THE NUMBER OF LINEAR TRANSFORMATIONS DEFINED ON A SUBSPACE WITH GIVEN INVARIANT FACTORS

SAMRITH RAM

Abstract. Given a finite-dimensional vector space $V$ over the finite field $\mathbb{F}_q$ and a subspace $W$ of $V$, we consider the problem of counting linear transformations $T : W \to V$ which have prescribed invariant factors. The case $W = V$ is a well-studied problem that is essentially equivalent to counting the number of square matrices over $\mathbb{F}_q$ in a conjugacy class and an explicit formula is known in this case. On the other hand, the case of general $W$ is also an interesting problem and there hasn’t been substantive progress in this case for over two decades, barring a special case where all the invariant factors of $T$ are of degree zero. We extend this result to the case of arbitrary $W$ by giving an explicit counting formula. As an application of our results, we give new proofs of some recent enumerative results in linear control theory and derive an extension of the Gerstenhaber-Reiner formula for the number of square matrices over $\mathbb{F}_q$ with given characteristic polynomial.

1. Introduction

Let $\mathbb{F}_q$ denote the finite field with $q$ elements and let $\mathbb{F}_q[x]$ denote the ring of polynomials in the indeterminate $x$ with coefficients in $\mathbb{F}_q$. Throughout this paper, $n, k$ denote positive integers with $n \geq k$. For any ring $R$, we denote by $M_{n,k}(R)$ the set of all $n \times k$ matrices with entries in $R$ and by $M_n(R)$ the set of all square $n \times n$ matrices with entries in $R$. Let $\text{GL}_n(\mathbb{F}_q)$ denote the general linear group of nonsingular matrices of $M_n(\mathbb{F}_q)$.

Given any nonzero polynomial matrix $A \in M_{n,k}(\mathbb{F}_q[x])$, there exist invertible matrices $P \in M_n(\mathbb{F}_q[x])$ and $Q \in M_k(\mathbb{F}_q[x])$ such that the product $PAQ$ is of the
form
\[
\begin{bmatrix}
p_1 & 0 & 0 & \cdots & 0 \\
0 & p_2 & 0 & \cdots & 0 \\
0 & 0 & \ddots & \ddots & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots \\
0 & \cdots & \cdots & \cdots & 0
\end{bmatrix}
\]
where all off-diagonal entries are zero and the nonzero entries on the diagonal are monic polynomials \(p_i (1 \leq i \leq t)\) satisfying \(p_i \mid p_{i+1}\) for \(1 \leq i < t\). This diagonal form is known as the Smith normal form of \(A\). We express this concisely by writing \(A \sim \text{diag}_{n,k}(p_1, \ldots, p_t)\). The \(p_i\) are known as the invariant factors of \(A\) and are given by
\[
p_i = \frac{\delta_i(A)}{\delta_{i-1}(A)},
\]
where \(\delta_i(A)\) denotes the \(i^{th}\) determinantal divisor of \(A\) and equals the greatest common divisor of all \(i \times i\) minors of the matrix \(A\).

Let us denote by \(I_{n,k}\) the element of \(M_{n,k}(\mathbb{F}_q)\) whose \((i, j)\)-th entry is 1 for \(i = j\) and zero otherwise. In this paper, we are mainly interested in counting the number of matrices \(B \in M_{n,k}(\mathbb{F}_q)\) for which the matrix polynomial \(xI_{n,k} - B\) has a prescribed list of invariant factors. More precisely, for any \(k\)-tuple of invariant factors, i.e., a tuple \(I = (p_1, \ldots, p_k)\) of monic polynomials in \(\mathbb{F}_q[x]\) such that \(p_i \mid p_{i+1}\) \((1 \leq i \leq k - 1)\), define
\[
N_q(n, k; I) := \# \{ B \in M_{n,k}(\mathbb{F}_q) : xI_{n,k} - B \sim \text{diag}_{n,k}(p_1, \ldots, p_k) \}.
\]
We are interested in a formula for \(N_q(n, k; I)\). In fact, the problem of determining \(N_q(n, n; I)\) is equivalent to counting the number of square \(n \times n\) matrices over \(\mathbb{F}_q\) in a conjugacy class. This problem has been studied by Kung [11] and Stong [17] among others, and an explicit formula (see (1)) due to Philip Hall can be found in Stanley [16].

On the other hand, rectangular matrices arise in many contexts in linear control theory and matrix completion problems and are also of considerable interest. We refer to Cravo [1, Thms. 15, 32] for specific examples of matrix completion problems. For another example, suppose \(k < n\) and \(\mathcal{I}_0\) denotes the \(k\)-tuple \((1, \ldots, 1)\). Determining \(N_q(n, k; \mathcal{I}_0)\) is equivalent to the problem of counting the number of zero kernel pairs over a finite field. We refer to Gohberg et al. [6, Sec. X.1] or [12] for the definition of zero kernel pairs. The problem of counting zero kernel pairs goes back at least to the work of Kociecki and Przyłuski [10] who considered the equivalent problem of determining the number of reachable pairs (see Definition 4.1) of matrices over a finite field. This problem has recently been settled by Helmke et al. [8, Thm. 1] who also solve the more general problem
of determining the number of matrix pairs with an \( r \)-dimensional reachability subspace (see Definition 4.2).

The main result of this paper is Theorem 3.6 which is a counting formula for the number of linear transformations defined on a subspace which have a given invariant subspace and prescribed invariant factors. We use this result to deduce Theorem 3.8 which gives a formula for \( N_q(n, k; \mathcal{I}) \) for an arbitrary \( k \)-tuple \( \mathcal{I} \) of invariant factors. We give further applications of our results in Sections 4 and 5. In Section 4 we give an alternate proof of the formula of Helmke et al. (Theorem 4.6) for the number of matrix pairs with an \( r \)-dimensional reachability subspace. Studying linear maps defined on a subspace allows us to extend some combinatorial results on square matrices to the setting of rectangular matrices. As an illustration, we prove an extension of the Gerstenhaber-Reiner formula for the number of square matrices with given characteristic polynomial in Section 5.

2. Preliminaries and Notation

It is a well-known fact that two matrices \( A, A' \in M_n(\mathbb{F}_q) \) are similar, i.e., there exists \( P \in \text{GL}_n(\mathbb{F}_q) \) with \( A' = PAP^{-1} \), if and only if they have the same invariant factors. More precisely, if \( xI_n - A \sim \text{diag}_{n,n}(p_1, \ldots, p_n) \) and \( xI_n - A' \sim \text{diag}_{n,n}(p'_1, \ldots, p'_n) \), then \( A \) is similar to \( A' \) if and only if \( p_i = p'_i \) for \( 1 \leq i \leq n \). Therefore, the conjugacy class of all matrices similar to \( A \) is uniquely determined by the invariant factors of \( A \). To give an expression for the number of matrices in the conjugacy class determined by an \( n \)-tuple of invariant factors, we introduce some notation.

Following Stanley [16, Sec. 1.10], we denote by \( \text{Par} \) the set of all partitions of all nonnegative integers. We write \( \lambda \vdash n \) if \( \lambda \) is a partition of \( n \). Given any partition \( \lambda \in \text{Par} \), let \( \lambda' = (\lambda'_1, \lambda'_2, \ldots) \) denote the conjugate partition to \( \lambda \) and let \( m_i = m_i(\lambda) = \lambda'_1 - \lambda'_{i+1} \) denote the number of parts of \( \lambda \) of size \( i \). Let

\[
h_i = h_i(\lambda) = \lambda'_1 + \lambda'_2 + \cdots + \lambda'_i.
\]

For \( \lambda \in \text{Par} \) and any irreducible \( f \in \mathbb{F}_q[x] \) of degree \( d \), define

\[
c_f(\lambda) := \prod_{i \geq 1} \prod_{j=1}^{m_i} (q^{dh_i} - q^{d(h_i-j)})
\]

For \( f, g \in \mathbb{F}_q[x] \) with \( f \) irreducible and \( g \) nonzero, set

\[
\nu_f(g) := \max\{e : f^e | g\}.
\]

Given a tuple \( \mathcal{I} = (p_1, \ldots, p_k) \) of invariant factors, let \( S(\mathcal{I}) \) denote the set of all monic irreducible polynomials which divide \( P = \prod_{i=1}^{k} p_i \). To each \( f \in S(\mathcal{I}) \), we associate a partition \( \Phi_f(P) \) by defining

\[
\Phi_f(P) := (\nu_f(p_k), \nu_f(p_{k-1}), \ldots, \nu_f(p_r)),
\]

where \( r \) is the largest integer such that \( f^{r+1} \) divides \( P \).
where \( r = \min \{ i : \nu_f(p_i) > 0 \} \). The following formula (essentially equivalent to [16, Eq. 1.107]) gives the size of the conjugacy class corresponding to a tuple of invariant factors:

\[
N_q(n, n; \mathcal{I}) = \frac{|\text{GL}_n(\mathbb{F}_q)|}{\prod_{f \in S(\mathcal{I})} e_f(\phi_f(\mathcal{I}))}
\]

for any \( n \)-tuple \( \mathcal{I} = (p_1, \ldots, p_n) \) of invariant factors with \( \deg(p_1 \cdots p_n) = n \). In the next section, we recall some basic facts about linear transformations defined on a subspace and then prove an extension (Theorem 3.8) of the above formula.

3. Maps defined on a subspace with prescribed invariant factors

Throughout this paper, we denote by \( V \) an \( n \)-dimensional vector space over \( \mathbb{F}_q \) and by \( W \) a fixed \( k \)-dimensional subspace of \( V \). We denote by \( L(W, V) \) the space of all linear transformations from \( W \) to \( V \). Let \( \mathcal{B}_1 = \{v_1, \ldots, v_k\} \) be a fixed ordered basis for \( W \) and let \( \mathcal{B}_2 = \{v_1, \ldots, v_n\} \) be an extension of this basis to an ordered basis of \( V \). We define the invariant factors of a linear transformation defined on a subspace in terms of its matrix with respect to the bases \( \mathcal{B}_1 \) and \( \mathcal{B}_2 \).

**Definition 3.1.** Suppose \( T \in L(W, V) \) and let \( B \in M_{n,k}(\mathbb{F}_q) \) denote the matrix of \( T \) with respect to the bases \( \mathcal{B}_1 \) and \( \mathcal{B}_2 \). The invariant factors of \( T \) are the invariant factors of \( xI_{n,k} - B \) viewed as a polynomial matrix in \( M_{n,k}(\mathbb{F}_q[x]) \).

Note that the above definition coincides with the usual definition of invariant factors of a linear transformation in the case \( W = V \). The invariant factors of \( T \in L(W, V) \) are independent of the specific choice of bases for \( W \) and \( V \) so long as the basis for \( V \) is obtained as an extension of the basis for \( W \). If \( p_1, \ldots, p_k \) are the invariant factors of \( T \), we write \( \mathcal{I}_T = (p_1, \ldots, p_k) \). Though the terms ‘invariant factors’ and ‘invariant polynomials’ are often used interchangeably, some authors refer to the nonunit invariant factors of \( T \) as the invariant polynomials of \( T \).

**Remark 3.2.** While it is possible for an arbitrary element of \( M_{n,k}(\mathbb{F}_q[x]) \) to have less than \( k \) invariant factors, a matrix polynomial of the form \( xI_{n,k} - B \) for \( B \in M_{n,k}(\mathbb{F}_q) \) always has \( k \) invariant factors as \( \delta_k(xI_{n,k} - B) \neq 0 \). In fact, it is easily seen that \( \delta_k(xI_{n,k} - B) \) is a polynomial of degree at most \( k \).

**Definition 3.3.** For any \( k \)-tuple \( \mathcal{I} \) of invariant factors, we define

\[
N_q(n, k; \mathcal{I}) := \# \{ T \in L(W, V) : \mathcal{I}_T = \mathcal{I} \}.
\]

The above definition of \( N_q(n, k; \mathcal{I}) \) is equivalent to the one given in the introduction. Given \( T \in L(W, V) \), define

\[
\mathcal{J}(T) := \text{the maximal } T\text{-invariant subspace contained in } W.
\]

The following proposition [6, Thm. III.1.2] relates the invariant factors of \( T \) to the invariant factors of its restriction to \( \mathcal{J}(T) \).
**Proposition 3.4.** For $T \in L(W, V)$, let $T'$ denote the restriction of $T$ to its maximal invariant subspace $\mathfrak{I}(T)$. Then $T$ and $T'$ have the same nonunit invariant factors.

**Corollary 3.5.** For $T \in L(W, V)$ with $\mathcal{I}_T = (p_1, \ldots, p_k)$, we have

$$\dim \mathfrak{I}(T) = \deg \prod_{i=1}^{k} p_i.$$  

**Proof.** Denote by $T'$ the restriction of $T$ to $\mathfrak{I}(T)$ and suppose $\dim \mathfrak{I}(T) = d$. Let $A \in M_{n,k}(F_q)$ denote the matrix of $T$ with respect to the bases $B_1$ and $B_2$ and let $A' \in M_d(F_q)$ denote the matrix of $T'$ with respect to some basis of $\mathfrak{I}(T)$. The dimension of $\mathfrak{I}(T)$ is equal to the degree of the characteristic polynomial of $T'$, i.e.,

$$\dim \mathfrak{I}(T) = d = \deg \det(xI_d - A')$$  

$$= \deg \delta_d(xI_d - A')$$  

$$= \deg \delta_k(xI_{n,k} - A).$$

The last equality follows from Proposition 3.4 since $T$ and $T'$ have the same nonunit invariant factors. Since

$$\delta_k(xI_{n,k} - A) = \prod_{i=1}^{k} p_i,$$

the corollary follows. \qed

For subspaces $U \subseteq W \subseteq V$ and a tuple $\mathcal{I}$ of invariant factors, we define

$$N(V, W, U; \mathcal{I}) := \# \{ T \in L(W, V) : \mathfrak{I}(T) = U \text{ and } \mathcal{I}_T = \mathcal{I} \}.$$ 

Then (1) is equivalent to

$$N(V, V, U; \mathcal{I}) = |GL_d(F_q)| \prod_{f \in S(\mathcal{I})} c_f(\Phi_\mathcal{I}(f))$$

for any $n$-tuple $\mathcal{I} = (p_1, \ldots, p_n)$ of invariant factors with $\deg \prod_{i=1}^{n} p_i = n$. Note that $N(V, V, U; \mathcal{I}) = 0$ for all $\mathcal{I}$ whenever $U$ is a proper subspace of $V$.

The next theorem gives an explicit formula for $N(V, W, U; \mathcal{I})$ and is our main result. A crucial ingredient in the proof is the formula for the number of zero kernel pairs over a finite field.

**Theorem 3.6.** Suppose $U$ is a subspace of $W$ and let $d, k, n$ denote the dimensions of $U, W, V$ respectively. Given a $k$-tuple $\mathcal{I} = (p_1, \ldots, p_k)$ of invariant factors with $\deg \prod_{i=1}^{k} p_i = d$, we have

$$N(V, W, U; \mathcal{I}) = \frac{|GL_d(F_q)|}{\prod_{f \in S(\mathcal{I})} c_f(\Phi_\mathcal{I}(f))} \prod_{i=d+1}^{k} (q^n - q^i).$$
Proof. The case $W = V$ follows easily from (2). Now suppose $W \not\subseteq V$. Let $T \in L(W, V)$ with $\mathcal{I}_T = (p_1, \ldots, p_k)$. We need to count the number of such $T$ for which $\mathcal{I}(T) = U$. Let $U'$ be a complementary subspace of $W$ with $W = U \oplus U'$. Let $\{v_1, \ldots, v_d\}$ be an ordered basis for $U$ and let $\{v_{d+1}, \ldots, v_k\}$ be an ordered basis for $U'$ so that $B = \{v_1, \ldots, v_k\}$ is a basis of $W$. We extend this basis further to a basis $B' = \{v_1, \ldots, v_n\}$ of $V$. We now count linear transformations by counting their matrices with respect to the bases $B$ and $B'$. Accordingly, let $[A_{ij}] \in M_{n,k}(\mathbb{F}_q)$ be the matrix of $T$ with respect to the bases $B$ and $B'$ for some $A \in M_k(\mathbb{F}_q)$ and $C \in M_{n-k,k}(\mathbb{F}_q)$. We partition $A$ and $C$ further into blocks corresponding to the dimensions of $U$ and $U'$ and write the matrix of $T$ in the form

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ C_{11} & C_{12} \end{bmatrix},$$

where $A_{11} \in M_d(\mathbb{F}_q)$. Since $U$ is $T$-invariant, it follows that $A_{21} = 0 \in M_{k-d,d}(\mathbb{F}_q)$ and $C_{11} = 0 \in M_{n-k,d}(\mathbb{F}_q)$. We thus have

$$\begin{bmatrix} xI_k - A \\ -C \end{bmatrix} = \begin{bmatrix} xI_d - A_{11} & -A_{12} \\ 0 & xI_{k-d} - A_{22} \\ 0 & -C_{12} \end{bmatrix}.$$

We now need to count the number of tuples $(A_{11}, A_{12}, A_{22}, C_{12})$ for which the invariant factors of (3) are precisely $p_1, \ldots, p_k$. The fact that $\deg(p_1 \cdots p_k) = d$ forces $p_i = 1$ for $i \leq k - d$. Since the matrix of $T$ restricted to $U$ is precisely $A_{11}$, it follows from Proposition 3.4 that the invariant factors of $xI_d - A_{11}$ are precisely $p_{k-d+1}, \ldots, p_k$, i.e.,

$$xI_d - A_{11} \sim \text{diag}_{d} (p_{k-d+1}, \ldots, p_{k}).$$

Note that $\det(xI_d - A_{11}) = \prod_{i=1}^{k} p_i$. Consider the $k \times k$ minors of the matrix on the right hand side of (3). For such a minor to be nonzero, it necessarily contains all rows of the submatrix $[xI_d - A_{11} \quad -A_{12}]$. Thus, the nonzero $k \times k$ minors of $[xI_k - A] - C$ are all divisible by $\det(xI_d - A_{11})$. Now observe that

$$\delta_{k} \left( \begin{bmatrix} xI_k - A \\ -C \end{bmatrix} \right) = \det(xI_d - A_{11}) \cdot \delta_{k-d} \left( \begin{bmatrix} xI_{k-d} - A_{22} \\ -C_{12} \end{bmatrix} \right).$$

Since $\det(xI_d - A_{11}) = \prod_{i=1}^{k} p_i$ which also equals the left hand side of (5), it follows that

$$\delta_{k-d} \left( \begin{bmatrix} xI_{k-d} - A_{22} \\ -C_{12} \end{bmatrix} \right) = 1.$$

From the preceding discussion, it is clear that if $A_{11}$ is chosen to satisfy (4), $A_{22}, C_{12}$ are chosen to satisfy (6) and $A_{12}$ is chosen arbitrarily in $M_{d,k-d}(\mathbb{F}_q)$, then $T$ has the desired invariant factors. Conversely, for any $T$ with invariant factors $p_1, \ldots, p_k$, the matrices $A_{11}, A_{12}, A_{22}, C_{12}$ necessarily satisfy (4) and (6).
The number of choices for \( A_{11} \) is precisely \( N_q(d, d; \mathcal{I}(d)) \) where \( \mathcal{I}(d) \) denotes the \( d \)-tuple formed by taking the last \( d \) coordinates of \( \mathcal{I} \). The pair \((C_{12}, A_{22}) \in M_{n-k,k-d}(\mathbb{F}_q) \times M_{k-d}(\mathbb{F}_q)\) may be chosen to be any zero kernel pair. This can be done in \( \prod_{i=1}^{k-d}(q^{n-d} - q^i) \) ways ([12, Cor. 3.9]) and \( A_{12} \) can be chosen in \( q^{d(k-d)} \) ways. Thus, the tuple \((A_{11}, A_{12}, A_{22}, C_{12}) \) can be chosen in

\[
(q^{d(k-d)}) N_q(d, d; \mathcal{I}(d)) \prod_{i=1}^{k-d}(q^{n-d} - q^i) = N_q(d, d; \mathcal{I}(d)) \prod_{i=d+1}^{k} (q^n - q^i)
\]

ways. It is easy to see that \( S(\mathcal{I}) = S(\mathcal{I}(d)) \) and \( \Phi_\mathcal{I}(f) = \Phi_{\mathcal{I}(d)}(f) \) for all \( f \in S(\mathcal{I}) \). The result now follows by substituting for \( N_q(d, d; \mathcal{I}(d)) \) from (1).

**Remark 3.7.** If we set \( U \) to be the zero subspace in the hypothesis of Theorem 3.6, then \( d = 0 \) and \( \mathcal{I} \) is necessarily the \( k \)-tuple \((1, \ldots, 1)\). In this case \( \mathcal{I}(0) \) denotes the null tuple and we adopt the convention that \( N_q(0, 0; \mathcal{I}(0)) = 1 \) and \( |\text{GL}_0(\mathbb{F}_q)| = 1 \) in the counting formula. We then obtain

\[
N(V, W, \{0\}, (1, \ldots, 1)) = \prod_{i=1}^{k} (q^n - q^i),
\]

which is an expression for the number of \( T \in L(W, V) \) which have no nonzero invariant subspace (see [12, Prob. 1.4]).

The number of \( d \)-dimensional subspaces of a \( k \)-dimensional vector space over \( \mathbb{F}_q \) is well-known and given by the \( q \)-binomial coefficient

\[
\binom{k}{d}_q := \frac{(q^k - 1)(q^{k-1} - 1) \cdots (q^{k-d+1} - 1)}{(q^d - 1)(q^{d-1} - 1) \cdots (q - 1)}.
\]

The following theorem gives a formula for the number of linear transformations defined on a subspace with a prescribed list of invariant factors.

**Theorem 3.8.** Let \( \mathcal{I} = (p_1, \ldots, p_k) \) be a \( k \)-tuple of invariant factors and suppose \( \deg \prod_{i=1}^{k} p_i = d \). We have

\[
N_q(n, k; \mathcal{I}) = \binom{k}{d}_q \frac{|\text{GL}_d(\mathbb{F}_q)|}{\prod_{f \in S(\mathcal{I})} c_f(\Phi_{\mathcal{I}}(f))} \prod_{i=d+1}^{k} (q^n - q^i).
\]

**Proof.** We sum over all \( d \)-dimensional subspaces \( U \) of \( W \) to obtain

\[
N_q(n, k; \mathcal{I}) = \sum_{\dim U = d} N(V, W; U; \mathcal{I}) = \binom{k}{d}_q N(V, W; U_0; \mathcal{I}),
\]

where \( U_0 \) is some fixed subspace of \( W \) of dimension \( d \). We now use Theorem 3.6 to deduce the result. \( \square \)
The following result gives an expression for the number of linear transformations which have a given invariant subspace and will be used later on in the proof of Theorem 4.6.

**Theorem 3.9.** If $U$ is a $d$-dimensional subspace of $W$, then

$$\# \{ T \in L(W, V) : \mathcal{I}(T) = U \} = q^{d^2} \prod_{i=d+1}^{k} (q^n - q^i).$$

**Proof.** Summing (7) over all possible $k$-tuples $\mathcal{I} = (p_1, \ldots, p_k)$ of invariant factors with $\deg \prod_{i=1}^{k} p_i = d$, we obtain

$$\# \{ T \in L(W, V) : \mathcal{I}(T) = U \} = \sum_{\mathcal{I}} N(V, W; \mathcal{I}) = \sum_{\mathcal{I}} N_q(d, d; \mathcal{I}(d)) \prod_{i=d+1}^{k} (q^n - q^i),$$

where $\mathcal{I}(d)$ denotes the $d$-tuple formed by the last $d$ coordinates of $\mathcal{I}$. This in turn equals

$$\prod_{i=d+1}^{k} (q^n - q^i) \sum_{\mathcal{I}} N_q(d, d; \mathcal{I}(d)) = q^{d^2} \prod_{i=d+1}^{k} (q^n - q^i).$$

The last equality follows from the fact that $\mathcal{I}(d)$ varies over invariant factors corresponding to all possible conjugacy classes in $M_d(\mathbb{F}_q)$ and hence the above sum counts all elements of $M_d(\mathbb{F}_q)$. \qed

**Remark 3.10.** In contrast with the case of linear operators, the invariant factors do not form a complete set of similarity invariants for a linear transformation defined only on a subspace. Recall that if $T \in L(W, V)$ and $T' \in L(W', V')$ denote two linear transformations defined on subspaces $W, W'$ of $V, V'$ respectively, then $T$ and $T'$ are similar if there exists a linear isomorphism $S : V \to V'$ with $S(W) = W'$ such that the following diagram commutes:

$$\begin{array}{ccc}
W & \xrightarrow{T} & V \\
\downarrow{s} & & \downarrow{s} \\
W' & \xrightarrow{T'} & V'
\end{array}$$

If $T$ and $T'$ are similar, then they necessarily have the same invariant factors but the converse is not true in general. We refer to Gohberg et al. [6, Sec. III.3] or Ferrer and Puerta [2] for a characterization of similarity invariants in this case.

### 4. Reachability subspace of a matrix pair

In this section we assume that $k, n$ are positive integers with $k < n$. Let $x_0$ denote a fixed column vector in $\mathbb{F}_q^k$ and consider the sequence $x_s$ in $\mathbb{F}_q^k$ defined by

$$x_{s+1} = Ax_s + Bu_s \quad (s \geq 0),$$
where \( A \in M_k(\mathbb{F}_q) \), \( B \in M_{k,n-k}(\mathbb{F}_q) \) are fixed and \( u_s(s \geq 0) \) is some sequence in \( \mathbb{F}_q^{n-k} \). This sequence arises frequently in diverse areas such as control systems, linear sequential machines, discrete automata theory and convolutional error correcting codes just to name a few (see [4], [7], [15]). The recurrence is readily solved for \( x_k \) by backward substitution and we obtain

\[
x_k = A^k x_0 + \begin{bmatrix} B & AB & \cdots & A^{k-1}B \end{bmatrix} \begin{bmatrix} u_{k-1} \\ u_{k-2} \\ \vdots \\ u_0 \end{bmatrix}.
\]

It is often required that \( x_k \) take all possible values in \( \mathbb{F}_q^k \) by suitably varying the ‘input sequence’ \( u_s(0 \leq s \leq k - 1) \). This happens if and only if the reachability matrix defined by

\[
C(A, B) := [B \ AB \ \cdots \ \ A^{k-1}B]
\]

has full rank \( k \). This motivates the definition of a reachable pair of matrices.

**Definition 4.1.** The pair \((A, B) \in M_k(\mathbb{F}_q) \times M_{k,n-k}(\mathbb{F}_q)\) is said to reachable if

\[
\text{rank} C(A, B) = k.
\]

The problem of counting the number of reachable pairs \((A, B)\) over a finite field is equivalent to counting the number of zero kernel pairs and is an interesting problem in its own right with connections to other combinatorial objects. We refer to [12] for more on this topic. When the pair \((A, B)\) is not reachable, \( x_k \) takes values in an affine subspace of dimension \( r \leq k \) determined by \( x_0 \) and the reachability matrix \( C(A, B) \). We have the following natural generalization of the notion of reachability.

**Definition 4.2.** The pair \((A, B) \in M_k(\mathbb{F}_q) \times M_{k,n-k}(\mathbb{F}_q)\) is said to have an \( r \)-dimensional reachability subspace if

\[
\text{rank} C(A, B) = r.
\]

A natural question at this point is to ask whether there is a simple formula for the number of matrix pairs with reachability subspace of a given dimension. This is the question addressed by Theorem 4.6. We first establish the connection between linear transformations defined on a subspace and matrix pairs with reachability subspace of a given dimension.

**Proposition 4.3.** Suppose \( T \in L(W, V) \) and let the matrix of \( T \) with respect to the bases \( B_1 \) and \( B_2 \) be \( \begin{bmatrix} A \\ C \end{bmatrix} \in M_{n,k}(\mathbb{F}_q) \) for some \( A \in M_k(\mathbb{F}_q) \) and \( C \in M_{n-k,k}(\mathbb{F}_q) \). We have an isomorphism of \( \mathbb{F}_q \)-vector spaces

\[
\mathcal{I}(T) \simeq \bigcap_{i=0}^{k-1} \ker(CA^i).
\]
Proof. The proof is similar to that of [12, Prop. 3.2]. For any \( v \in W \), let \( x_v \in \mathbb{F}_q^k \) denote the coordinate column vector of \( v \) with respect to the basis \( B_1 \). If \( v \in \mathcal{I}(T) \), then \( T^i v \) lies in \( \mathcal{I}(T) \) and hence in \( W \) for all \( i \). The coordinate vector of \( T^i v \) with respect to \( B_2 \) is

\[
\begin{bmatrix}
A \\
C
\end{bmatrix} x_v = \begin{bmatrix} A x_v \\ C x_v \end{bmatrix}.
\]

Since \( T^i v \in W \), it follows that \( C x_v = 0 \) and the coordinate vector of \( T^i v \) with respect to \( B_1 \) is simply \( A x_v \). By considering \( T^2 v \) it follows similarly that \( CA x_v = 0 \) and hence the coordinate vector of \( T^2 v \) with respect to \( B_1 \) is \( A^2 x_v \). Continuing this line of reasoning, we find that \( CA^i x_v = 0 \) for \( 0 \leq i \leq k - 1 \). Thus

\[
x_v \in \bigcap_{i=0}^{k-1} \ker(CA^i),
\]

which proves that

\[
\{ x_v : v \in \mathcal{I}(T) \} \subseteq \bigcap_{i=0}^{k-1} \ker(CA^i). \tag{8}
\]

Conversely, suppose \( x \in \bigcap_{i=0}^{k-1} \ker(CA^i) \). Let \( \chi_A(x) \) denote the characteristic polynomial of \( A \). If \( u \) is the vector whose coordinates with respect to \( B_1 \) are precisely \( x \), then it is easily verified that \( \chi_A(T)u \) is well-defined and equal to 0. It follows that the \( \mathbb{F}_q \)-linear subspace spanned by \( u, Tu, \ldots, T^{k-1}u \) is invariant under \( T \). Therefore \( u \) generates a \( T \)-invariant subspace and hence \( u \in \mathcal{I}(T) \). Thus the reverse inclusion in (8) holds as well. We thus have

\[
\{ x_v : v \in \mathcal{I}(T) \} = \bigcap_{i=0}^{k-1} \ker(CA^i),
\]

from which the proposition immediately follows. \( \square \)

**Corollary 4.4.** Suppose \( \begin{bmatrix} A \\ C \end{bmatrix} \in M_{n,k}(\mathbb{F}_q) \) denotes the matrix of \( T \in L(W, V) \). We have

\[
\dim \mathcal{I}(T) = \dim \bigcap_{i=0}^{k-1} \ker(CA^i)
\]

\[
= k - \text{rank} \begin{bmatrix}
C \\
CA \\
\vdots \\
CA^{k-1}
\end{bmatrix}.
\]

Proof. The second equality follows from the rank-nullity theorem since the null space of the block matrix above is precisely \( \bigcap_{i=0}^{k-1} \ker(CA^i) \). \( \square \)
Corollary 4.5. The pair \((A, B) \in M_k(\mathbb{F}_q) \times M_{k,n-k}(\mathbb{F}_q)\) has a reachability subspace of dimension \(k - \dim \mathcal{I}(T)\) where \(T : W \to V\) is the linear transformation defined by declaring the matrix of \(T\) with respect to \(B_1\) and \(B_2\) to be \(ATB\).

Proof. Follows from Definition 4.2 and Corollary 4.4. \(\square\)

We are now ready to prove the result of Helmke et al. [8, Thm. 2] on matrix pairs with reachability subspace of a given dimension.

Theorem 4.6. The number of pairs \((A, B) \in M_k(\mathbb{F}_q) \times M_{k,n-k}(\mathbb{F}_q)\) with \(r\)-dimensional reachability subspace is given by

\[
\binom{k}{r}_q \cdot q^{(k-r)^2} \prod_{i=k-r+1}^{k} (q^n - q^i).
\]

Proof. The number of pairs \((A, B)\) with \(r\)-dimensional reachability subspace is equal to

\[
\# \{(A, B) \in M_k(\mathbb{F}_q) \times M_{k,n-k}(\mathbb{F}_q) : \text{rank} \mathcal{C}(A, B) = r\}.
\]

By Corollary 4.5, this equals

\[
\# \{T \in L(W, V) : \dim \mathcal{I}(T) = d\},
\]

where \(d = k - r\). As \(\mathcal{I}(T)\) could be any \(d\)-dimensional subspace of \(W\), the number of possibilities for \(\mathcal{I}(T)\) is clearly \(\binom{k}{d}_q = \binom{k}{r}_q\). The result now follows from Theorem 3.9. \(\square\)

Setting \(r = k\) in the above theorem, we find that the number of reachable pairs \((A, B) \in M_k(\mathbb{F}_q) \times M_{k,n-k}(\mathbb{F}_q)\) is given by

\[
\prod_{i=1}^{k} (q^n - q^i).
\]

Remark 4.7. Let \(\mathcal{I}_1\) denote the \(k\)-tuple \((1, \ldots, 1)\). The formula for the number of reachable pairs is equivalent to

\[
N_q(n, k; \mathcal{I}_1) = \prod_{i=1}^{k} (q^n - q^i).
\]

5. An Extension of the Gerstenhaber-Reiner Formula

It is a remarkable fact that the number of nilpotent matrices in \(M_n(\mathbb{F}_q)\) is given by a rather simple formula, namely, \(q^{n(n-1)}\). This result is often attributed to Fine and Herstein [3] but appears to have been known earlier to Philip Hall who gave two different proofs in a series of lectures at the Edinburgh mathematical colloquium at St. Andrews in 1955 (see [14, p. 618]). Subsequently, the result was generalized independently by Gerstenhaber [5] and Reiner [13, Thm. 2] who gave a formula for the number of matrices with a given characteristic polynomial.
In this section, we extend the result of Gerstenhaber and Reiner to the setting of rectangular matrices.

Let $\Phi_n : M_n(\mathbb{F}_q) \to \mathbb{F}_q[x]$ denote the map which associates to a square $n \times n$ matrix its characteristic polynomial, i.e., $\Phi_n(A) = \det(xI_n - A)$. The fibers of $\Phi_n$ can be computed explicitly as the following theorem shows.

**Theorem 5.1** (Gerstenhaber-Reiner). For a monic polynomial $f(x) \in \mathbb{F}_q[x]$ of degree $n$, let $f = f_1^{e_1} \cdots f_t^{e_t}$ denote the canonical factorization of $f$ into distinct irreducible polynomials $f_i$ of degree $d_i (1 \leq i \leq t)$ in $\mathbb{F}_q[x]$. The number of matrices in $M_n(\mathbb{F}_q)$ with characteristic polynomial $f$ is given by

$$|\Phi_n^{-1}(f)| = q^{n^2-n} \frac{F(q,n)}{\prod_{i=1}^t F(q^{d_i}, e_i)},$$

where $F(q, r) = \prod_{i=1}^r (1 - q^{-i})$ for each nonnegative integer $r$.

We extend this result to rectangular matrices as follows. Define

$$\Phi_{n,k} : M_{n,k}(\mathbb{F}_q) \to \mathbb{F}_q[x]$$

by

$$\Phi_{n,k}(A) = \delta_n(xI_{n,k} - A),$$

where $\delta_n(\cdot)$ denotes the $k$th determinantal divisor as before. It is clear that $\Phi_{n,n}$ is just the map $\Phi_n$. The following result gives a formula for the size of the fibers of $\Phi_{n,k}$ and may be viewed as an extension of the Gerstenhaber-Reiner formula.

**Theorem 5.2.** For a monic polynomial $f(x) \in \mathbb{F}_q[x]$ of degree $d \leq k$, let $f = f_1^{e_1} \cdots f_t^{e_t}$ be the canonical factorization of $f$ into distinct irreducibles $f_i$ of degree $d_i (1 \leq i \leq t)$. We have

$$|\Phi_{n,k}^{-1}(f)| = \begin{bmatrix} k \\ d \end{bmatrix} q^{d^2-d} \frac{F(q,d)}{\prod_{i=1}^t F(q^{d_i}, e_i)} \prod_{i=d+1}^k (q^n - q^i).$$

**Proof.** We have $\Phi_{n,k}(A) = f$ precisely when the product of the diagonal elements appearing in the Smith form of $xI_{n,k} - A$ is $f$. Therefore, if we sum over all possible $k$-tuples $(p_1, \ldots, p_k)$ of monic polynomials in $\mathbb{F}_q[x]$ with $p_i \mid p_{i+1} (1 \leq i \leq k-1)$ and $p_1 \cdots p_k = f$, we obtain

$$|\Phi_{n,k}^{-1}(f)| = \sum_{\substack{p_1 \cdots p_k = f \\ p_i \mid p_{i+1}}} N_q(n,k; (p_1, \ldots, p_k))$$

$$= \begin{bmatrix} k \\ d \end{bmatrix} \prod_{i=d+1}^k (q^n - q^i) \sum_{\substack{p_1 \cdots p_k = f \\ p_i \mid p_{i+1}}} N_q(d,d; (p_{k-d+1}, \ldots, p_k)), $$

where the second equality follows from Theorem 3.8 and (1). The sum in the last expression counts all $d \times d$ matrices for which the product of invariant factors
is \( f \) and, therefore, is clearly equal to the number of matrices in \( M_d(\mathbb{F}_q) \) with characteristic polynomial \( f \). Thus

\[
|\Phi_{n,k}^{-1}(f)| = \binom{k}{d}_q |\Phi_{d}^{-1}(f)| \prod_{i=d+1}^{k} (q^n - q^i),
\]

and the theorem follows by applying the Gerstenhaber-Reiner formula to \( |\Phi_{d}^{-1}(f)| \).

\[ \square \]

We conclude this section with an interesting application of Theorem 5.2. Consider the following combinatorial problem: Given an \( n \)-dimensional vector space \( V \) over \( \mathbb{F}_q \) and a \( k \)-dimensional subspace \( W \) of \( V \), how many linear transformations \( T : W \to V \) have the property that \( T \) can be extended to a nilpotent transformation on all of \( V \)? We shall show that the answer is \( q^{n(k-1)}(q^n - q^k + 1) \). Note that this extends Hall’s result which says that the answer for \( k = n \) must be \( q^{n(n-1)} \). Before we proceed, we need a lemma.

**Lemma 5.3.** For any nonnegative integer \( d \) and indeterminate \( y \),

\[
y^d = \sum_{j=0}^{d} q^j \prod_{i=j+1}^{d} (y - q^i) \binom{d}{j}_q.
\]

**Proof.** We begin with the identity (see [9, Ch. 4, Eq. 4.4])

\[
z^d = \sum_{j=0}^{d} \binom{d}{j}_q \prod_{l=0}^{j-1} (z - q^l)
\]

\[
= \sum_{j=0}^{d} q^{j(d-j)} \binom{d}{j}_q \prod_{l=0}^{j-1} (z - q^l)
\]

\[
= \sum_{j=0}^{d} q^{j(d-j)} \binom{d}{j}_q \prod_{i=d-j+1}^{d} (z - q^{d-i}).
\]

Since the left hand side is independent of \( q \), we replace \( q \) by \( 1/q \) to obtain

\[
z^d = \sum_{j=0}^{d} q^{j-d} \binom{d}{j}_q \prod_{i=d-j+1}^{d} (z - q^{i-d})
\]

\[
= \sum_{j=0}^{d} q^{j-2d} \binom{d}{j}_q \prod_{i=d-j+1}^{d} (q^d z - q^i).
\]
Put $q^dz = y$ to obtain

$$y^d = q^{d^2} \sum_{j=0}^{d} q^{j(2d-j)} \prod_{i=d-j+1}^{d} (y - q^i)$$

$$= \sum_{j=0}^{d} q^{(d-j)^2} \prod_{i=d-j+1}^{d} (y - q^i)$$

$$= \sum_{j=0}^{d} q^{j^2} \prod_{i=j+1}^{d} (y - q^i)$$

as desired. □

**Theorem 5.4.** Let $V$ be an $n$-dimensional vector space over $\mathbb{F}_q$ and let $W$ be a $k$-dimensional subspace of $V$. The number of linear transformations $T : W \to V$ that can be extended to a nilpotent transformation on all of $V$ is given by

$$q^{n(k-1)}(q^n - q^k + 1).$$

**Proof.** Let $\Delta_q(k, n)$ denote the number of linear transformations $T : W \to V$ which can be extended to a nilpotent transformation on $V$. It follows from Wimmer’s theorem [1, Th. 15] that the product of invariant factors of any such $T$ is a divisor of $x^n$ and is of degree at most $k$. Therefore,

$$\Delta_q(k, n) = \sum_{l=0}^{k} |\Phi_{n,k}^{-1}(x^l)|$$

$$= \sum_{l=0}^{k} |\Phi^{-1}(x^l)| \prod_{i=l+1}^{k} (q^n - q^i)$$

$$= \sum_{l=0}^{k} q^{l(l-1)} \prod_{i=l+1}^{k} (q^n - q^i),$$

where we have used Theorem 5.2 for the second equality. We set $q^n = y$ and use Lemma 5.3 to see that the last expression is equivalent to

$$\Delta_q(k, n) = y^k - \sum_{l=0}^{k} (q^{l^2} - q^{l(l-1)}) \prod_{i=l+1}^{k} (y - q^i)$$

$$= y^k - \sum_{l=1}^{k} q^{l(l-1)}(q^k - 1) \prod_{i=l+1}^{k} (y - q^i)$$

$$= y^k - \sum_{m=0}^{k-1} q^{m(m+1)}(q^k - 1) \prod_{i=m+2}^{k} (y - q^i)$$

as desired.
\[
= y^k - (q^k - 1) \sum_{m=0}^{k-1} q^{m(m+1)} \binom{k - 1}{m} q^{k-m-1} \prod_{i=m+2}^{k} (y/q - q^{-i}) \\
= y^k - (q^k - 1)q^{k-1} \sum_{m=0}^{k-1} q^{m^2} \binom{k - 1}{m} \prod_{j=m+1}^{k} (y/q - q^{-j}) \\
= y^k - (q^k - 1)q^{k-1}(y/q)^{k-1} \\
= y^k - (q^k - 1)y^{k-1},
\]

where the penultimate equality again follows from Lemma 5.3. We now resubstitute \( y \) by \( q^n \) to obtain the result. \( \Box \)

ACKNOWLEDGEMENTS

This work was partly done while the author was a postdoctoral fellow at the Harish-Chandra Research Institute, Allahabad and partly at the Indraprastha Institute of Information Technology, Delhi.

REFERENCES

[1] Glória Cravo. Matrix completion problems. *Linear Algebra Appl.*, 430(8-9):2511–2540, 2009.

[2] Josep Ferrer and Ferran Puerta. Similarity of non-everywhere defined linear maps. *Linear Algebra Appl.*, 168:27–55, 1992.

[3] N. J. Fine and I. N. Herstein. The probability that a matrix be nilpotent. *Illinois J. Math.*, 2:499–504, 1958.

[4] G. David Forney, Jr. Convolutional codes. I. Algebraic structure. *IEEE Trans. Information Theory*, IT-16:720–738, 1970.

[5] Murray Gerstenhaber. On the number of nilpotent matrices with coefficients in a finite field. *Illinois J. Math.*, 5:330–333, 1961.

[6] Israel Gohberg, Marinus A. Kaashoek, and Frederik van Schagen. *Partially specified matrices and operators: classification, completion, applications*, volume 79 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 1995.

[7] Michael A. Harrison. *Lectures on linear sequential machines*. Academic Press, New York-London, 1969.

[8] Uwe Helmke, Jens Jordan, and Julia Lieb. Reachability of random linear systems over finite fields. In Raquel Pinto, Paula Rocha Malonek, and Paolo Vettori, editors, *Coding Theory and Applications*, volume 3 of *CIM Series in Mathematical Sciences*, pages 217–225. Springer International Publishing, 2015.

[9] Victor Kac and Pokman Cheung. *Quantum calculus*. Universitext. Springer-Verlag, New York, 2002.

[10] M. Kocięcki and K. M. Przyłuski. On the number of controllable linear systems over a finite field. *Linear Algebra Appl.*, 122/123/124:115–122, 1989.

[11] Joseph P. S. Kung. The cycle structure of a linear transformation over a finite field. *Linear Algebra Appl.*, 36:141–155, 1981.

[12] Samrith Ram. Counting zero kernel pairs over a finite field. *Linear Algebra Appl.*, 495:1–10, 2016.
[13] Irving Reiner. On the number of matrices with given characteristic polynomial. Illinois J. Math., 5:324–329, 1961.
[14] J. E. Roseblade. Obituary: Philip Hall. Bull. London Math. Soc., 16(6):603–626, 1984. With contributions by J. G. Thompson and J. A. Green.
[15] Eduardo D. Sontag. Mathematical control theory, volume 6 of Texts in Applied Mathematics. Springer-Verlag, New York, second edition, 1998. Deterministic finite-dimensional systems.
[16] Richard P. Stanley. Enumerative combinatorics. Volume 1, volume 49 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, second edition, 2012.
[17] Richard Stong. Some asymptotic results on finite vector spaces. Adv. in Appl. Math., 9(2):167–199, 1988.

Indraprastha Institute of Information Technology Delhi (IIIT-Delhi), New Delhi, India.
E-mail address: samrith@gmail.com