ACTA ARITHMETICA
XX (1972)

Matrix divisors of mI

by

EDWARD SPENCE (Glasgow)

1. Introduction. In [3] canonical forms were found for a certain class of $2n\times 2n$ matrices $\begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix}$ with rational integer entries under the equivalence relation of premultiplication by symplectic unimodular matrices. It was found that each such matrix could be reduced to the form $\begin{bmatrix} A_2 & B_2 \\ 0 & D_2 \end{bmatrix}$ where $A_2D_2^T = mI$ and A_2 is in Hermite's normal form (described below) with positive determinant. B_2 satisfies conditions which need not be repeated here. In order to find the number of such canonical forms it is therefore necessary to investigate matrix divisors of mI. This leads to an interesting result about the index of a certain subgroup of the unimodular group.

2. Preliminaries. Throughout this paper Ω_n will denote the semigroup of all $n\times n$ $(n\geqslant 2)$ matrices with rational integer entries and Γ is the group of all matrices in Ω_n with determinant unity (the unimodular group). As usual, if $A = [a_{ij}]$ and $B = [b_{ij}]\epsilon\Omega_n$, write $A\equiv B \pmod q$ if and only if $a_{ij}\equiv b_{ij} \pmod q$ for $1\leqslant i,j\leqslant n$.

Γ_q is the subgroup of Γ defined by

$$\Gamma_q = \{U\epsilon\Gamma\colon\ U\equiv I \pmod q\}.$$

It is well known ([2]) that Γ_q is a normal subgroup of Γ with finite index

$$[\Gamma\colon\Gamma_q] = q^{n^2-1}\prod_{p|q}\Big\{\prod_{j=2}^{n}(1-p^{-j})\Big\}.$$

The following results have been known for some time and will be used extensively throughout. If $A\epsilon\Omega_n$ has positive determinant, then there is a $U\epsilon\Gamma$ such that $UA = B = [b_{ij}]$ is an upper triangular matrix (Hermite's normal form of A), the entries of which satisfy $b_{ij} = 0$ $(1\leqslant j < i\leqslant n)$, $b_{ii} > 0$ $(1\leqslant i\leqslant n)$ and $0\leqslant b_{ij} < b_{jj}$ $(1\leqslant i < j\leqslant n)$. This form is unique. The reduction can be taken a stage further, for there exist

$V, W \in \Gamma$ such that $VAW = \operatorname{diag}\{d_1, d_2, \ldots, d_n\}$ where $d_{i-1} \mid d_i$ $(1 < i \leqslant n)$ and $d_j > 0$ $(1 \leqslant j \leqslant n)$ (Smith's normal form of A). Here again the normal form is unique, but V and W are not. A more detailed discussion of these results may be found in [4].

3. The matrix equation $AB = mI$. In this section I discuss the solutions of

$$AB = mI \tag{1}$$

where $A, B \in \Omega_n$, A is in Hermite's normal form with positive determinant, I is the identity $n \times n$ matrix, and m is a positive integer. The main result is that $N(m)$, the number of matrices A satisfying (1) and the associated conditions, is multiplicative. This is proved by looking at a certain subset of the solutions of (1). Let $N(m; d)$ $(d > 0)$ be the number of solutions A of (1) with $\det A = d$. It will be proved that if $(d_1, d_2) = 1$, then

$$N(m; d_1 d_2) = N(m; d_1) N(m; d_2) \tag{2}$$

and from this it will be deduced that $N(m)$ is multiplicative. Clearly, since any solution of (1) must have $\det A \mid m^n$, it may be assumed that both d_1 and d_2 divide m^n (for $d_1 \nmid m^n$ implies that $d_1 d_2 \nmid m^n$ and (2) is obviously satisfied). As a first step in the proof of (2) it is shown that $N(m; d_1 d_2) \geqslant N(m; d_1) N(m; d_2)$. To do this the following lemma is required.

LEMMA 1. *Suppose that*

$$A_1 B_1 = A_2 B_2 = mI \tag{3}$$

where $A_1, A_2, B_1, B_2 \in \Omega_n$, A_1, A_2 *are in Hermite's normal form, and* $\det A_1 = d_1$, $\det A_2 = d_2$, $(d_1, d_2) = 1$, *then* $B_1 B_2 \equiv 0 \pmod{m}$.

Proof. The proof is by induction on the order of the matrices. The hypothesis is made that the result is true for all $(n-1) \times (n-1)$ matrices satisfying the conditions of the lemma. Suppose therefore that A_1, A_2, B_1, B_2 are partitioned in the form

$$A_1 = \begin{bmatrix} C_1 & X_1 \\ 0 & \alpha \end{bmatrix}, \quad B_1 = \begin{bmatrix} D_1 & Y_1 \\ 0 & m/\alpha \end{bmatrix}, \quad A_2 = \begin{bmatrix} C_2 & X_2 \\ 0 & \beta \end{bmatrix}, \quad B_2 = \begin{bmatrix} D_2 & Y_2 \\ 0 & m/\beta \end{bmatrix},$$

where $C_1, C_2, D_1, D_2 \in \Omega_{n-1}$ and $\alpha \mid m$, $\beta \mid m$. From the conditions of the lemma it is clear that

$$C_1 D_1 = C_2 D_2 = mI$$

with C_1 and C_2 in Hermite's normal form and $(\det C_1, \det C_2) = 1$. It follows therefore from the induction hypothesis that $D_1 D_2 \equiv 0 \pmod{m}$. From $(d_1, d_2) = 1$ it is deduced that $(\alpha, \beta) = 1$ and consequently $m \mid m^2/\alpha\beta$.

Thus

$$B_1 B_2 = \begin{bmatrix} D_1 D_2 & D_1 Y_2 + (m/\beta) Y_1 \\ 0 & m^2/\alpha\beta \end{bmatrix} \equiv 0 \pmod{m}$$

if and only if

$$D_1 Y_2 + (m/\beta) Y_1 \equiv 0 \pmod{m}. \tag{4}$$

To see that (4) is satisfied observe that

$$\begin{aligned} \beta\big(D_1 Y_2 + (m/\beta) Y_1\big) &= D_1(\beta Y_2) + m Y_1 \\ &= -D_1 D_2 X_2 + m Y_1, \quad \text{from } B_2 A_2 = mI, \\ &\equiv 0 \pmod{m}, \end{aligned}$$

by the induction hypothesis. Thus

$$D_1 Y_2 + (m/\beta) Y_1 \equiv 0 \pmod{m/\beta}. \tag{5}$$

Also

$$\begin{aligned} C_1\big(D_1 Y_2 + (m/\beta) Y_1\big) &= m Y_2 + (m/\beta) C_1 Y_1 \\ &= m Y_2 + (m/\beta)\big(-(m/\alpha) X_1\big), \quad \text{from } A_1 B_1 = mI, \\ &\equiv 0 \pmod{m}, \end{aligned}$$

whence

$$D_1 Y_2 + (m/\beta) Y_1 \equiv 0 \left(\operatorname{mod} \frac{m}{(m, \det C_1)}\right). \tag{6}$$

Combine (5) and (6) to obtain

$$D_1 Y_2 + (m/\beta) Y_1 \equiv 0 \left(\operatorname{mod} \left\{ m/\beta, \frac{m}{(m, \det C_1)} \right\}\right), \tag{7}$$

where the curly brackets denote the least common multiple of m/β and $m/(m, \det C_1)$. Call this t and write $d = (m, \det C_1)$ so that $(d, \beta) = 1$. Since both d and β divide m it follows that $d \mid m/\beta$. Thus $(m/\beta, m/d) = (m/\beta d)(d, \beta) = m/\beta d$ and so

$$t = \frac{(m/\beta)(m/d)}{(m/\beta, m/d)} = m.$$

Consequently $B_1 B_2 \equiv 0 \pmod{m}$ and the result is true for all $n \times n$ matrices satisfying the conditions of the lemma. This establishes the lemma since the result clearly holds for $n = 1$.

LEMMA 2.

$$N(m; d_1 d_2) \geqslant N(m; d_1) N(m; d_2).$$

Proof. From Lemma 1, $A_2 A_1$ and $m^{-1} B_1 B_2 \in \Omega_n$. Further, there exists $U \in \Gamma$ such that $U A_2 A_1$ is in Hermite's normal form. Then if $A = U A_2 A_1$ and $B = m^{-1} B_1 B_2 U^{-1}$, $AB = mI$ where $\det A = d_1 d_2$.

The lemma will be proved if it can be shown that the matrix $A = UA_1A_2$ can arise only from the given A_1, A_2. To this end, suppose that $A_1, A_2, A_1^*, A_2^* \in \Omega_n$ are in Hermite's normal form, that

$$A_1B_1 = A_2B_2 = A_1^*B_1^* = A_2^*B_2^* = mI,$$

$$\det A_1 = \det A_1^* = d_1, \quad \det A_2 = \det A_2^* = d_2, \quad (d_1, d_2) = 1,$$

and that there exists $V \in \Gamma$ such that

(8) $$VA_2A_1 = A_2^*A_1^*.$$

If A_1, A_2, A_1^*, A_2^* are partitioned as previously,

$$A_2A_1 = \begin{bmatrix} C_2 & X_2 \\ 0 & b \end{bmatrix}\begin{bmatrix} C_1 & X_1 \\ 0 & a \end{bmatrix} = \begin{bmatrix} C_2C_1 & C_2X_1 + aX_2 \\ 0 & ab \end{bmatrix},$$

$$A_2^*A_1^* = \begin{bmatrix} C_2^*C_1^* & C_2^*X_1^* + a^*X_2^* \\ 0 & a^*b^* \end{bmatrix}.$$

From (8) it is seen that V must have the form

$$V = \begin{bmatrix} V_1 & V_2 \\ 0 & 1 \end{bmatrix}$$

where $V_1 \in \Omega_{n-1}$ is unimodular, so that $V_1C_2C_1 = C_2^*C_1^*$. Also, it is clear that $ab = a^*b^*$, and since $(a, b^*) = (a^*, b) = 1$, $a = a^*$ and $b = b^*$ from which it follows that $\det C_1 = \det C_1^*$ and $\det C_2 = \det C_2^*$.

It is now possible to adopt an induction hypothesis that if $P_1, P_2, P_1^*, P_2^* \in \Omega_{n-1}$ are in Hermite's normal form, and if

$$P_1Q_1 = P_2Q_2 = P_1^*Q_1^* = P_2^*Q_2^* = mI,$$

$$\det P_1 = \det P_1^*, \quad \det P_2 = \det P_2^*, \quad (\det P_1, \det P_2) = 1$$

$(Q_1, Q_2, Q_1^*, Q_2^* \in \Omega_{n-1})$, and if $UP_2P_1 = P_2^*P_1^*$ for some unimodular U, then $U = I$ and $P_1^* = P_1, P_2^* = P_2$. Using this hypothesis it is clear that $V_1 = I$ and $C_1^* = C_1$, $C_2^* = C_2$ and so from (8)

$$C_2X_1 + aX_2 + V_2ab = C_2X_1^* + aX_2^*.$$

This implies that $C_2(X_1 - X_1^*) \equiv 0 \pmod{a}$ which in turn yields

$$X_1 - X_1^* \equiv 0 \pmod{a/(a, \det C_2)}.$$

But $(a, \det C_2) = 1$ and $X_1 - X_1^* \equiv 0 \pmod{a}$ implies that $X_1 = X_1^*$ since A_1, A_1^* are in Hermite's normal form. Since the induction hypothesis is clearly valid for matrices of order 1 it is immediate that

$$N(m; d_1)N(m; d_2) \leqslant N(m; d_1d_2).$$

Before the opposite inequality to this can be proved a definition is required. Let $P, Q \in \Omega_n$. Call the pair (P, Q) *coprime* if the matrix products GP and $GQ \in \Omega_n$ if and only if $G \in \Omega_n$. It is easily shown ([5]) that this definition of a coprime pair is equivalent to the existence of $X, Y \in \Omega_n$ such that $PX + QY = I$. This result will be used in the next lemma.

LEMMA 3. *Given $A \in \Omega_n$ in Hermite's normal form with $\det A = d_1d_2$, $(d_1, d_2) = 1$, and which satisfies*

$$AB = mI, \quad (B \in \Omega_n),$$

there is a procedure which gives uniquely matrices A_1, A_2 in Hermite's normal form such that

$$A_1B_1 = A_2B_2 = mI,$$

$\det A_1 = d_1$, $\det A_2 = d_2$; *and $A = UA_1A_2$ for some $U \in \Gamma$.*

Proof. If $A = [a_{ij}]$, then $a_{11}a_{22}\ldots a_{nn} = d_1d_2$. Let $\alpha_i = (a_{ii}, d_1)$, $\beta_i = a_{ii}/(a_{ii}, d_1)$ $(1 \leqslant i \leqslant n)$ so that $\alpha_1\alpha_2\ldots\alpha_n\beta_1\beta_2\ldots\beta_n = d_1d_2$. However, $(d_1, d_2) = 1$ and this gives $\alpha_1\alpha_2\ldots\alpha_n = d_1$, $\beta_1\beta_2\ldots\beta_n = d_2$.

Consider the partitions of A and B

$$A = \begin{bmatrix} C & X \\ 0 & a_{nn} \end{bmatrix}, \quad B = \begin{bmatrix} D & Y \\ 0 & b_{nn} \end{bmatrix} \quad (C, D \in \Omega_{n-1}),$$

so that $d_1d_2 = a_{nn}\det C$. Then $\det C = (d_1/\alpha_n)(d_2/\beta_n)$ and $(d_1/\alpha_n, d_2/\beta_n) = 1$. Clearly $CD = mI$ and C is in Hermite's normal form. It is therefore possible to apply the induction hypothesis that there exist $C_1, C_2, D_1, D_2 \in \Omega_{n-1}$ such that $C = C_1C_2$ with C_2 in Hermite's normal form, $C_1D_1 = C_2D_2 = mI$ and $\det C_1 = d_1/\alpha_n$, $\det C_2 = d_2/\beta_n$. Then

$$A = \begin{bmatrix} C_1C_2 & X \\ 0 & \alpha_n\beta_n \end{bmatrix}$$

which can be written as

$$\begin{bmatrix} C_1 & Y_1 \\ 0 & \alpha_n \end{bmatrix}\begin{bmatrix} C_2 & Z_1 \\ 0 & \beta_n \end{bmatrix}$$

if and only if

(9) $$X = C_1Z_1 + \beta_nY_1.$$

Now the matrices (C_1, β_nI_{n-1}) form a coprime pair, for $(\det C_1, \beta_n) = 1$ and so there exist integers u and v such that $u\det C_1 + v\beta_n = 1$, and consequently (GC_1, β_nG) both integral implies that $(G\det C_1, \beta_nG)$ are both integral which in turn yields $u\det C_1G + v\beta_nG = G$ is integral. As a result there exist $P, Q \in \Omega_{n-1}$ such that $C_1P + \beta_nQ = I_{n-1}$. Thus $C_1(PX) + \beta_n(QX) = X$ and if Z_1 and Y_1 are chosen so that $Z_1 = PX + \beta_nZ$, $Y_1 = QX - C_1Z$ for any $(n-1)\times 1$ vector Z, then (9) is satisfied, and hence

$$A = \begin{bmatrix} C_1 & Y_1 \\ 0 & \alpha_n \end{bmatrix}\begin{bmatrix} C_2 & Z_1 \\ 0 & \beta_n \end{bmatrix} = A_1^*A_2^*,$$

say, where $\det A_1^* = d_1$, $\det A_2^* = d_2$. Choose the vector Z so that every entry of $Z_1 = \{z_1, z_2, \ldots, z_{n-1}\}$ satisfies $0 \leqslant z_i < \beta_n$ $(1 \leqslant i \leqslant n-1)$ and A_2^* will then be in Hermite's normal form (denote this by A_2). Also, there exists $U \in \Gamma$ such that $A_1 = U^{-1}A_1^*$ is in Hermite's normal form, so that $A = UA_1A_2$ and $A_1(A_2BU^{-1}) = mI$, i.e. $A_1B_1 = mI$ where $B_1 = A_2BU^{-1}$. Moreover, $A_2B_2 = mI$ where $B_2 = BUA_1$. Since the induction hypothesis is trivially true for matrices of order 1, the procedure defined above is clearly valid for all matrices of order $\geqslant 1$ satisfying the conditions. Moreover, the A_1, A_2 obtained from A are unique, by an argument similar to the one given in Lemma 2.

Thus $N(m; d_1d_2) = N(m; d_1)N(m; d_2)$ if $(d_1, d_2) = 1$.

LEMMA 4. *Let $m = m_1m_2$ where $(m_1, m_2) = 1$. Then if $d \mid m_1^n$,*

$$N(m; d) = N(m_1; d).$$

Proof. For

$$\begin{aligned} AB = mI \quad \text{with } \det A = d &\Leftrightarrow dB = m_1m_2 \operatorname{adj} A \\ &\Leftrightarrow B_1 = (m_1/d)\operatorname{adj} A \in \Omega_n \quad (\text{since } (d, m_2) = 1) \\ &\Leftrightarrow AB_1 = m_1 I, \end{aligned}$$

which proves the lemma.

THEOREM 1. *$N(m)$ is multiplicative.*

Proof. Suppose $(m_1, m_2) = 1$. Then

$$\begin{aligned} N(m_1m_2) &= \sum_{d \mid (m_1m_2)^n} N(m_1m_2; d) \\ &= \sum_{d_1 \mid m_1^n} \sum_{d_2 \mid m_2^n} N(m_1m_2; d_1d_2) \\ &= \sum_{d_1 \mid m_1^n} N(m_1; d_1) \sum_{d_2 \mid m_2^n} N(m_2; d_2), \quad \text{by Lemma 4,} \\ &= N(m_1)N(m_2). \end{aligned}$$

4. Evaluation of $N(m)$. Before utilizing fully the fact that $N(m)$ is multiplicative, an alternative formulation is given for it in terms of the indices of certain subgroups of Γ in Γ.

Suppose that, as in (1), $AB = mI$. Then a reduction of A to Smith's normal form shows that there exist positive integers $d_1, d_2, \ldots, d_n$ and $U, V \in \Gamma$ such that $UAV = \operatorname{diag}\{d_1, d_2, \ldots, d_n\}$ with $d_{i-1} \mid d_i$ $(1 < i \leqslant n)$. It follows that $(UAV)(V^{-1}BU^{-1}) = mI$, and consequently, since $V^{-1}BU^{-1} \in \Omega_n$, $d_i \mid m$ $(1 \leqslant i \leqslant n)$.

The converse to this is obviously true. If

$$D = \operatorname{diag}\{d_1, d_2, \ldots, d_n\} \tag{10}$$

with $d_{i-1} \mid d_i$ $(1 < i \leqslant n)$ and $d_n \mid m$, then $A = U_1DV_1$ and $B = V_1^{-1}mD^{-1}U_1^{-1}$ satisfy (1), for any $U_1, V_1 \in \Gamma$. Since all solutions of (1) are obtained from a matrix of the form (10) it is natural to ask the following question. Given D defined by (10), how many of U_1DV_1, $U_1, V_1 \in \Gamma$, have different Hermite's normal form? Now U_1DV_1 and U_2DV_2 have the same Hermite's normal form if and only if

$$\begin{aligned} D^{-1}(UU_2)^{-1}U_1D = V_2V_1^{-1} \quad &\text{for some } U \in \Gamma, \\ \Leftrightarrow V_2V_1^{-1} \in D^{-1}\Gamma D \cap \Gamma. & \end{aligned}$$

Since $D^{-1}\Gamma D \cap \Gamma$ is a subgroup of Γ, this last condition is equivalent to requiring that V_1 and V_2 belong to the same right coset of $D^{-1}\Gamma D \cap \Gamma$ in Γ. It follows that the number of matrices of the form U_1DV_1 with $U_1, V_1 \in \Gamma$ which have different Hermite's normal form is $[\Gamma: D^{-1}\Gamma D \cap \Gamma]$. Denote this number by $N_n(d_1, d_2, \ldots, d_n)$. Then, by the above discussion

$$N(m) = \sum_{d_{i-1} \mid d_i \mid m} N_n(d_1\, d_2, \ldots, d_n).$$

For each prime p dividing m let $m(p) = \max\{k\colon p^k \mid m\}$, so that, since $N(m)$ is multiplicative,

$$N(m) = \prod_{p \mid m} \Big\{ \sum_{0 \leqslant \alpha_1 \leqslant \ldots \leqslant \alpha_n \leqslant m(p)} N_n(p^{\alpha_1}, p^{\alpha_2}, \ldots, p^{\alpha_n}) \Big\}.$$

Thus to evaluate $N(m)$ it is sufficient to consider

$$N_n(p^{\alpha_1}, p^{\alpha_2}, \ldots, p^{\alpha_n})$$

where p is a prime divisor of m and $0 \leqslant \alpha_1 \leqslant \alpha_2 \leqslant \ldots \leqslant \alpha_n \leqslant m(p)$. Let D be the $n \times n$ matrix $\operatorname{diag}\{p^{\alpha_1}, p^{\alpha_2}, \ldots, p^{\alpha_n}\}$ and let $H = H(\alpha_1, \alpha_2, \ldots, \alpha_n)$ denote the subgroup $D^{-1}\Gamma D \cap \Gamma$. Then

$$N_n(p^{\alpha_1}, p^{\alpha_2}, \ldots, p^{\alpha_n}) = [\Gamma : H].$$

LEMMA 5. *Suppose $\beta \geqslant \alpha_n$ and write $q = p^\beta$. Then*

$$[\Gamma : H] = [\Gamma : \Gamma_q]/[H : \Gamma_q].$$

Proof. It is trivial to verify that $\Gamma_q \subseteq H$ and the result follows.

Since, as was mentioned earlier, $[\Gamma : \Gamma_q]$ is known, evaluation of $[\Gamma : H]$ has been reduced to the evaluation of $[H : \Gamma_q]$. This can be done in a way similar to the derivation of $[\Gamma : \Gamma_q]$ in [2].

Let $U = [u_{ij}] \in H$; then $U = D^{-1}VD$ for some $V = [v_{ij}] \in \Gamma$, and a little manipulation shows that $u_{ij} = v_{ij}p^{\alpha_j - \alpha_i}$ $(1 \leqslant i, j \leqslant n)$. Thus $U = [u_{ij}] \in H$ if and only if $\det U = 1$ and $u_{ij} \equiv 0 \pmod{p^{\alpha_j - \alpha_i}}$ $(1 \leqslant i < j \leqslant n)$ (since $0 \leqslant \alpha_1 \leqslant \alpha_2 \leqslant \ldots \leqslant \alpha_n$).

LEMMA 6. *$[H:\Gamma_q]$ is the number of matrices $U = [u_{ij}] \epsilon \Omega_n$ which are incongruent modulo q, where $u_{ij} \equiv 0 \pmod{p^{\alpha_j-\alpha_i}}$ $(1 \leqslant i < j \leqslant n)$ and* $\det U \equiv 1 \pmod{q}$.

Proof. Let $U, V \epsilon H$. Then $U \equiv V \pmod{q}$ if and only if $UV^{-1} \equiv I \pmod{q}$, i.e. if and only if U and V belong to the same right coset of Γ_q in H. The lemma will therefore be proved if it can be shown that, given $U = [u_{ij}] \epsilon \Omega_n$ with $u_{ij} \equiv 0 \pmod{p^{\alpha_j-\alpha_i}}$ $(1 \leqslant i < j \leqslant n)$ and $\det U \equiv 1 \pmod{q}$, there exists $W \epsilon H$ such that $W \equiv U \pmod{q}$. This is done in

LEMMA 7. *If $A = [a_{ij}] \epsilon \Omega_n$ is such that $a_{ij} \equiv 0 \pmod{p^{\alpha_j-\alpha_i}}$ $(1 \leqslant i < j \leqslant n)$ and $\det A \equiv a \pmod{q}$ where $(a, q) = 1$, then there exists $B = [b_{ij}] \epsilon \Omega_n$ such that $\det B = a$, $b_{ij} \equiv 0 \pmod{p^{\alpha_j-\alpha_i}}$ $(1 \leqslant i < j \leqslant n)$ and $B \equiv A \pmod{q}$.*

Proof. First observe that if there exists $B \equiv A \pmod{q}$ then necessarily $b_{ij} \equiv 0 \pmod{p^{\alpha_j-\alpha_i}}$ $(1 \leqslant i < j \leqslant n)$. For

$$b_{ij} \equiv a_{ij} \pmod{q} \Rightarrow b_{ij} \equiv a_{ij} \pmod{p^{\alpha_j-\alpha_i}}$$
$$\equiv 0 \bmod (p^{\alpha_j-\alpha_i} \quad (1 \leqslant i < j \leqslant n)$$

since $\beta \geqslant \alpha_n \geqslant \alpha_j - \alpha_i$ $(1 \leqslant i < j \leqslant n)$.

It is therefore sufficient to consider the lemma without the hypothesis that $a_{ij} \equiv 0 \pmod{p^{\alpha_j-\alpha_i}}$ $(1 \leqslant i < j \leqslant n)$. Make the induction hypothesis that the result is true for all $(n-1)\times(n-1)$ matrices with integer entries and all a such that $(a, q) = 1$.

Let $\mathrm{diag}\{a_1, a_2, \ldots, a_n\}$ be Smith's normal form of A $(n > 1)$ so that there exist $U_1, V_1 \epsilon \Gamma$ such that

$$(11) \qquad U_1AV_1 = \mathrm{diag}\{a_1, a_2, \ldots, a_n\}, \qquad \text{and} \qquad a_1a_2\ldots a_n \equiv a \pmod{q}$$

where $a_{i-1} | a_i$ $(1 < i \leqslant n)$. Postmultiply U_1AV_1 by the unimodular matrix

$$V_2 = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ 1 & 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot\;\cdot\;\cdot & \cdot \\ 1 & 0 & 0 & \ldots & 1 \end{bmatrix}$$

to obtain

$$U_1AV_1V_2 = \begin{bmatrix} a_1 & 0 & 0 & \ldots & 0 \\ a_2 & a_2 & 0 & \ldots & 0 \\ a_3 & 0 & a_3 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot\;\cdot\;\cdot & \cdot \\ a_n & 0 & 0 & \ldots & a_n \end{bmatrix} \equiv \begin{bmatrix} a_1 & 0 & 0 & \ldots & 0 \\ a_2+q & a_2 & 0 & \ldots & 0 \\ a_3 & 0 & a_3 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot\;\cdot\;\cdot & \cdot \\ a_n & 0 & 0 & \ldots & a_n \end{bmatrix} \pmod{q}.$$

Call this latter matrix D. Then since $a_1 | a_2$ and $(a_1, q) = 1$ from (11), it is seen that $\mathrm{g.c.d.}(a_1, a_2+q, a_3, \ldots, a_n) = 1$. Premultiplying D by a suitable unimodular matrix makes the (1, 1) entry 1 and then by subtracting suitable multiples of the first column from every other column, and then suitable multiples of the first row from every other row, the existence is established of $U, V \epsilon \Gamma$ such that

$$UAV \equiv \begin{bmatrix} 1 & 0 \\ 0 & A_1 \end{bmatrix} \pmod{q},$$

where A_1 is an $(n-1)\times(n-1)$ matrix with integer entries and $\det A_1 \equiv a \pmod{q}$. Therefore, by the induction hypothesis there exists $(n-1)\times\times(n-1)$ matrix B_1 such that $A_1 \equiv B_1 \pmod{q}$ and $\det B_1 = a$. It follows that

$$A \equiv U^{-1}\begin{bmatrix} 1 & 0 \\ 0 & A_1 \end{bmatrix} V^{-1} \pmod{q},$$

and if B is taken to be $U^{-1}\begin{bmatrix} 1 & 0 \\ 0 & B_1 \end{bmatrix} V^{-1}$ the lemma is proved, since the result is clearly valid for $n = 1$.

Before proving the main theorem the following two lemmas are required.

LEMMA 8. *If $J_r(1, p^t)$ is the number of r-tuples $(a_1, a_2, \ldots, a_r)$ of integers $a_1, a_2, \ldots, a_r$ in a complete system of residues* (mod p^t) $(t \geqslant 1)$ *such that $(a_1, a_2, \ldots, a_r, p^t) = 1$, then*

$$(12) \qquad J_r(1, p^t) = p^{rt}(1-p^{-r}).$$

Proof. $J_r(1, p^t)$ is Jordan's generalization of Euler's φ-function. The proof of (12) is well known and can be found in [1].

LEMMA 9. *If $a_1, a_2, \ldots, a_r$ are r integers such that $(a_1, a_2, \ldots, a_r, m) = 1$ $(m > 1)$, then there exist $\lambda_1, \lambda_2, \ldots, \lambda_r$ with $(\lambda_1, m) = 1$ such that*

$$\lambda_1a_1 + \lambda_2a_2 + \ldots + \lambda_ra_r \equiv 1 \pmod{m}.$$

Proof. This is proved in [2].

Let $U_n(\beta; \alpha_1, \alpha_2, \ldots, \alpha_n)$ denote the set of matrices $U = [u_{ij}] \epsilon \Omega_n$ satisfying $u_{ij} \equiv 0 \pmod{p^{\alpha_j-\alpha_i}}$ $(1 \leqslant i < j \leqslant n)$, $0 \leqslant u_{ij} < q$ $(1 \leqslant i, j \leqslant n)$ and $\det U \equiv 1 \pmod{q}$, so that $[H:\Gamma_q] = |U_n(\beta; \alpha_1, \alpha_2, \ldots, \alpha_n)|$. Also, at this stage make the assumption that

$$(p^{\alpha_1}, p^{\alpha_2}, \ldots, p^{\alpha_n}) \equiv (\underbrace{p^{a_1}, \ldots, p^{a_1}}_{r_1 \text{ factors}}, \underbrace{p^{a_2}, \ldots, p^{a_2}}_{r_2 \text{ factors}}, \ldots, \underbrace{p^{a_k}, \ldots, p^{a_k}}_{r_k \text{ factors}}),$$

where $\alpha_1 = a_1 < a_2 < \ldots < a_k = \alpha_n$, $r_1 + r_2 + \ldots + r_k = n$, $r_i \geqslant 1$ $(1 \leqslant i \leqslant k)$ and $n \geqslant 2$. Suppose, further, that $k \geqslant 2$ since the case $k \doteq 1$ is trivial.

THEOREM 2. *If* $n \geqslant 2$,

$$|U_n(\beta;\alpha_1,\alpha_2,\ldots,\alpha_n)| = q^{n^2-1}p^{\sum\limits_1^n(n+1-2j)\alpha_j}\prod_{i=1}^{k}\Big\{\prod_{j=1}^{r_i}(1-p^{-j})\Big\}(1-p^{-1})^{-1}.$$

Proof. Let $U = [u_{ij}] \epsilon U_n(\beta;\alpha_1,\alpha_2,\ldots,\alpha_n)$. Then $(u_{11},\ldots,u_{1n},q) = 1$ and by Lemma 9 there exist $\lambda_1,\lambda_2,\ldots,\lambda_n$ with $(\lambda_1,q) = 1$ such that $\lambda_1u_{11}+\lambda_2u_{12}+\ldots+\lambda_nu_{1n} \equiv 1 \pmod{q}$. Postmultiply U by the matrix

$$V = \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ \lambda_2 & 1 & 0 & \ldots & 0 \\ \lambda_3 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \lambda_n & 0 & 0 & \ldots & \lambda_1^{-1} \end{bmatrix}$$

where λ_1^{-1} is the integer x such that $0 < x < q$ and $x\lambda_1 \equiv 1 \pmod{q}$, to get a matrix W with entries reduced $(\mathrm{mod}\ q)$, $W = [w_{ij}] \equiv UV \pmod{q}$, so that

$$w_{11} = 1,\quad w_{i1} \equiv u_{i1}\lambda_1+u_{i2}\lambda_2+\ldots+u_{in}\lambda_n \pmod{q}\quad (i>1),$$
$$w_{ij} = u_{ij}\ (1<j<n)\quad\text{and}\quad w_{in} \equiv \lambda_1^{-1}u_{in} \pmod{q}\ (1\leqslant i\leqslant n).$$

Postmultiply W by

$$W_1 = \begin{bmatrix} 1 & -w_{12} & -w_{13} & \ldots & w_{1n} \\ 0 & 1 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}$$

to get a matrix A with entries reduced $(\mathrm{mod}\ q)$

$$A = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & a_{n3} & \ldots & a_{nn} \end{bmatrix} \equiv UVW_1 \pmod{q},$$

where $a_{i1} = w_{i1}$ $(1 < i \leqslant n)$ and $a_{ij} \equiv w_{ij}-w_{i1}w_{1j} \pmod{q}$ $(1 < i, j \leqslant n)$. Since $w_{1j} \equiv 0 \ (\mathrm{mod}\ p^{\alpha_j-\alpha_1})$, it follows that

$$w_{1j} \equiv 0\ (\mathrm{mod}\ p^{\alpha_j-\alpha_i})\quad (1\leqslant i<j),$$

and consequently

$$(13)\qquad a_{ij} \equiv 0\ (\mathrm{mod}\ p^{\alpha_j-\alpha_i})\quad (1\leqslant i<j\leqslant n).$$

Thus $A \epsilon U_n(\beta;\alpha_1,\alpha_2,\ldots,\alpha_n)$. Let $B = [b_{ij}]$ be the $(n-1)\times(n-1)$ matrix defined by $b_{ij} = a_{i+1,j+1}$ $(1\leqslant i, j<n)$. So that

$$\det A \equiv 1 \pmod{q} \Leftrightarrow \det B \equiv 1 \pmod{q}.$$

This, in conjunction with (13), implies that $B \epsilon U_{n-1}(\beta;\alpha_2,\alpha_3,\ldots,\alpha_n)$. Thus the number of matrices in $U_n(\beta;\alpha_1,\alpha_2,\ldots,\alpha_n)$ of the same form as A is

$$q^{n-1}|U_{n-1}(\beta;\alpha_2,\alpha_3,\ldots,\alpha_n)|$$

(since there are q^{n-1} choices for the first column of A).

It will now be shown that for a given matrix A of the above form and a given first row of U, the other rows of U are uniquely determined $(\mathrm{mod}\ q)$. For if $i \geqslant 2$

$$a_{i1} \equiv w_{i1} \pmod{q},$$
$$a_{i2} \equiv w_{i2}-w_{i1}w_{12} \pmod{q},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_{in} \equiv w_{in}-w_{i1}w_{1n} \pmod{q},$$

i.e.

$$a_{i1} \equiv u_{i1}\lambda_1+u_{i2}\lambda_2+\ldots+u_{in}\lambda_n \pmod{q},$$
$$a_{i2} \equiv u_{i2}-a_{i1}u_{12} \pmod{q},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_{i,n-1} \equiv u_{i,n-1}-a_{i1}u_{1,n-1} \pmod{q},$$
$$a_{in} \equiv \lambda_1^{-1}u_{in}-\lambda_1^{-1}a_{i1}u_{1n} \pmod{q},$$

and the determinant of this transformation is

$$\begin{vmatrix} \lambda_1 & \lambda_2 & \ldots & \lambda_{n-1} & \lambda_n \\ -u_{12}\lambda_1 & 1-u_{12}\lambda_2 & \ldots & -u_{12}\lambda_{n-1} & -u_{12}\lambda_n \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ -u_{1,n-1}\lambda_1 & -u_{1,n-1}\lambda_2 & \ldots & 1-u_{1,n-1}\lambda_{n-1} & -u_{1,n-1}\lambda_n \\ -u_{1n} & -u_{1n}\lambda_1^{-1}\lambda_2 & \ldots & -u_{1n}\lambda_1^{-1}\lambda_{n-1} & \lambda_1^{-1}(1-u_{1n}\lambda_n) \end{vmatrix}$$

which is easily transformed into the following determinant by adding suitable multiples of the first row to every other row,

$$\begin{vmatrix} \lambda_1 & \lambda_2 & \ldots & \lambda_{n-1} & \lambda_n \\ 0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & \ldots & 0 & \lambda_1^{-1} \end{vmatrix}.$$

Since this determinant is congruent to $1 \pmod{q}$, it follows that, if for all $i \geqslant 2$ $(a_{i1},a_{i2},\ldots,a_{in})$ and $(u_{11},u_{12},\ldots,u_{1n})$ are given, then $(u_{i1},u_{i2},\ldots,u_{in})$ $(i\geqslant 2)$ is uniquely determined.

The next step is to determine the number of possible choices of $(u_{11}, u_{12}, \ldots, u_{1n})$ as the first row of $U \epsilon U_n(\beta; a_1, a_2, \ldots, a_n)$. Clearly it is necessary that

$$(u_{11}, u_{12}, \ldots, u_{1r_1}, u_{1,r_1+1}, \ldots, u_{1n}, q) = 1. \tag{14}$$

But $u_{1j} \equiv 0 \pmod p$ for $j > r_1$ and so condition (14) is equivalent to

$$(u_{11}, u_{12}, \ldots, u_{1r_1}, q) = 1, \quad 0 \leqslant u_{1j} < q \ (1 \leqslant j \leqslant n). \tag{15}$$

Using Lemma 8, the number of $(u_{11}, u_{12}, \ldots, u_{1n})$ satisfying (15) (bearing in mind the fact that $u_{1j} \equiv 0 \pmod{p^{a_j - a_1}}$) is

$$p^{\beta r_1}(1-p^{-r_1})p^{\sum\limits_{j>r_1}(\beta - a_j + a_1)}.$$

This may be rewritten as

$$p^{\sum\limits_1^n (a_1 - a_j)}(1-p^{-r_1})q^n.$$

Consequently

$$|U_n(\beta; a_1, a_2, \ldots, a_n)| = q^{2n-1}(1-p^{-r_1})p^{\sum\limits_1^n (a_1-a_j)}|U_{n-1}(\beta; a_2, \ldots, a_n)|. \tag{16}$$

It is now a simple exercise to use (16) to prove the theorem by induction:

Corollary 1.

$$N_n(p^{a_1}, \ldots, p^{a_n}) = [\Gamma : H] = p^{\sum\limits_1^n (2j-n-1)a_j} \frac{\prod\limits_{j=1}^{n}(1-p^{-j})}{\prod\limits_{i=1}^{k}\{\prod\limits_{j=1}^{r_i}(1-p^{-j})\}}.$$

Proof. The result follows immediately from Theorem 1 and Lemma 5.

Using this corollary $N(m)$ may be calculated in a finite number of steps from the formula

$$N(m) = \prod_{p|m}\Big\{\sum_{0 \leqslant a_1 \leqslant \ldots \leqslant a_n \leqslant m(p)} N_n(p^{a_1}, p^{a_2}, \ldots, p^{a_n})\Big\}.$$

The contents of the above paper formed part of a Ph. D. thesis presented to the University of Glasgow in 1968; the author gratefully acknowledges the debt he owes to his supervisor Professor R. A. Rankin for his advice and encouragement.

References

[1] L. E. Dickson, *History of the Theory of Numbers*, Vol. 1, Washington 1918.

[2] N. J. Fine and I. Niven, *The probability that a determinant be congruent to a* (mod m), Bull. Amer. Math. Soc. 50 (1944), pp. 89–93.

[3] H. Maass, *Die Primzahlen in der Theorie der Siegelschen Modulfunktionen*, Math. Ann. 124 (1951), pp. 87–122.

[4] C. C. MacDuffee, *The Theory of Matrices*, New York 1946.

[5] C. L. Siegel, *Einführung in die Theorie der Modulfunktionen n-ten Grades*, Math. Ann. 116 (1939), pp. 615–657.

UNIVERSITY OF GLASGOW
Glasgow, W.2.

Received on 28. 1. 1971 (139)