$$\langle K, W\rangle = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{Z}) \middle| a \equiv b \pmod 3 \right\}.$$

It is obvious that $E, 2E \notin \langle K, W\rangle$. Therefore, $\operatorname{coker}\psi = M_2(\mathbb{Z})/\langle K, W\rangle = \{\overline{0}, \overline{E}, \overline{2E}\}$. By choosing $P = -I \in C(K)$, $Q = I \in C(W)$, we see that

$$EQ - P(2E) = \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix} \in \langle K, W\rangle,$$

and thus, $E \cong 2E \pmod{K, W}$. Note that $|\mathcal{S}(K, W)| > 1$, hence $\mathcal{S}(K, W) = \{\overline{0}, \overline{E}\}$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 478-493, September 2015

When $A = U$ and $B = W$. We also have $r = 3$. This time the Sylvester equation is equivalent to

$$\begin{cases} t_{11} - \ \ t_{12} + t_{21} = a \\ t_{11} + 2t_{12} + t_{22} = b \\ \qquad - \ \ t_{21} - t_{22} = c \\ \qquad\qquad\quad t_{21} = d. \end{cases}$$

It is clear that

$$\langle U, W \rangle = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{Z}) \middle| a + d \equiv b + c \,(\mathrm{mod}\ 3) \right\}.$$

So, $\operatorname{coker} \psi = \{\overline{0}, \overline{E}, \overline{2E}\}$ and then $\mathcal{S}(U, W) = \{\overline{0}, \overline{E}\}$.

Since $\langle -A, -B \rangle = \langle A, B \rangle$ for any A and B, we obtain that $\mathcal{S}(-K, -W) = \mathcal{S}(-U, -W) = \{\overline{0}, \overline{E}\}$.

Similarly, we have

$$\langle K, J \rangle = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{Z}) \middle| a + b \equiv c + d \equiv 0 \,(\mathrm{mod}\ 2) \right\}, \ \mathcal{S}(K, J) = \{\overline{0}, \overline{E}, \overline{I}, \overline{I - E}\},$$
$$\langle U, J \rangle = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{Z}) \middle| a + b + c \equiv c + d \equiv 0 \,(\mathrm{mod}\ 2) \right\}, \ \mathcal{S}(U, J) = \{\overline{0}, \overline{E}, \overline{I}\},$$
$$\langle W, -W \rangle = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{Z}) \middle| a + b + c \equiv a \equiv d \,(\mathrm{mod}\ 2) \right\}, \ \mathcal{S}(W, -W) = \{\overline{0}, \overline{E}\}.$$

In summary, we have the following table

A	B	$\mathcal{S}(A, B)$
K or U	W	$\overline{0}, \overline{E}$
$-K$ or $-U$	$-W$	$\overline{0}, \overline{E}$
K	J	$\overline{0}, \overline{E}, \overline{I}, \overline{I-E}$
U	J	$\overline{0}, \overline{E}, \overline{I}$
W	$-W$	$\overline{0}, \overline{E}$

We can use the results in this table, Lemma 1.1 and Lemma 3.1 to complete the proof. $\square$

From above theorems, and some simple calculations, all torsion in $GL_4(\mathbb{Z})$ up to

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 478-493, September 2015

conjugation are listed as follows:

$d=1\quad I_4;$

$d=2\quad -I_4;\quad K\oplus(-I),\quad U\oplus(-I);$

$\qquad I\oplus(-I),\quad K\oplus U,\quad U\oplus U;\quad I\oplus K,\quad I\oplus U;$

$d=3\quad W\oplus W;\quad I\oplus W,\quad \begin{bmatrix} I & E\\ 0 & W\end{bmatrix};$

$d=4\quad J\oplus J;\quad I\oplus J,\quad \begin{bmatrix} I & E\\ 0 & J\end{bmatrix};\quad (-I)\oplus J,\quad \begin{bmatrix} -I & E\\ 0 & J\end{bmatrix};$

$\qquad K\oplus J,\quad \begin{bmatrix} K & E\\ 0 & J\end{bmatrix},\quad \begin{bmatrix} K & I\\ 0 & J\end{bmatrix},\quad \begin{bmatrix} K & I-E\\ 0 & J\end{bmatrix};$

$\qquad U\oplus J,\quad \begin{bmatrix} U & E\\ 0 & J\end{bmatrix},\quad \begin{bmatrix} U & I\\ 0 & J\end{bmatrix};$

$d=5\quad C_5;$

$d=6\quad -(W\oplus W);\quad I\oplus(-W);\quad (-I)\oplus W;\quad -(I\oplus W),\quad \begin{bmatrix} -I & E\\ 0 & -W\end{bmatrix};$

$\qquad K\oplus W,\quad \begin{bmatrix} K & E\\ 0 & W\end{bmatrix};\quad U\oplus W,\quad \begin{bmatrix} U & E\\ 0 & W\end{bmatrix};$

$\qquad -(K\oplus W),\quad \begin{bmatrix} -K & E\\ 0 & -W\end{bmatrix};\quad -(U\oplus W),\quad \begin{bmatrix} -U & E\\ 0 & -W\end{bmatrix};$

$\qquad W\oplus(-W),\quad \begin{bmatrix} W & E\\ 0 & -W\end{bmatrix};$

$d=8\quad C_8;$

$d=10\quad -C_5;$

$d=12\quad C_{12};\quad J\oplus W;\quad J\oplus(-W).$

Acknowledgment. The author would like to thank the referee for helpful comments.

REFERENCES

[1] E.C. Dade. The maximal finite groups of 4×4 integral matrices. *Illinois J. Math.*, 9:99–122, 1965.
[2] R. Guralnick. Roth's theorems and decomposition of modules. *Linear Algebra Appl.*, 39:155–165, 1981.
[3] R. Guralnick. Roth's Theorems for sets of matrices. *Linear Algebra Appl.*, 71:113–117, 1985.
[4] L.K. Hua and I. Reiner. Automorphisms of the unimodular group. *Trans. Amer. Math. Soc.*, 71(3):331–348, 1951.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 478-493, September 2015

ELA

[5] J. Kuzmanovich and A. Pavlichenkov. Finite groups of matrices whose entries are integers. *Amer. Math. Monthly*, 109(2):173–186, 2002.
[6] T.J. Laffey. *Lectures on Integer Matrices.* Unpublished lecture notes, 1997.
[7] S. Lang. *Algebra*. Graduate Texts in Mathematics, Vol. 211, Springer-Verlag, New York, 2002.
[8] G. Mackiw. Finite groups of 2×2 integer matrices. *Math. Mag.*, 69(5):356–361, 1996.
[9] M. Newman. *Integral Matrices.* Academic Press, New York, 1972.
[10] W. Roth. The equations $AX - YB = C$ and $AX - XB = C$ in matrices. *Proc. Amer. Math. Soc.*, 3(3):392–396, 1952.
[11] O. Taussky. On a theorem of Latimer and Macduffee. *Canadian J. Math.*, 1:300–302, 1949.
[12] Q. Yang. Conjugacy classes of Torsion in 4×4 integral symplectic group. *J.Math. Res. Exposition*, 28(1):177–191, 2008.