THE INVARIANT SUBSPACE LATTICE OF A LINEAR TRANSFORMATION

L. BRICKMAN AND P. A. FILLMORE

1. Introduction. The purpose of this paper is to study the lattice of invariant subspaces of a linear transformation on a finite-dimensional vector space over an arbitrary field. Among the topics discussed are structure theorems for such lattices, implications between linear-algebraic properties and lattice-theoretic properties, nilpotent transformations, and the conditions for the isomorphism of two such lattices. These topics correspond roughly to §§2, 3, 4, and 5 respectively.

Before summarizing our results, we shall introduce some notation and recall some pertinent notions and properties. Let A be a linear transformation on a finite-dimensional vector space V over a field F. We denote by $L_F(A)$, or simply $L(A)$, the set of all subspaces M of V such that $AM \subset M$. (The symbol "$\subset$" allows the possibility of equality.) Such subspaces are called *invariant*. We denote by m_A the minimum polynomial of A. If $m_A = p^n$ for some irreducible polynomial p and some positive integer n, we shall say that A is *primary* or, if necessary, *p-primary*. If $m_A = \prod_i p_i^{n_i}$ is the prime factorization of m_A, if $V_i = \ker p_i(A)^{n_i}$, and if $A_i = A|V_i$, then we shall call the A_i the *primary summands* of A. (We recall that the V_i are invariant, $V = \sum_i \oplus V_i$, $A = \sum_i \oplus A_i$, and $m_{A_i} = p_i^{n_i}$.) If $x_1, x_2, \ldots \in V$, $\langle x_1, x_2, \ldots \rangle$ will denote the subspace of V spanned by $x_1, x_2, \ldots$. If M is a subspace of V, and if $M = \langle x, Ax, A^2x, \ldots \rangle$ for some $x \in V$, then M is called a *cyclic subspace* and x a *cyclic vector* for M. If V is a cyclic subspace, we say that A is a *cyclic transformation*. We recall that A is cyclic if and only if A_i is cyclic for all i if and only if $A|\ker p_i(A)$ is cyclic for all i. A *semi-linear transformation* from V over F to V' over F' is a pair (T, σ) such that $T\colon V \to V'$, σ is an isomorphism of F onto F', and $T(\alpha x + \beta y) = \alpha^\sigma Tx + \beta^\sigma Ty$ for all $\alpha, \beta \in F$ and $x, y \in V$.

A *lattice* is a partially ordered system in which each pair of elements M, N has a meet (greatest lower bound), denoted $M \cap N$, and a join (least upper bound), denoted $M + N$. Clearly $L(A)$ is a lattice with inclusion as order, with intersection as meet, and with linear sum as join. If $M \subsetneq N$, we shall say that N *covers* M if there is no lattice element strictly between M and N. All lattices considered in the paper will have a *zero element* $\{0\}$ and a *unit element* V such that $\{0\} \subset M \subset V$ for all lattice elements M. Such a lattice is *complemented* if for any element M there exists at least one element N with $M \cap N = \{0\}$ and $M + N = V$. If M and N are any lattice elements, we denote by

Received November 1, 1965.

$[M, N]$ the set of lattice elements P with $M \subset P \subset N$. If each interval sublattice $[M, N]$ is complemented, then the lattice is said to be *relatively complemented.* A lattice is distributive if $(M + N) \cap P = (M \cap P) + (N \cap P)$ for all elements M, N, P, and *modular* if this identity holds whenever $M \subset P$. It is well known that the lattice of all subspaces of V is modular, and therefore so is its sublattice $L(A)$. A *Boolean algebra* is a distributive and complemented lattice. A lattice L is said to be the *direct sum* of sublattices L_1 and L_2 (notation: $L = L_1 \oplus L_2$) if each $M \in L$ is uniquely representable in the form $M = M_1 + M_2$ with $M_1 \in L_1$ and $M_2 \in L_2$ (notation: $M = M_1 \oplus M_2$) in such a way that the lattice operations can be performed "coordinate-wise." It follows that if V is the unit element of L, and $V = V_1 \oplus V_2$, then V_1 and V_2 are complementary, and $L_i = \{M \in L |\ M \subset V_i\}$ $(i = 1, 2)$. A lattice that cannot be written as a direct sum (except trivially) will be called *irreducible.* A *lattice homomorphism* is a mapping between lattices which preserves meets and joins. (Such a mapping is necessarily order-preserving.) Two lattices are *isomorphic* (*anti-isomorphic*) if there exists a one-to-one correspondence between them which preserves (reverses) order. A lattice is *self-dual* if it is anti-isomorphic to itself. Finally, a lattice is called *simple* if it admits only trivial homomorphisms (isomorphisms and constant maps). We note that a simple lattice is necessarily irreducible.

Our main results may be outlined as follows. In §2, $L(A)$ is investigated for the general linear transformation A. We find at once that $L(A) = \sum_i \oplus L(A_i)$ (the A_i being the primary summands of A) and that the $L(A_i)$ are irreducible. Further study proves that each $L(A_i)$ is either simple or a finite chain. Finally, it is observed that $L(A)$ is always self-dual. Section 3 contains the following information: $L(A)$ is distributive if and only if A is cyclic; $L(A)$ is a Boolean algebra if and only if A is cyclic and m_A is a product of distinct primes; $L(A)$ is a chain if and only if A is cyclic and primary; $L(A)$ is simple but not $\{\{0\}, V\}$ if and only if A is non-cyclic and primary. In §4 we obtain a formula for the lattice of an arbitrary nilpotent transformation. The use of the formula is illustrated by two examples, and the resulting lattices are sketched. Thus, the following question is of interest: Given a p-primary transformation A, does there exist a nilpotent transformation with the same lattice? We find that this is so if p is separable. (Here we permit an enlargement of the scalar field F. More specifically, we adjoin to F a root of p, make V a vector space over the resulting field K, and find a K-linear nilpotent transformation Q such that $L_K(Q) = L_F(A)$.) If p is not separable, the answer to the above question is probably "no". In the final section we present some necessary and sufficient conditions in order that two primary transformations have isomorphic lattices, and a lattice inclusion theorem for two commuting transformations.

2. General structure theorems. The main object of this section is to analyse $L(A)$ as far as possible assuming nothing about A beyond linearity. Consequences of further assumptions are considered in §3.

LEMMA 1. *Let V_1 and V_2 be non-trivial finite-dimensional vector spaces over the field F, and let A_1 and A_2 be linear transformations on V_1 and V_2 respectively. Then*

$$L(A_1 \oplus A_2) = L(A_1) \oplus L(A_2) \Leftrightarrow (m_{A_1}, m_{A_2}) = 1.$$

Proof. In any case we have the inclusion $L(A_1) \oplus L(A_2) \subset L(A_1 \oplus A_2)$. Suppose that $(m_{A_1}, m_{A_2}) = 1$. Let $N \in L(A_1 \oplus A_2)$, and let N_1 and N_2 be the projections of N on V_1 (along V_2) and on V_2 (along V_1) respectively. Clearly $N_1 \in L(A_1)$, $N_2 \in L(A_2)$, and $N \subset N_1 \oplus N_2$. To prove that $N \supset N_1 \oplus N_2$, let r_1 and r_2 be polynomials (coefficients in F) such that $r_1 m_{A_1} + r_2 m_{A_2} = 1$, and let $q_1 = r_1 m_{A_1}$. Then $N \supset q_1(A_1 \oplus A_2)N = (0 \oplus q_1(A_2))N = N_2$. Similarly $N \supset N_1$, so $N = N_1 \oplus N_2 \in L(A_1) \oplus L(A_2)$; cf. (**4**, p. 213).

Conversely, suppose m_{A_1} and m_{A_2} have a common prime factor q. Then for $i = 1, 2$, there exist non-zero vectors $x_i \in V_i$ such that $q(A_i)x_i = 0$. Let

$$M = \{r(A_1)x_1 + r(A_2)x_2 | \deg r < \deg q\}.$$

Then $M \in L(A_1 \oplus A_2)$. If $M = M_1 \oplus M_2$ with $M_i \in L(A_i)$, we should have $M_1 = M \cap V_1$ and $M_2 = M \cap V_2$. But $r(A_1)x_1 + r(A_2)x_2 \in M \cap V_1$ implies that $r(A_2)x_2 = 0$ and therefore $r = 0$. Thus $M \cap V_1 = \{0\}$, and similarly $M \cap V_2 = \{0\}$. Hence $M = \{0\}$, a contradiction.

THEOREM 1. *Let A be a linear transformation on V with primary summands A_i. Then*

$$L(A) = \textstyle\sum_i \oplus L(A_i)$$

and each direct summand $L(A_i)$ is irreducible.

Proof. The asserted equation follows from Lemma 1 by induction. To prove the last statement suppose that A is primary and that $L(A)$ is not irreducible. Then, as explained in §1, there exist non-trivial complementary subspaces $V_1, V_2 \in L(A)$ such that $L(A) = L(A|V_1) \oplus L(A|V_2)$. But the minimum polynomials of $A|V_1$ and $A|V_2$ are divisors of m_A. Therefore they are not relatively prime, and this contradicts Lemma 1.

Remark. The last statement of this theorem will be superseded by the deeper Theorem 2, which asserts that each $L(A_i)$ is either simple or a chain.

COROLLARY. *If F is algebraically closed, then each irreducible summand of $L(A)$ is of the form $L(Q)$ for a suitable nilpotent transformation Q.*

Proof. The hypothesis implies that m_A has the form $\prod_i (t - \lambda_i)^{n_i}$. Hence A_i satisfies $(A_i - \lambda_i)^{n_i} = 0$. Clearly $L(A_i) = L(A_i - \lambda_i)$. Thus $Q = A_i - \lambda_i$ is the required nilpotent transformation.

LEMMA 2. *$L(A)$ is a chain if and only if A is cyclic and primary.*

Proof. Assume A is cyclic and $m_A = p^n$. We shall show that

$$L(A) = \{\ker p(A)^k | k = 0, 1, \ldots, n\}.$$

https://doi.org/10.4153/CJM-1967-075-4 Published online by Cambridge University Press

Indeed, if $\{0\} \neq M \in L(A)$, then $A|M$ has minimum polynomial p^k for some $k \geqslant 1$. Hence $M \subset \ker p(A)^k$. Now, the restriction of a cyclic transformation to any invariant subspace is again cyclic (**5**, p. 129). Thus both $A|M$ and $A|\ker p(A)^k$ are cyclic with minimum polynomial p^k. Therefore

$$\dim M = \deg p^k = \dim \ker p(A)^k,$$

and so $M = \ker p(A)^k$.

Conversely, if m_A contains distinct irreducible factors, then by Theorem 1, $L(A)$ is not irreducible and is therefore not a chain. Again, if A is not cyclic, then it is a direct sum of two or more cyclic transformations. Consequently there exist non-trivial disjoint invariant subspaces, and so $L(A)$ is not a chain.

COROLLARY. *$L(A) = \{\{0\}, V\}$ if and only if A is cyclic and m_A is irreducible.*

The next lemma will be used repeatedly throughout the paper.

LEMMA 3. *Let m_A be irreducible, and let K be the algebra of polynomials in A with coefficients in F. Then*

(a) *K is a field isomorphic to that obtained by adjoining a root of m_A to F,*

(b) *V is naturally a vector space over K of K-dimension equal to the number of summands in a representation of A as a direct sum of cyclic transformations,*

(c) *A is K-linear, and*

(d) *$L_F(A)$ is the lattice of all K-linear subspaces of V.*

Proof. We shall prove only the dimensionality assertion of (b). Let $V = \sum_i \oplus V_i$, where $A|V_i$ is cyclic. If x_i is a cyclic vector for $A|V_i$, then $V_i = \{f(A)x_i | \deg f < \deg m_A\}$. From this it is clear that the K-dimension of V_i is 1, and (b) follows.

LEMMA 4. *Let A be p-primary, and let $d = \deg p$. If $M, N \in L(A)$ and N covers M, then*

(a) $p(A)N \subset M$,

(b) $\dim N = d + \dim M$.

Consequently

(c) *$d|\dim M$ for every $M \in L(A)$.*

Proof. Let A' be the quotient transformation induced by A on V/M. Then $N' = N/M$ is a minimal non-zero element of $L(A')$. But $p(A')N' \subset N'$ and $p(A')N' \in L(A')$. Since $p(A')$ is nilpotent, $p(A')N' \neq N'$. Hence $p(A')$ annihilates N', and this is equivalent to (a). To prove (b) we apply Lemma 3 to the transformation $A'|\ker p(A')$. Since N' is a minimal lattice element for this transformation, we can conclude from (d) that N' has K-dimension 1. Since $[K: F] = d$, N' has F-dimension d, and (b) follows. Finally, (c) follows from (b) by construction of a maximal chain in $L(A)$ extending from $\{0\}$ to M.

LEMMA 5. *Let A be p-primary, and let $M \in L(A)$. Then*
(a) $p(A)^{-1}M \in L(A)$,
(b) *the interval* $[M, p(A)^{-1}M]$ *in* $L(A)$ *is a simple sublattice, and*
(c) $M \subset \operatorname{rng} p(A)$ *implies* $\dim p(A)^{-1}M - \dim M = \dim \ker p(A)$.

Proof. Part (a) is immediate. Let A' be the quotient transformation induced by A on $p(A)^{-1}M/M$. Then $p(A') = 0$. Consequently Lemma 3 implies that $L(A')$ is the lattice of all subspaces of the vector space $p(A)^{-1}M/M$ over the field of polynomials in A'. Therefore $L(A')$ is simple (**3**, p. 121). But $L(A')$ and $[M, p(A)^{-1}M]$ are isomorphic, and so (b) is proved. For (c) we apply the equation

$$\dim \ker B + \dim \operatorname{rng} B = \dim \operatorname{dom} B$$

with $B = p(A)|p(A)^{-1}M$. Since $M \subset \operatorname{rng} p(A)$, we have $\operatorname{rng} B = M$. Moreover, since $\ker p(A) \subset p(A)^{-1}M$, we also have $\ker B = \ker p(A)$.

LEMMA 6. *Let A be p-primary and non-cyclic. If $M, N \in L(A|\operatorname{rng} p(A))$ and if N covers M, then $N \subsetneqq p(A)^{-1}M$.*

Proof. The weak inclusion $N \subset p(A)^{-1}M$ was established in Lemma 4. Now $\dim N - \dim M = d$ by the same lemma, and since $M \subset N \subset p(A)^{-1}M$, the desired conclusion will follow if $\dim p(A)^{-1}M - \dim M > d$. By Lemma 5(c), this is equivalent to $\dim \ker p(A) > d$. But $A|\ker p(A)$ has the same number of cyclic summands as A. Therefore Lemma 3 implies that $\dim \ker p(A) \geqslant 2d$.

THEOREM 2. *Let A be a linear transformation on V with primary summands A_i. Then*

$$L(A) = \textstyle\sum_i \oplus L(A_i),$$

and each direct summand $L(A_i)$ is simple or a chain according as A_i is non-cyclic or cyclic.

Proof. By Lemma 2 and Theorem 1 we need only prove that if a linear transformation A is primary and non-cyclic, then $L(A)$ is simple. For this we suppose given a homomorphism h of $L(A)$ which identifies two distinct elements. It follows easily that h is constant on the entire interval determined by the meet and join of these two elements. Hence there exist $N_1, N_2 \in L(A)$ with $h(N_1) = h(N_2)$, and which have the property that N_2 covers N_1. If $M = p(A)N_2$, it follows from Lemma 4 that $N_1, N_2 \in [M, p(A)^{-1}M]$. By Lemma 5, such an interval of $L(A)$ is a simple sublattice, and so h identifies all its elements. Now we select a maximal chain

$$\{0\} = M_0 \subsetneqq M_1 \subsetneqq \ldots \subsetneqq M_k = \operatorname{rng} p(A)$$

of invariant subspaces which includes M and which extends from $\{0\}$ to $\operatorname{rng} p(A)$. Let $I_i = [M_i, p(A)^{-1}M_i]$ $(i = 0, 1, \ldots, k)$. Then

$$I_{i-1} \cap I_i = [M_i, p(A)^{-1}M_{i-1}] \quad (i = 1, \ldots, k).$$

Since M_i covers M_{i-1}, we can use Lemma 6 to conclude that

$$M_i \subsetneq p(A)^{-1}M_{i-1}.$$

Thus adjacent members of the sequence $I_0, I_1, \ldots, I_k$ intersect in at least two elements. This together with the fact that h is constant on one of the I_i implies that h is constant on $\bigcup_i I_i$. Since $M_0 = \{0\}$ and $p(A)^{-1}M_k = V$, we can conclude that h is constant on $L(A)$.

THEOREM 3. *If A is any linear transformation on V, then $L(A)$ is self-dual.*

Proof. Let V^* and A^* be the duals of V and A respectively. Then A is similar to A^* (**5**, p. 98), and therefore $L(A)$ and $L(A^*)$ are isomorphic. But if $M \in L(A)$ and M^0 is the annihilator of M in V^*, the mapping $M \to M^0$ is evidently an anti-isomorphism of $L(A)$ onto $L(A^*)$.

3. Special structure theorems. The results in this section relate properties of a linear transformation A to properties of its invariant subspace lattice $L(A)$. For illustration and convenience, we begin by collecting the results of this nature already found in §2.

1. $L(A)$ is irreducible if and only if A is primary.
2. $L(A)$ is a chain if and only if A is cyclic and primary.
3. $L(A)$ is trivial if and only if A is cyclic and m_A is irreducible.
4. $L(A)$ is simple but non-trivial if and only if A is non-cyclic and primary.

THEOREM 4. *The following statements are equivalent:*

(a) *A is cyclic,*

(b) *$L(A)$ is a (finite) direct sum of chains, and*

(c) *$L(A)$ is distributive.*

Each of these conditions implies that

(d) *$L(A)$ is finite,*

and if F is infinite all the conditions (a)–(d) *are equivalent* (cf., (**5**, p. 129, Ex. 3) *for the equivalence of* (a) *and* (d)).

Proof. If A is cyclic, then so are its primary summands A_i. Hence the irreducible summands $L(A_i)$ of $L(A)$ are chains by Lemma 2. Thus we obtain (b) by Theorem 1.

Since a chain is finite and distributive, so is a (finite) product of chains, and therefore (b) implies (c) and (d).

To prove that (c) implies (a) let us suppose that A is not cyclic. It follows that at least one of the A_i, say A_1, is not cyclic. If A_1 decomposes into $m > 1$ cyclic summands, the same is true of $B = A_1|\ker p_1(A_1)$. By Lemma 3, $\ker p_1(A_1)$ has K-dimension m, where K is the field of polynomials in B. It follows from (d) of Lemma 3 that $L(B)$ is not distributive. Since $L(B)$ is a sublattice of $L(A)$, the latter is not distributive either. Hence (c) implies (a).

Finally, let us assume that F is infinite and that A is not cyclic. Then we may again suppose that A_1 is not cyclic. It follows that K, $L(B)$, and $L(A)$ are all infinite. Thus (d) implies (a).

COROLLARY. *The cyclic invariant subspaces of A are precisely the elements* $M \in L(A)$ *such that* $[\{0\}, M]$ *is a distributive sublattice.*

COROLLARY. *If F is algebraically closed, each of the conditions* (a)–(d) *is equivalent to the statement:*

(e) *All the eigenspaces of A are one-dimensional.*

Proof. Since F is algebraically closed, F is infinite, and so (a)–(d) are equivalent. We complete the proof by observing that (e) is equivalent to the statement that all eigenspaces are cyclic, that this is equivalent to all generalized eigenspaces being cyclic, and this is equivalent to (a).

The next theorem is well known (**4**, p. 214; **5**, p. 129); we include it for completeness and for the possible interest of our proof.

THEOREM 5. *$L(A)$ is complemented if and only if m_A is a product of distinct irreducible polynomials.*

Proof. We observe that $L(A)$ is complemented if and only if the direct summands $L(A_i)$ are complemented. Now if $m_A = \prod_i p_i$, then $p_i(A_i) = 0$, and therefore $L(A_i)$ is the lattice of all subspaces of a certain vector space (Lemma 3(d)). Hence $L(A_i)$ is complemented.

Conversely, let us suppose that $L(A)$ is complemented and (if possible) that $m_{A_i} = p_i^{n_i}$ with $n_i \geqslant 2$ for some i. Then A_i has a direct summand A'_i which is cyclic and which has the same minimum polynomial. By (the proof of) Lemma 2, $L(A'_i)$ is a chain of $n_i + 1$ elements. Since $n_i + 1 \geqslant 3$, $L(A'_i)$ is not complemented. Since $L(A'_i)$ is an interval in $L(A_i)$, $L(A_i)$ is not relatively complemented. But a complemented modular lattice is necessarily relatively complemented (**3**, Theorem 1, p. 114). Having reached this contradiction, we may conclude that $n_i = 1$ for all i.

Remark. The polynomial m_A is a product of distinct irreducible polynomials if and only if the algebra of polynomials in A is semi-simple. Such transformations are called semi-simple.

COROLLARY. *If F is algebraically closed, $L(A)$ is complemented if, and only if, A can be reduced to diagonal form.*

COROLLARY. *$L(A)$ is a Boolean algebra if and only if A is cyclic and m_A is a product of distinct irreducible polynomials. $L(A)$ is then a (finite) direct sum of two-element chains.*

4. Nilpotent transformations. In the proof of Lemma 5(b) we showed that if A is p-primary, and $M \in L(A)$, then the interval $[M, p(A)^{-1}M]$ in $L(A)$ is isomorphic to the lattice of all subspaces of a certain vector space over the field K obtained by adjoining to F a root of p. Moreover, it is easy to see that $L(A) = \cup [M, p(A)^{-1}M]$, where the union is over all

$$M \in L(A)|\text{rng}\, p(A)).$$

Thus in a vague sense all the invariant subspaces of A are K-linear. Under the additional assumption that p is separable we shall now make this precise by causing K to act on V in such a way that V becomes a K-vector space, and A and all its invariant subspaces are indeed K-linear. Moreover, we shall obtain (Theorem 6) a K-linear nilpotent transformation Q such that $L_F(A) = L_K(Q)$. The value of reducing considerations to the lattice of a nilpotent transformation will be apparent from Theorem 7 and the subsequent remarks.

THEOREM 6. *Let A be p-primary with p separable. Let $A = S + Q$ be the decomposition of A into its semi-simple and nilpotent parts* (**4**, p. 217, Th. 8). *Let K be the algebra of polynomials in S over F. Then K is a field, V is naturally a K-vector space, A is K-linear, and $L_F(A) = L_K(A) = L_K(Q)$.*

Proof. We first remark that although Theorem 8 of (**4**, p. 217) is stated with the assumption that F is a subfield of the complex numbers, the proof is valid in the more general situation that the irreducible factors of the minimum polynomial are all separable. Taylor's formula for polynomials is not valid over fields of finite characteristic, but the proof in (**4**) requires only an equation of the form $f(a + b) \equiv f(a) + f'(a)b \pmod{b^2}$ and this is always valid. From (**4**, p. 218) we obtain the equation $p(S) = 0$. Hence our first two assertions are immediate (as in Lemma 3). Since Q and S are polynomials in A (**4**, Th. 8), $AS = SA$, and therefore A is K-linear. If $M \in L_F(A)$, then clearly $M \in L_F(S)$. Thus M is K-linear, and so $L_F(A) = L_K(A)$. Finally, $M \in L_K(A)$ is clearly equivalent to $M \in L_K(A - S)$. Hence $L_K(A) = L_K(Q)$.

Remarks. If A is regarded as a K-linear transformation, the equation $A = Q + S$ exhibits A as the sum of a nilpotent transformation and a scalar. Therefore there is a basis $\{e_1, \ldots, e_m\}$ for V over K with respect to which the matrix of A is in Jordan canonical form. If degree $p = d$, then

$$\{e_1, Se_1, \ldots, S^{d-1}e_1; \ldots ; e_m, Se_m, \ldots, S^{d-1}e_m\}$$

is a basis for V over F. With respect to this basis, A, now regarded as an F-linear transformation, has for its matrix a direct sum of matrices of the form

$$\begin{bmatrix} C & & & & & \\ I & C & & & & \\ & I & \cdot & & & \\ & & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & C \\ & & & & \cdot & I & C \end{bmatrix},$$

where C is the companion matrix of p and I is the identity matrix of order d. Such matrices are more convenient for computation than those occurring in the classical canonical form (**5**) because the diagonal and off-diagonal parts commute.

THEOREM 7. *If Q is nilpotent on V, then*

$$L(Q) = \bigcup_{M \in L(Q|QV)} [M, Q^{-1}M],$$

where $[M, Q^{-1}M]$ is an interval in the lattice of all subspaces of V. Each interval satisfies the equation

$$\dim Q^{-1}M - \dim M = \dim \ker Q.$$

Proof. If $M \in L(Q)$ and if N is any subspace of V with $M \subset N \subset Q^{-1}M$, then $QN \subset QQ^{-1}M \subset M \subset N$. Hence $N \in L(Q)$. Conversely if $N \in L(Q)$, then $QN \in L(Q|QV)$ and $QN \subset N \subset Q^{-1}QN$. Thus our formula is established. The final statement is a special case of that of Lemma 5(c).

Remarks. Since the intervals $[M, Q^{-1}M]$ are taken in the subspace lattice of V, the above formula does not contain the "circularity" present in the more general formula $L(A) = \bigcup [M, p(A)^{-1}M]$. It is true that $L(A)$ is given in terms of $L(Q|QV)$ but this is clearly a reduction in complexity, for $\dim QV < \dim V$, and also the index of nilpotence of $Q|QV$ is less (by 1) than that of Q. (Indeed, one can easily "iterate" the above formula and obtain a multiple-union formula for $L(Q)$ which involves nothing but Q and the lattice of all subspaces of V.)

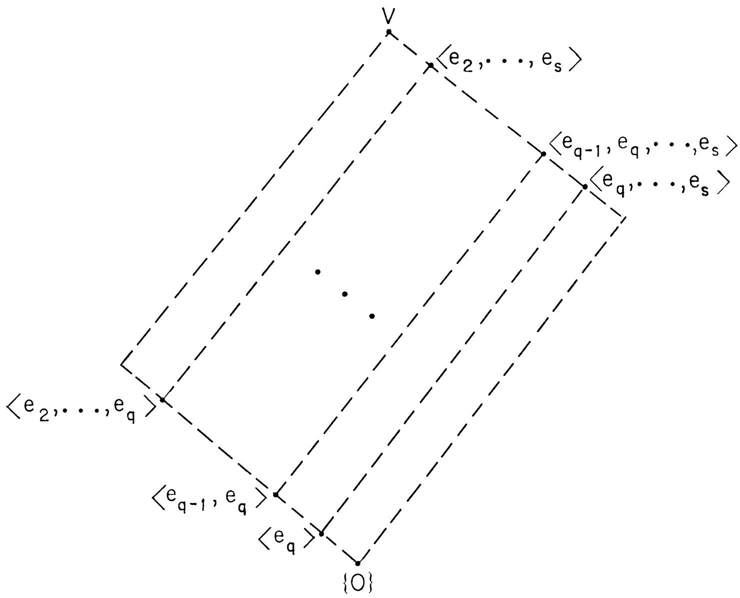

FIGURE 1

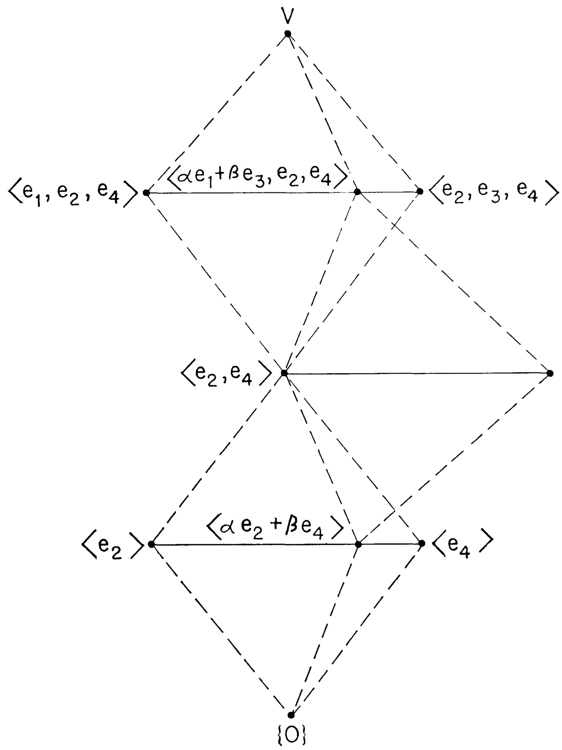

FIGURE 2

We conclude §4 by presenting two examples in each of which the lattice of a nilpotent transformation is computed using the formula of Theorem 7.

1. Let $V = \langle e_1, \ldots, e_s\rangle$, $Q: e_1 \rightarrow e_2 \rightarrow \ldots \rightarrow e_q \rightarrow 0$, $e_{q+1} \rightarrow 0, \ldots, e_s \rightarrow 0$. Then $QV = \langle e_2, \ldots, e_q\rangle$, $Q|QV: e_2 \rightarrow e_3 \rightarrow \ldots \rightarrow e_q \rightarrow 0$, and

$$L(Q|QV) = \{0\}, \langle e_q\rangle, \langle e_{q-1}, e_q\rangle, \ldots, \langle e_2, \ldots, e_q\rangle\}.$$

(See Lemma 2.) Hence Theorem 7 gives

$$L(Q) = [\{0\}, \langle e_q, \ldots, e_s\rangle] \cup [\langle e_q\rangle, \langle e_{q-1}, e_q, \ldots, e_s\rangle]$$
$$\cup[\langle e_{q-1}, e_q\rangle, \langle e_{q-2}, e_{q-1}, \ldots, e_s\rangle] \cup \ldots \cup [\langle e_2, \ldots, e_q\rangle, V].$$

2. Let $V = \langle e_1, e_2, e_3, e_4\rangle$ and let Q be defined on V by $Qe_1 = e_2$, $Qe_2 = 0$, $Qe_3 = e_4$, $Qe_4 = 0$. Thus

$$\text{matrix } Q = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Then by Theorem 7

$$L(Q) = \bigcup_{M \subset \langle e_2, e_4 \rangle} [M, Q^{-1}M].$$

Hence

$$L(Q) = [\{0\}, \langle e_2, e_4\rangle] \cup [\langle e_2, e_4\rangle, V] \cup (\cup_{\alpha,\beta} [\langle \alpha e_2 + \beta e_4\rangle, \langle \alpha e_1 + \beta e_3, e_2, e_4\rangle]).$$

This lattice is pictured above, dotted lines indicating some of the inclusion relations. There is a one-parameter family of intervals extending from dimension 1 to dimension 3, only one of which is drawn.

5. Isomorphism theorems. We now present our results concerning isomorphism and equality of invariant subspace lattices.

THEOREM 8. *For $i = 1, 2$ let A_i be a p_i-primary linear transformation on the vector space V_i over the field F_i. If there is a non-singular semi-linear transformation (T, σ) of V_1 over F_1 onto V_2 over F_2 such that $TA_1 = A_2 T$, then $p_1^\sigma = p_2$ and T induces an isomorphism of $L(A_1)$ onto $L(A_2)$. Conversely, if $L(A_1) \cong L(A_2)$ and if σ is any isomorphism of F_1 onto F_2 with $p_1^\sigma = p_2$, then there exists a non-singular semi-linear transformation (T, σ) such that $TA_1 = A_2 T$.*

Proof. The proof of the first statement is a routine computation. We suppose then, that $M \to M'$ is an isomorphism of $L(A_1)$ onto $L(A_2)$, and that σ is an isomorphism of F_1 onto F_2 with $p_1^\sigma = p_2$. We can select $W_1, \ldots, W_k \in L(A_1)$ such that $V_1 = W_1 \oplus \ldots \oplus W_k$, and the restrictions $A_1|W_i$ are cyclic (and primary). By Lemma 2 the interval $[\{0\}, W_i]$ is a chain, and hence so is $[\{0\}, W'_i]$. Therefore $A_2|W'_i$ is cyclic. Let x_i and x'_i be cyclic vectors for W_i and W'_i respectively. Then each element of W_i is of the form $f(A_1)x_i$ for a unique polynomial f of degree less than that of the minimum polynomial of $A_1|W_i$. We define T on W_i by $Tf(A_1)x_i = f^\sigma(A_2)x'_i$, and extend T to V_1 "by additivity." It follows easily that (T, σ) is as required.

COROLLARY. *Let A_1 and A_2 be p-primary linear transformations on the vector space V. Then $L(A_1) \cong L(A_2)$ if and only if A_1 and A_2 are similar.*

THEOREM 9. *For $i = 1, 2$ let p_i be an irreducible and separable polynomial over the field F_i, and let A_i be a linear transformation on the vector space V_i over F_i which is a direct sum of at least three cyclic p_i-primary transformations. Let $A_i = S_i + Q_i$ be the decomposition of A_i into its semi-simple and nilpotent parts, and let K_i be the field of polynomials in S_i over F_i* (See Theorem 6). *Then $L(A_1) \cong L(A_2)$ if and only if there exists a non-singular semi-linear transformation (T, σ) of V_1 over K_1 onto V_2 over K_2 such that $TQ_1 = Q_2 T$.*

Proof. Suppose that the semi-linear transformation (T, σ) satisfies $TQ_1 = Q_2 T$. Then $L_{K_1}(Q_1) \cong L_{K_2}(Q_2)$. But $L(A_i) = L_{K_i}(Q_i)$ by Theorem 6, so that $L(A_1) \cong L(A_2)$.

Suppose conversely that $L(A_1) \cong L(A_2)$. We show that this implies that $K_1 \cong K_2$. The minimum polynomial of $A_1|\ker p_1(A_1)$ is p_1, and so by Lemma 3 the interval $[\{0\}, \ker p_1(A_1)\}]$ in $L(A_1)$ is the lattice of all subspaces of a vector space over K_1. The dimension of this space is the number of summands in a decomposition of A_1 as a direct sum of cyclic transformations, and thus is at least 3. Since this interval is complemented, so is its image $[\{0\}, N]$ in $L(A_2)$. Theorem 5 now implies that p_2 is the minimum polynomial of $A_2|N$, and Lemma 3 implies that $[\{0\}, N]$ is the lattice of all subspaces of a vector space over K_2. By one of the fundamental theorems of projective geometry **(2**, p. 51**)** the isomorphism of $[\{0\}, \ker p_1(A_1)]$ and $[\{0\}, N]$ is induced by a semi-linear transformation. In particular $K_1 \cong K_2$. Now since $L(A_1) \cong L(A_2)$, Theorem 6 implies that $L_{K_1}(Q_1) \cong L_{K_2}(Q_2)$. Hence the existence of the required semi-linear transformation follows from Theorem 8.

Remark. By way of generalizing the result from projective geometry used in the above proof, it would be interesting to determine when an isomorphism $L(A_1) \cong L(A_2)$ is induced by a semi-linear transformation on the underlying vector space. Baer **(1**, Th. II.3.1**)** has investigated similar questions.

THEOREM 10. *Let A and B be commuting linear transformations on V. Then $L(A) \subset L(B)$ if and only if B is a polynomial in A.*

Proof. The other implication being trivial, suppose that $L(A) \subset L(B)$. We can write $V = V_1 \oplus \ldots \oplus V_K$ such that the V_i are cyclic and invariant relative to A, and such that the minimum polynomials m_i of $A|V_i$ have the property: $m_{i+1}|m_i$ for $i = 1, \ldots, k - 1$. By assumption the V_i are B-invariant. If e_1 is a cyclic vector for V_1, there is a polynomial q_1 such that $Be_1 = q_1(A)e_1$. Any vector $x \in V_1$ is of the form $x = r(A)e_1$ for some polynomial r, and therefore $Bx = Br(A)e_1 = r(A)Be_1 = r(A)q_1(A)e_1 = q_1(A)x$, so that $B = q_1(A)$ on V_1. In like manner, if e_2 is a cyclic vector for V_2, then $Be_2 = q_2(A)e_2$ and $B = q_2(A)$ on V_2. Consider now the vector $f = e_1 + e_2$. The subspace $\langle f, Af, \ldots \rangle$ is A-invariant, hence B-invariant, and so $Bf = s(A)f$ for some polynomial s. We then have $Be_i = s(A)e_i$ for $i = 1, 2$, and therefore

$$s = q_1 + k_1 m_1 = q_2 + k_2 m_2$$

for suitable polynomials k_1 and k_2. Since $m_2|m_1$, we conclude that $q_1(A) = q_2(A)$ on V_2. Hence $B = q_1(A)$ on $V_1 \oplus V_2$. Iteration of this procedure yields $B = q_1(A)$ on all of V, and completes the proof.

Remark. The theorem is false if it is not assumed that A and B commute. However, an elaboration of the above argument shows that this hypothesis may be dropped if $m_1 = m_2$ in the notation of the above proof; cf., **(1**, Th. II.2.2.**)**.

References

1. R. Baer, *A unified theory of projective spaces and finite abelian groups*, Trans. Amer. Math. Soc. *52* (1942), 283–343.

2. ——— *Linear algebra and projective geometry* (New York, 1952).

3. G. Birkhoff, *Lattice theory*, Amer. Math. Soc. Colloquium Publications, Vol. XXV (1948).

4. K. Hoffman and R. Kunze, *Linear algebra* (Englewood Cliffs, N.J., 1961).

5. N. Jacobson, *Lectures in abstract algebra*, Vol. II (Princeton, N.J., 1953).

Indiana University,
Bloomington, Indiana