ON THE NUMBER OF MATRICES WITH GIVEN CHARACTERISTIC POLYNOMIAL

BY
IRVING REINER[1]

1. Introduction

Let K be a finite field with q elements, and let K_n denote the ring of all $n \times n$ matrices with entries in K. Recently Fine and Herstein proved[2]

The number of nilpotent matrices in K_n is q^{n^2-n}.

We shall prove here the following generalizations.

THEOREM 1. *Let $f(x)$ be an irreducible polynomial in $K[x]$ of degree $d \geqq 1$. Then the number of matrices $X \in K_{nd}$ for which $f(X)$ is nilpotent is*

$$q^{n^2d^2-nd} \cdot \frac{(1-q^{-1})(1-q^{-2}) \cdots (1-q^{-nd})}{(1-q^{-d})(1-q^{-2d}) \cdots (1-q^{-nd})}. \tag{1}$$

Before stating the second result to be proved here, which includes the above theorem as a special case, we introduce some notation. Define

$$F(u, r) = (1-u^{-1})(1-u^{-2}) \cdots (1-u^{-r}), \tag{2}$$

where $F(u, 0) = 1$. Then we have[3]

THEOREM 2. *Let $g(x) \in K[x]$ be a polynomial of degree n, and let*

$$g(x) = f_1^{n_1}(x) \cdots f_k^{n_k}(x) \tag{3}$$

be its factorization in $K[x]$ into powers of distinct irreducible polynomials $f_1(x), \cdots, f_k(x)$. Set

$$d_i = \text{degree of } f_i(x), \qquad 1 \leqq i \leqq k.$$

Then the number of matrices $X \in K_n$ with characteristic polynomial $g(x)$ is

$$q^{n^2-n} \cdot \frac{F(q, n)}{\prod_{i=1}^{k} F(q^{d_i}, n_i)}. \tag{4}$$

The proofs of these theorems do not require a knowledge of the Fine-Herstein paper, except for the following combinatorial lemma which they establish and which we state without proof.

Received July 2, 1960.

[1] This research was supported by a contract with the Office of Naval Research.

[2] N. J. FINE AND I. N. HERSTEIN, *The probability that a matrix be nilpotent*, Illinois J. Math., vol. 2 (1958), pp. 499–504.

[3] Another proof of Theorems 1 and 2 is given by M. GERSTENHABER, *On the number of nilpotent matrices with coefficients in a finite field*, Illinois J. Math., vol. 5 (1961), pp. 330–333.

LEMMA 1 (Fine-Herstein). *Let u be any complex number which is not a root of unity. Let $\{r_1, \cdots, r_n\}$ range over all n-tuples of non-negative integers for which*

$$r_1 + 2r_2 + \cdots + nr_n = n,$$

and set

$$s_j = r_j + r_{j+1} + \cdots + r_n, \qquad 1 \leqq j \leqq n.$$

Then

$$\sum_{\{r_1,\cdots,r_n\}} \frac{u^{s_1^2+s_2^2+\cdots+s_n^2}}{F(u^{-1}, r_1)F(u^{-1}, r_2) \cdots F(u^{-1}, r_n)} = \frac{u^n}{F(u^{-1}, n)}. \tag{5}$$

2. Automorphisms of modules over local rings

Throughout this section we let R be a commutative local ring with unity element, and let πR be its unique maximal ideal. Suppose further that π is nilpotent, say $\pi^n = 0$, and let

$$t = \text{number of elements in the field } R/\pi R.$$

Then

$$R \supset \pi R \supset \pi^2 R \supset \cdots \supset \pi^{n-1}R \supset \pi^n R = (0)$$

is a descending chain of ideals of R in which every ideal occurs, and each quotient is isomorphic (as R-module) to the field $R/\pi R$. Thus R contains t^n elements, and more generally $R/\pi^j R$ contains t^j elements.

We shall restrict our attention to R-modules which are finitely generated. Since R is a principal ideal domain, each such R-module V is a direct sum of cyclic R-modules. Moreover every nonzero cyclic R-module is a homomorphic image of R, hence is of the form $R/\pi^j R$ for some j, $1 \leqq j \leqq n$. Set

$$V_j = R/\pi^j R, \qquad 1 \leqq j \leqq n.$$

Then V_j contains t^j elements, and is indecomposable since it contains a unique minimal submodule $\pi^{j-1}V_j$. Thus every R-module V is expressible as

$$\begin{aligned} V &= W_1 \oplus \cdots \oplus W_n, \\ W_j &= V_j \oplus \cdots \oplus V_j \qquad (r_j \text{ summands}), \end{aligned} \tag{6}$$

and such an expression is unique by the Krull-Schmidt Theorem.

LEMMA 2. *Let W_j be given in* (6). *The number of R-automorphisms of W_j is precisely*

$$t^{jr_j^2}F(t, r_j).$$

Proof. Since π^j annihilates W_j, we may regard W_j as an R'-module, where

$$R' = R/\pi^j R.$$

The number of R-automorphisms of W_j is then the same as the number of nonsingular $r_j \times r_j$ matrices X with entries in R'. Now a matrix X over R' is nonsingular if and only if $\bar{X}$ is nonsingular, where $\bar{X}$ is obtained from X by mapping each entry α of X onto its image $\bar{\alpha}$ in $R'/\pi R'$. Since $\bar{X}$ has its entries in the field $R'/\pi R' \cong R/\pi R$, there are

$$t^{r_j^2} F(t, r_j)$$

possible choices for $\bar{X}$. But for given $\bar{\alpha}$ there are t^{j-1} choices for $\alpha \in R'$, and thus the number of nonsingular matrices X over R' of size $r_j \times r_j$ is

$$(t^{j-1})^{r_j^2} \cdot t^{r_j^2} F(t, r_j).$$

This proves the lemma.

LEMMA 3. *Let V be given by* (6), *and set*

$$s_j = r_j + r_{j+1} + \cdots + r_n, \qquad 1 \leqq j \leqq n. \tag{7}$$

The number of R-automorphisms of V is then

$$N_V = \prod_{j=1}^{n} t^{s_j^2} F(t, r_j). \tag{8}$$

Proof. For convenience we rewrite (6) as

$$V = \sum_{j=1}^{n} \sum_{i=1}^{r_j} V_j e_{ji},$$

where e_{ji} is just an indexing mark, say

$$e_{ji} = (0, \cdots, 0, 1, 0, \cdots, 0)$$

with the 1 in an appropriate position. Any R-homomorphism is completely determined by its effect on the $\{e_{ji}\}$.

For $v \in V$, $v \neq 0$, define the *order* of v to be the smallest integer s for which $\pi^s v = 0$. Let us say that 0 has order zero. The elements of W_j have order $\leqq j$, clearly.

Now let θ be an R-automorphism of V. Then θ preserves order, so that for $1 \leqq j \leqq n$ we have

$$\theta(W_j) \subset W_1 + \cdots + W_{j-1} + W_j + \pi W_{j+1} + \cdots + \pi^{n-j} W_n.$$

Hence if we set

$$\theta(e_{ji}) = \sum_m a_{ji}^{(m)}, \qquad a_{ji}^{(m)} \in W_m, \tag{9}$$

then we see that for $m > j$ we have

$$a_{ji}^{(m)} \in \pi^{m-j} W_m. \tag{10}$$

Furthermore, for fixed j the mapping

$$e_{ji} \to a_{ji}^{(j)}, \qquad 1 \leqq i \leqq r_j, \tag{11}$$

must be an R-automorphism of W_j. It is easy to see that conversely if we define an R-homomorphism θ by means of (9) and (10), where for each j ($1 \leqq j \leqq n$) equation (11) gives an R-automorphism of W_j, then θ is indeed an R-automorphism of V.

For fixed j, $1 \leqq j \leqq n$, the elements $\{a_{ji}^{(m)}: m < j\}$ may be chosen arbitrarily. Since there are r_j choices to be made, and $W_1 + \cdots + W_{j-1}$ contains

$$t^{1r_1+2r_2+\cdots+(j-1)r_{j-1}}$$

elements, this gives

$$t^{r_j(1r_1+2r_2+\cdots+(j-1)r_{j-1})} \tag{12}$$

possibilities for the $\{a_{ji}^{(m)}: m < j, 1 \leqq i \leqq r_j\}$.

Next the set of elements $\{a_{ji}^{(j)}: 1 \leqq i \leqq r_j\}$ may be chosen in

$$t^{jr_j^2}F(t, r_j) \tag{13}$$

ways, by Lemma 2. Finally, since for $m > j$ there are exactly t^{jr_m} elements in $\pi^{m-j}W_m$, there are

$$t^{jr_j(r_{j+1}+\cdots+r_n)} \tag{14}$$

choices for the elements $\{a_{ji}^{(m)}: m > j, 1 \leqq i \leqq r_j\}$. The number of R-automorphisms of V is therefore

$$N_V = \prod_{j=1}^{n} \{t^{u_j}F(t, r_j)\},$$

where for each j,

$$u_j = \sum_{m=1}^{j} mr_m r_j + jr_j \sum_{m=j+1}^{n} r_m .$$

If we define the symbols $\{s_j\}$ by (7), a routine calculation establishes (8).

(The above generalizes the formula for N_V obtained by Fine-Herstein in pp. 500–502, loc. cit., where N_V is referred to as μ in their paper.)

Now let V range over a full set of non-isomorphic R-modules having exactly t^n elements, so that $\{r_1, \cdots, r_n\}$ range over all n-tuples of non-negative integers for which

$$n = r_1 + 2r_2 + \cdots + nr_n .$$

LEMMA 4. *As V ranges over the above-mentioned R-modules, we have*

$$\sum_V 1/N_V = 1/t^n F(t, n). \tag{15}$$

Proof. Use the formula (8) for N_V, and then apply Lemma 1 with $u = t^{-1}$.

3. Nilpotent matrix polynomials

Let K be a field with q elements, $f(x) \in K[x]$ an irreducible polynomial of degree $d \geqq 1$, and let n be a fixed integer. We wish to determine the number of matrices $X \in K_{nd}$ for which $f(X)$ is nilpotent. We remark that $f(X)$ is nilpotent if and only if $f^n(X) = 0$, since $f(X)$ is nilpotent if and only if the characteristic polynomial of X is $f^n(x)$.

Define the ring R by

$$R = K[x]/(f^n(x)),$$

and for each polynomial $g(x) \in K[x]$ let $\overline{g(x)}$ denote its image in R. Then R is a commutative ring of the type discussed in the preceding section, with maximal ideal πR, where $\pi = \overline{f(x)}$. We have $\pi^n = 0$, and the number t of elements in the field $R/\pi R$ is given by

$$t = q^d, \tag{16}$$

since $R/\pi R \cong K[x]/(f(x))$.

If V is any R-module of K-dimension nd, then V contains t^n elements. Furthermore V gives rise to a representation of R by matrices in K_{nd}, and the matrix X corresponding to $\bar{x}$ satisfies $f^n(X) = 0$. Conversely each such matrix X is obtainable in this way from some R-module with t^n elements.

For the rest of the proof we restrict ourselves to R-modules V with t^n elements. Each V gives rise to a set of equivalent matrix representations, and hence gives not only one matrix X corresponding to $\bar{x}$, but a system of matrices

$$\{P^{-1}XP : P \in K_{nd},\ P \text{ nonsingular}\}.$$

The number of distinct matrices in this system is just the number $q^{n^2d^2}F(q, nd)$ of nonsingular matrices in K_{nd}, divided by the number of nonsingular matrices $P \in K_{nd}$ satisfying

$$P^{-1}XP = X.$$

But since $\bar{x}$ generates the ring R, any such P yields an R-automorphism of V, and so there are N_V such nonsingular P's, where N_V is given by (8) with $t = q^d$.

On the other hand it is clear that non-isomorphic R-modules V, V^* give rise to matrices X, X^* which are not connected by any relation

$$X^* = P^{-1}XP, \qquad P \in K_{nd},\quad P \text{ nonsingular}.$$

The above discussion shows therefore that the number of matrices $X \in K_{nd}$ for which $f(X)$ is nilpotent is precisely

$$\sum\nolimits_V q^{n^2d^2}F(q, nd)/N_V,$$

where V ranges over a full set of non-isomorphic R-modules having t^n elements. By using (15), the above is just

$$q^{n^2d^2}F(q, nd)/q^{nd}F(q^d, n),$$

that is,

$$q^{n^2d^2-nd}\cdot\frac{(1-q^{-1})(1-q^{-2})\cdots(1-q^{-nd})}{(1-q^{-d})(1-q^{-2d})\cdots(1-q^{-nd})}.$$

This completes the proof of Theorem 1.

4. Matrices with given characteristic polynomial

We are now ready to prove Theorem 2. Let $g(x)$ be given by (3), and let

$$S = K[x]/(g(x)) = R_1 \oplus \cdots \oplus R_k,$$

where

$$R_i = K[x]/(f_i^{n_i}(x)), \qquad 1 \leqq i \leqq k.$$

Any S-module V can be decomposed into a direct sum

$$V = V_1 \oplus \cdots \oplus V_k ,$$

in which V_i is a left R_i-module, $1 \leqq i \leqq k$. We obtain all matrices $X \in K_n$ with characteristic polynomial $g(x)$ by letting V range over a full set of non-isomorphic S-modules of dimension n over K, chosen in such a way that

$$(V_1{:}K) = n_1 d_1 , \quad \cdots , \quad (V_k{:}K) = n_k d_k ,$$

and then for each such module V taking the set of matrices which correspond to $\bar{x} \in S$ (the image of $x \in K[x]$). Thus the number of matrices $X \in K_n$ with characteristic polynomial $g(x)$ is just

$$\textstyle\sum_V q^{n^2} F(q, n)/N_V .$$

It follows readily from the fact that the $\{f_i(x)\}$ are pairwise relatively prime that any S-automorphism of V maps each V_i onto itself, and thus is composed of a set of k automorphisms $\{\theta_i{:}1 \leqq i \leqq k\}$, where $\theta_i{:}V_i \rightarrow V_i$. Therefore

$$N_V = N_{V_1} \cdots N_{V_k} .$$

Furthermore, a full set of non-isomorphic S-modules V of the type described above is obtained by letting each V_i range independently over a full set of non-isomorphic R_i-modules with $(V_i{:}K) = n_i d_i$, for $i = 1, \cdots , k$. Thus the number of matrices $X \in K_n$ with characteristic polynomial $g(x)$ is

$$\begin{aligned} q^{n^2} F(q, n) \textstyle\sum_V 1/N_{V_1} \cdots N_{V_k} &= q^{n^2} F(q, n) \textstyle\prod_{i=1}^k \{ \textstyle\sum_{V_i} 1/N_{V_i} \} \\ &= q^{n^2} F(q, n) \cdot \{ \textstyle\prod_{i=1}^k q^{d_i n_i} F(q^{d_i}, n_i) \}^{-1}. \end{aligned}$$

Using the relation $n = \sum d_i n_i$, we obtain formula (4). This completes the proof of Theorem 2.

UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS