CHAPTER XIV

Representation of One Endomorphism

We deal here with one endomorphism of a module, actually a free module, and especially a finite dimensional vector space over a field k. We obtain the Jordan canonical form for a representing matrix, which has a particularly simple shape when k is algebraically closed. This leads to a discussion of eigenvalues and the characteristic polynomial. The main theorem can be viewed as giving an example for the general structure theorem of modules over a principal ring. In the present case, the principal ring is the polynomial ring $k[X]$ in one variable.

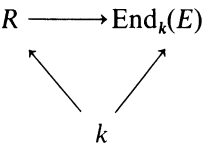
§1. REPRESENTATIONS

Let k be a commutative ring and E a module over k. As usual, we denote by $\mathrm{End}_k(E)$ the ring of k-endomorphisms of E, i.e. the ring of k-linear maps of E into itself.

Let R be a k-algebra (given by a ring-homomorphism $k \to R$ which allows us to consider R as a k-module). By a **representation** of R in E one means a k-algebra homomorphism $R \to \mathrm{End}_k(E)$, that is a ring-homomorphism

$$\rho : R \to \mathrm{End}_k(E)$$

which makes the following diagram commutative:



[As usual, we view $\mathrm{End}_k(E)$ as a k-algebra; if I denotes the identity map of E, we have the homomorphism of k into $\mathrm{End}_k(E)$ given by $a \mapsto aI$. We shall also use I to denote the unit matrix if bases have been chosen. The context will always make our meaning clear.]

We shall meet several examples of representations in the sequel, with various types of rings (both commutative and non-commutative). In this chapter, the rings will be commutative.

We observe that E may be viewed as an $\mathrm{End}_k(E)$ module. Hence E may be viewed as an R-module, defining the operation of R on E by letting

$$(x, v) \mapsto \rho(x)v$$

for $x \in R$ and $v \in E$. We usually write xv instead of $\rho(x)v$.

A subgroup F of E such that $RF \subset F$ will be said to be an **invariant** submodule of E. (It is both R-invariant and k-invariant.) We also say that it is invariant under the representation.

We say that the representation is **irreducible**, or **simple**, if $E \neq 0$, and if the only invariant submodules are 0 and E itself.

The purpose of representation theories is to determine the structure of all representations of various interesting rings, and to classify their irreducible representations. In most cases, we take k to be a field, which may or may not be algebraically closed. The difficulties in proving theorems about representations may therefore lie in the complication of the ring R, or the complication of the field k, or the complication of the module E, or all three.

A representation ρ as above is said to be **completely reducible** or **semi-simple** if E is an R-direct sum of R-submodules E_i,

$$E = E_1 \oplus \cdots \oplus E_m$$

such that each E_i is irreducible. We also say that E is completely reducible. It is not true that all representations are completely reducible, and in fact those considered in this chapter will not be in general. Certain types of completely reducible representations will be studied later.

There is a special type of representation which will occur very frequently. Let $v \in E$ and assume that $E = Rv$. We shall also write $E = (v)$. We then say that E is **principal** (over R), and that the representation is **principal**. If that is the case, the set of elements $x \in R$ such that $xv = 0$ is a left ideal $\mathfrak{a}$ of R (obvious). The map of R onto E given by

$$x \mapsto xv$$

induces an isomorphism of R-modules,

$$R/\mathfrak{a} \to E$$

(viewing R as a left module over itself, and $R/\mathfrak{a}$ as the factor module). In this map, the unit element 1 of R corresponds to the generator v of E.

As a matter of notation, if $v_1, \ldots, v_n \in E$, we let $(v_1, \ldots, v_n)$ denote the submodule of E generated by $v_1, \ldots, v_n$.

Assume that E has a decomposition into a direct sum of R-submodules

$$E = E_1 \oplus \cdots \oplus E_s.$$

Assume that each E_i is free and of dimension $\geqq 1$ over k. Let $\mathfrak{B}_1, \ldots, \mathfrak{B}_s$ be bases for $E_1, \ldots, E_s$ respectively over k. Then $\{\mathfrak{B}_1, \ldots, \mathfrak{B}_s\}$ is a basis for E. Let $\varphi \in R$, and let φ_i be the endomorphism induced by φ on E_i. Let M_i be the matrix of φ_i with respect to the basis $\mathfrak{B}_i$. Then the matrix M of φ with respect to $\{\mathfrak{B}_1, \ldots, \mathfrak{B}_s\}$ looks like

$$\begin{pmatrix} M_1 & 0 & \cdots & 0 \\ 0 & M_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & 0 \\ 0 & \cdots & 0 & M_s \end{pmatrix}.$$

A matrix of this type is said to be decomposed into **blocks**, $M_1, \ldots M_s$. When we have such a decomposition, the study of φ or its matrix is completely reduced (so to speak) to the study of the blocks.

It does not always happen that we have such a reduction, but frequently something almost as good happens. Let E' be a submodule of E, invariant under R. Assume that there exists a basis of E' over k, say $\{v_1, \ldots, v_m\}$, and that this basis can be completed to a basis of E,

$$\{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}.$$

This is always the case if k is a field.

Let $\varphi \in R$. Then the matrix of φ with respect to this basis has the form

$$\begin{pmatrix} M' & * \\ 0 & M'' \end{pmatrix}.$$

Indeed, since E' is mapped into itself by φ, it is clear that we get M' in the upper left, and a zero matrix below it. Furthermore, for each $j = m + 1, \ldots, n$ we can write

$$\varphi v = c_{j1}v_1 + \ldots + c_{jm}v_m + c_{j,m+1}v_{m+1} + \ldots + c_{jn}v_n.$$

The transpose of the matrix (c_{ji}) then becomes the matrix

$$\begin{pmatrix} * \\ M'' \end{pmatrix}$$

occurring on the right in the matrix representing φ.

Furthermore, consider an exact sequence

$$0 \to E' \to E \to E'' \to 0.$$

Let $\bar{v}_{m+1}, \ldots, \bar{v}_n$ be the images of $v_{m+1}, \ldots, v_n$ under the canonical map $E \to E''$. We can define a linear map

$$\varphi'' : E'' \to E''$$

in a natural way so that $(\overline{\varphi v}) = \varphi''(\bar{v})$ for all $v \in E$. Then it is clear that the matrix of φ'' with respect to the basis $\{\bar{v}_1, \ldots, \bar{v}_n\}$ is M''.

§2. DECOMPOSITION OVER ONE ENDOMORPHISM

Let k be a field and E a finite-dimensional vector space over k, $E \neq 0$. Let $A \in \mathrm{End}_k(E)$ be a linear map of E into itself. Let t be transcendental over k. We shall define a representation of the polynomial ring $k[t]$ in E. Namely, we have a homomorphism

$$k[t] \to k[A] \subset \mathrm{End}_k(E)$$

which is obtained by substituting A for t in polynomials. The ring $k[A]$ is the subring of $\mathrm{End}_k(E)$ generated by A, and is commutative because powers of A commute with each other. Thus if $f(t)$ is a polynomial and $v \in E$, then

$$f(t)v = f(A)v.$$

The kernel of the homomorphism $f(t) \mapsto f(A)$ is a principal ideal of $k[t]$, which is $\neq 0$ because $k[A]$ is finite dimensional over k. It is generated by a unique polynomial of degree > 0, having leading coefficient 1. This polynomial will be called the **minimal polynomial** of A over k, and will be denoted by $q_A(t)$. It is of course not necessarily irreducible.

Assume that there exists an element $v \in E$ such that $E = k[t]v = k[A]v$. This means that E is generated over k by the elements

$$v, Av, A^2v, \ldots.$$

We called such a module **principal**, and if $R = k[t]$ we may write $E = Rv = (v)$.

If $q_A(t) = t^d + a_{d-1}t^{d-1} + \cdots + a_0$ then the elements

$$v, Av, \ldots, A^{d-1}v$$

constitute a basis for E over k. This is proved in the same way as the analogous statement for finite field extensions. First we note that they are linearly inde pendent, because any relation of linear dependence over k would yield a poly-

nomial $g(t)$ of degree less than deg q_A and such that $g(A) = 0$. Second, they generate E because any polynomial $f(t)$ can be written $f(t) = g(t)q_A(t) + r(t)$ with deg $r <$ deg q_A. Hence $f(A) = r(A)$.

With respect to this basis, it is clear that the matrix of A is of the following type:

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & -a_{d-2} \\ 0 & 0 & 0 & \cdots & 1 & -a_{d-1} \end{pmatrix}.$$

If $E = (v)$ is principal, then E is isomorphic to $k[t]/(q_A(t))$ under the map $f(t) \mapsto f(A)v$. The polynomial q_A is uniquely determined by A, and does not depend on the choice of generator v for E. This is essentially obvious, because if f_1, f_2 are two polynomials with leading coefficient 1, then $k[t]/(f_1(t))$ is isomorphic to $k[t]/(f_2(t))$ if and only if $f_1 = f_2$. (Decompose each polynomial into prime powers and apply the structure theorem for modules over principal rings.)

If E is principal then we shall call the polynomial q_A above the **polynomial invariant of** E, with respect to A, or simply its **invariant**.

Theorem 2.1. *Let E be a non-zero finite-dimensional space over the field k, and let $A \in \mathrm{End}_k(E)$. Then E admits a direct sum decomposition*

$$E = E_1 \oplus \cdots \oplus E_r,$$

where each E_i is a principal $k[A]$-submodule, with invariant $q_i \neq 0$ such that

$$q_1 | q_2 | \cdots | q_r.$$

The sequence $(q_1, \ldots, q_r)$ is uniquely determined by E and A, and q_r is the minimal polynomial of A.

Proof. The first statement is simply a rephrasing in the present language for the structure theorem for modules over principal rings. Furthermore, it is clear that $q_r(A) = 0$ since $q_i | q_r$ for each i. No polynomial of lower degree than q_r can annihilate E, because in particular, such a polynomial does not annihilate E_r. Thus q_r is the minimal polynomial.

We shall call $(q_1, \ldots, q_r)$ the **invariants** of the pair (E, A). Let $E = k^{(n)}$, and let A be an $n \times n$ matrix, which we view as a linear map of E into itself. The invariants $(q_1, \ldots, q_r)$ will be called the **invariants** of A (over k).

Corollary 2.2. *Let k' be an extension field of k and let A be an $n \times n$ matrix in k. The invariants of A over k are the same as its invariants over k'.*

Proof. Let $\{v_1, \ldots, v_n\}$ be a basis of $k^{(n)}$ over k. Then we may view it also as a basis of $k'^{(n)}$ over k'. (The unit vectors are in the k-space generated by $v_1, \ldots, v_n$; hence $v_1, \ldots, v_n$ generate the n-dimensional space $k'^{(n)}$ over k'.) Let $E = k^{(n)}$. Let L_A be the linear map of E determined by A. Let L'_A be the linear map of $k'^{(n)}$ determined by A. The matrix of L_A with respect to our given basis is the same as the matrix of L'_A. We can select the basis corresponding to the decomposition

$$E = E_1 \oplus \cdots \oplus E_r$$

determined by the invariants $q_1, \ldots, q_r$. It follows that the invariants don't change when we lift the basis to one of $k'^{(n)}$.

Corollary 2.3. *Let A, B be $n \times n$ matrices over a field k and let k' be an extension field of k. Assume that there is an invertible matrix C' in k' such that $B = C'AC'^{-1}$. Then there is an invertible matrix C in k such that $B = CAC^{-1}$.*

Proof. Exercise.

The structure theorem for modules over principal rings gives us two kinds of decompositions. One is according to the invariants of the preceding theorem. The other is according to prime powers.

Let $E \neq 0$ be a finite dimensional space over the field k, and let $A : E \to E$ be in $\mathrm{End}_k(E)$. Let $q = q_A$ be its minimal polynomial. Then q has a factorization,

$$q = p_1^{e_1} \cdots p_s^{e_s} \qquad (e_i \geqq 1)$$

into prime powers (distinct). Hence E is a direct sum of submodules

$$E = E(p_1) \oplus \cdots \oplus E(p_s),$$

such that each $E(p_i)$ is annihilated by $p_i^{e_i}$. Furthermore, each such submodule can be expressed as a direct sum of submodules isomorphic to $k[t]/(p^e)$ for some irreducible polynomial p and some integer $e \geqq 1$.

Theorem 2.4. *Let $q_A(t) = (t - \alpha)^e$ for some $\alpha \in k$, $e \geqq 1$. Assume that E is isomorphic to $k[t]/(q)$. Then E has a basis over k such that the matrix of A relative to this basis is of type*

$$\begin{pmatrix} \alpha & 0 & \cdots & 0 \\ 1 & \alpha & & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & & \ddots & 0 \\ 0 & \cdots & 1 & \alpha \end{pmatrix}.$$

Proof. Since E is isomorphic to $k[t]/(q)$, there exists an element $v \in E$ such that $k[t]v = E$. This element corresponds to the unit element of $k[t]$ in the isomorphism

$$k[t]/(q) \to E.$$

We contend that the elements

$$v, (t - \alpha)v, \ldots, (t - \alpha)^{e-1}v,$$

or equivalently,

$$v, (A - \alpha)v, \ldots, (A - \alpha)^{e-1}v,$$

form a basis for E over k. They are linearly independent over k because any relation of linear dependence would yield a relation of linear dependence between

$$v, Av, \ldots, A^{e-1}v,$$

and hence would yield a polynomial $g(t)$ of degree less than $\deg q$ such that $g(A) = 0$. Since $\dim E = e$, it follows that our elements form a basis for E over k. But $(A - \alpha)^e = 0$. It is then clear from the definitions that the matrix of A with respect to this basis has the shape stated in our theorem.

Corollary 2.5. *Let k be algebraically closed, and let E be a finite-dimensional non-zero vector space over k. Let $A \in \mathrm{End}_k(E)$. Then there exists a basis of E over k such that the matrix of A with respect to this basis consists of blocks, and each block is of the type described in the theorem.*

A matrix having the form described in the preceding corollary is said to be in **Jordan canonical form**.

Remark 1. A matrix (or an endomorphism) N is said to be **nilpotent** if there exists an integer $d > 0$ such that $N^d = 0$. We see that in the decomposition of Theorem 2.4 or Corollary 2.5, the matrix M is written in the form

$$M = B + N$$

where N is nilpotent. In fact, N is a triangular matrix (i.e. it has zero coefficients on and above the diagonal), and B is a diagonal matrix, whose diagonal elements are the roots of the minimal polynomial. Such a decomposition can always be achieved whenever the field k is such that all the roots of the minimal polynomial lie in k. We observe also that the only case when the matrix N is 0 is when all the roots of the minimal polynomial have multiplicity 1. In this case, if $n = \dim E$, then the matrix M is a diagonal matrix, with n distinct elements on the diagonal.

Remark 2. The main theorem of this section can also be viewed as falling under the general pattern of decomposing a module into a direct sum as far as possible, and also giving normalized bases for vector spaces with respect to various structures, so that one can tell in a simple way the effect of an endomorphism. More formally, consider the category of pairs (E, A), consisting of a finite dimensional vector space E over a field k, and an endomorphism $A: E \to E$. By a morphism of such pairs

$$f: (E, A) \to (E', A')$$

we mean a k-homomorphism $f: E \to E'$ such that the following diagram is commutative:

$$\begin{array}{ccc} E & \xrightarrow{f} & E' \\ {\scriptstyle A}\downarrow & & \downarrow{\scriptstyle A'} \\ E & \xrightarrow[f]{} & E' \end{array}$$

It is then immediate that such pairs form a category, so we have the notion of isomorphism. One can reformulate Theorem 2.1 by stating:

Theorem 2.6. *Two pairs (E, A) and (F, B) are isomorphic if and only if they have the same invariants.*

You can prove this as Exercise 19. The Jordan basis gives a normalized form for the matrix associated with such a pair and an appropriate basis.

In the next chapter, we shall find conditions under which a normalized matrix is actually diagonal, for hermitian, symmetric, and unitary operators over the complex numbers.

As an example and application of Theorem 2.6, we prove:

Corollary 2.7. *Let k be a field and let K be a finite separable extension of degree n. Let V be a finite dimensional vector space of dimension n over k, and let ρ, $\rho' : K \to \mathrm{End}_k(V)$ be two representations of K on V; that is, embeddings of K in $\mathrm{End}_k(V)$. Then ρ, ρ' are conjugate; that is, there exists $B \in \mathrm{Aut}_k(V)$ such that*

$$\rho'(\xi) = B\rho(\xi)B^{-1} \text{ for all } \xi \in K.$$

Proof. By the primitive element theorem of field theory, there exists an element $\alpha \in K$ such that $K = k[\alpha]$. Let $p(t)$ be the irreducible polynomial of α over k. Then $(V, \rho(\alpha))$ and $(V, \rho'(\alpha))$ have the same invariant, namely $p(t)$. Hence these pairs are isomorphic by Theorem 2.6, which means that there exists $B \in \mathrm{Aut}_k(V)$ such that

$$\rho'(\alpha) = B\rho(\alpha)B^{-1}.$$

But all elements of K are linear combinations of powers of α with coefficients in k, so it follows immediately that $\rho'(\xi) = B\rho(\xi)B^{-1}$ for all $\xi \in K$, as desired.

To get a representation of K as in corollary 2.7, one may of course select a basis of K, and represent multiplication of elements of K on K by matrices with respect to this basis. In some sense, Corollary 2.7 tells us that this is the only way to get such representations. We shall return to this point of view when considering Cartan subgroups of GL_n in Chapter XVIII, §12.

§3. THE CHARACTERISTIC POLYNOMIAL

Let k be a commutative ring and E a free module of dimension n over k. We consider the polynomial ring $k[t]$, and a linear map $A : E \to E$. We have a homomorphism

$$k[t] \to k[A]$$

as before, mapping a polynomial $f(t)$ on $f(A)$, and E becomes a module over the ring $R = k[t]$. Let M be any $n \times n$ matrix in k (for instance the matrix of A relative to a basis of E). We define the **characteristic polynomial** $P_M(t)$ to be the determinant

$$\det(tI_n - M)$$

where I_n is the unit $n \times n$ matrix. It is an element of $k[t]$. Furthermore, if N is an invertible matrix in R, then

$$\det(tI_n - N^{-1}MN) = \det(N^{-1}(tI_n - M)N) = \det(tI_n - M).$$

Hence the characteristic polynomial of $N^{-1}MN$ is the same as that of M. We may therefore define the characteristic polynomial of A, and denote by P_A, the characteristic polynomial of any matrix M associated with A with respect to some basis. (If $E = 0$, we **define the characteristic polynomial to be** 1.)

If $\varphi : k \to k'$ is a homomorphism of commutative rings, and M is an $n \times n$ matrix in k, then it is clear that

$$P_{\varphi M}(t) = \varphi P_M(t)$$

where φP_M is obtained from P_M by applying φ to the coefficients of P_M.

Theorem 3.1. (Cayley-Hamilton). *We have* $P_A(A) = 0$.

Proof. Let $\{v_1, \ldots, v_n\}$ be a basis of E over k. Then

$$tv_j = \sum_{i=1}^{n} a_{ij} v_i$$

where $(a_{ij}) = M$ is the matrix of A with respect to the basis. Let $\tilde{B}(t)$ be the matrix with coefficients in $k[t]$, defined in Chapter XIII, such that

$$\tilde{B}(t)B(t) = P_A(t)I_n.$$

Then

$$\tilde{B}(t)B(t)\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} P_A(t)v_1 \\ \vdots \\ P_A(t)v_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

because

$$B(t)\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Hence $P_A(t)E = 0$, and therefore $P_A(A)E = 0$. This means that $P_A(A) = 0$, as was to be shown.

Assume now that k is a field. Let E be a finite-dimensional vector space over k, and let $A \in \mathrm{End}_k(E)$. By an **eigenvector** w of A in E one means an element $w \in E$, such that there exists an element $\lambda \in k$ for which $Aw = \lambda w$. If $w \neq 0$, then λ is determined uniquely, and is called an **eigenvalue** of A. Of course, distinct eigenvectors may have the same eigenvalue.

Theorem 3.2. *The eigenvalues of A are precisely the roots of the characteristic polynomial of A.*

Proof. Let λ be an eigenvalue. Then $A - \lambda I$ is not invertible in $\mathrm{End}_k(E)$, and hence $\det(A - \lambda I) = 0$. Hence λ is a root of P_A. The arguments are reversible, so we also get the converse.

For simplicity of notation, we often write $A - \lambda$ instead of $A - \lambda I$.

Theorem 3.3. *Let $w_1, \ldots, w_m$ be non-zero eigenvectors of A, having distinct eigenvalues. Then they are linearly independent.*

Proof. Suppose that we have

$$a_1 w_1 + \cdots + a_m w_m = 0$$

with $a_i \in k$, and let this be a shortest relation with not all $a_i = 0$ (assuming such exists). Then $a_i \neq 0$ for all i. Let $\lambda_1, \ldots, \lambda_m$ be the eigenvalues of our vectors. Apply $A - \lambda_1$ to the above relation. We get

$$a_2(\lambda_2 - \lambda_1)w_2 + \cdots + a_m(\lambda_m - \lambda_1)w_m = 0,$$

which shortens our relation, contradiction.

Corollary 3.4. *If A has n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$ belonging to eigenvectors $v_1, \ldots, v_n$, and* $\dim E = n$, *then $\{v_1, \ldots, v_n\}$ is a basis for E. The matrix*

of A with respect to this basis is the diagonal matrix:

$$\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}.$$

Warning. It is not always true that there exists a basis of E consisting of eigenvectors!

Remark. Let k be a subfield of k'. If M is a matrix in k, we can define its characteristic polynomial with respect to k, and also with respect to k'. It is clear that the characteristic polynomials thus obtained are equal. If E is a vector space over k, we shall see later how to extend it to a vector space over k'. A linear map A extends to a linear map of the extended space, and the characteristic polynomial of the linear map does not change either. Actually, if we select a basis for E over k, then $E \approx k^{(n)}$, and $k^{(n)} \subset k'^{(n)}$ in a natural way. Thus selecting a basis allows us to extend the vector space, but this seems to depend on the choice of basis. We shall give an invariant definition later.

Let $E = E_1 \oplus \cdots \oplus E_r$ be an expression of E as a direct sum of vector spaces over k. Let $A \in \mathrm{End}_k(E)$, and assume that $AE_i \subset E_i$ for all $i = 1, \ldots, r$. Then A induces a linear map on E_i. We can select a basis for E consisting of bases for $E_1, \ldots, E_r$, and then the matrix for A consists of blocks. Hence we see that

$$P_A(t) = \prod_{i=1}^{r} P_{A_i}(t).$$

Thus the characteristic polynomial is multiplicative on direct sums.

Our condition above that $AE_i \subset E_i$ can also be formulated by saying that E is expressed as a $k[A]$-direct sum of $k[A]$-submodules, or also a $k[t]$-direct sum of $k[t]$-submodules. We shall apply this to the decomposition of E given in Theorem 2.1.

Theorem 3.5. *Let E be a finite-dimensional vector space over a field k, let $A \in \mathrm{End}_k(E)$, and let $q_1, \ldots, q_r$ be the invariants of (E, A). Then*

$$P_A(t) = q_1(t) \cdots q_r(t).$$

Proof. We assume that $E = k^{(n)}$ and that A is represented by a matrix M. We have seen that the invariants do not change when we extend k to a larger field, and neither does the characteristic polynomial. Hence we may assume that k is algebraically closed. In view of Theorem 2.1 we may assume that M has a

single invariant q. Write

$$q(t) = (t - \alpha_1)^{e_1} \cdots (t - \alpha_s)^{e_s}$$

with distinct $\alpha_1, \ldots, \alpha_s$. We view M as a linear map, and split out vector space further into a direct sum of submodules (over $k[t]$) having invariants

$$(t - \alpha_1)^{e_1}, \ldots, (t - \alpha_s)^{e_s}$$

respectively (this is the prime power decomposition). For each one of these submodules, we can select a basis so that the matrix of the induced linear map has the shape described in Theorem 2.4. From this it is immediately clear that the characteristic polynomial of the map having invariant $(t - \alpha)^e$ is precisely $(t - \alpha)^e$, and our theorem is proved.

Corollary 3.6. *The minimal polynomial of A and its characteristic polynomial have the same irreducible factors.*

Proof. Because q_r is the minimal polynomial, by Theorem 2.1.

We shall generalize our remark concerning the multiplicativity of the characteristic polynomial over direct sums.

Theorem 3.7. *Let k be a commutative ring, and in the following diagram,*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & E' & \longrightarrow & E & \longrightarrow & E'' & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle A'} & & \downarrow{\scriptstyle A} & & \downarrow{\scriptstyle A''} & & \\ 0 & \longrightarrow & E' & \longrightarrow & E & \longrightarrow & E'' & \longrightarrow & 0 \end{array}$$

let the rows be exact sequences of free modules over k, of finite dimension, and let the vertical maps be k-linear maps making the diagram commutative. Then

$$P_A(t) = P_{A'}(t)P_{A''}(t).$$

Proof. We may assume that E' is a submodule of E. We select a basis $\{v_1, \ldots, v_m\}$ for E'. Let $\{\bar{v}_{m+1}, \ldots, \bar{v}\}$ be a basis for E'', and let $v_{m+1}, \ldots, v_n$ be elements of E mapping on $\bar{v}_{m+1}, \ldots, \bar{v}_n$ respectively. Then

$$\{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}$$

is a basis for E (same proof as Theorem 5.2 of Chapter III), and we are in the situation discussed in §1. The matrix for A has the shape

$$\begin{pmatrix} M' & * \\ 0 & M'' \end{pmatrix}$$

where M' is the matrix for A' and M'' is the matrix for A''. Taking the characteristic polynomial with respect to this matrix obviously yields our multiplicative property.

Theorem 3.8. *Let k be a commutative ring, and E a free module of dimension n over k. Let $A \in \mathrm{End}_k(E)$. Let*

$$P_A(t) = t^n + c_{n-1}t^{n-1} + \cdots + c_0.$$

Then

$$\mathrm{tr}(A) = -c_{n-1} \quad \textit{and} \quad \det(A) = (-1)^n c_0.$$

Proof. For the determinant, we observe that $P_A(0) = c_0$. Substituting $t = 0$ in the definition of the characteristic polynomial by the determinant shows that $c_0 = (-1)^n \det(A)$.

For the trace, let M be the matrix representing A with respect to some basis, $M = (a_{ij})$. We consider the determinant $\det(tI_n - a_{ij})$. In its expansion as a sum over permutations, it will contain a diagonal term

$$(t - a_{11}) \cdots (t - a_{nn}),$$

which will give a contribution to the coefficient of t^{n-1} equal to

$$-(a_{11} + \cdots + a_{nn}).$$

No other term in this expansion will give a contribution to the coefficient of t^{n-1}, because the power of t occurring in another term will be at most t^{n-2}. This proves our assertion concerning the trace.

Corollary 3.9. *Let the notation be as in Theorem* 3.7. *Then*

$$\mathrm{tr}(A) = \mathrm{tr}(A') + \mathrm{tr}(A'') \quad \textit{and} \quad \det(A) = \det(A')\det(A'').$$

Proof. Clear.

We shall now interpret our results in the Euler-Grothendieck group.

Let k be a commutative ring. We consider the category whose objects are pairs (E, A), where E is a k-module, and $A \in \mathrm{End}_k(E)$. We define a morphism

$$(E', A') \to (E, A)$$

to be a k-linear map $E' \xrightarrow{f} E$ making the following diagram commutative:

$$\begin{array}{ccc} E' & \xrightarrow{f} & E \\ {\scriptstyle A'}\downarrow & & \downarrow{\scriptstyle A} \\ E' & \xrightarrow[f]{} & E \end{array}$$

Then we can define the kernel of such a morphism to be again a pair. Indeed, let E'_0 be the kernel of $f: E' \to E$. Then A' maps E'_0 into itself because

$$fA'E'_0 = AfE'_0 = 0.$$

We let A'_0 be the restriction of A' on E'_0. The pair (E'_0, A'_0) is defined to be the kernel of our morphism.

We shall denote by f again the morphism of the pair $(E', A') \to (E, A)$. We can speak of an exact sequence

$$(E', A') \to (E, A) \to (E'', A''),$$

meaning that the induced sequence

$$E' \to E \to E''$$

is exact. We also write 0 instead of (0, 0), according to our universal convention to use the symbol 0 for all things which behave like a zero element.

We observe that our pairs now behave formally like modules, and they in fact form an abelian category.

Assume that k is a field. Let $\mathfrak{A}$ consist of all pairs (E, A) where E is finite dimensional over k.

Then Theorem 3.7 *asserts that the characteristic polynomial is an Euler-Poincaré map defined for each object in our category* $\mathfrak{A}$, *with values into the multiplicative monoid of polynomials with leading coefficient* 1.

Since the values of the map are in a monoid, this generalizes slightly the notion of Chapter III, §8, when we took the values in a group. Of course when k is a field, which is the most frequent application, we can view the values of our map to be in the multiplicative group of non-zero rational functions, so our previous situation applies.

A similar remark holds now for the trace and the determinant. *If k is a field, the trace is an Euler map into the additive group of the field, and the determinant is an Euler map into the multiplicative group of the field.* We note also that all these maps (like all Euler maps) are defined on the isomorphism classes of pairs, and are defined on the Euler-Grothendieck group.

Theorem 3.10. *Let k be a commutative ring, M an $n \times n$ matrix in k, and f a polynomial in $k[t]$. Assume that $P_M(t)$ has a factorization,*

$$P_M(t) = \prod_{i=1}^{n} (t - \alpha_i)$$

into linear factors over k. Then the characteristic polynomial of $f(M)$ is given by

$$P_{f(M)}(t) = \prod_{i=1}^{n} (t - f(\alpha_i)),$$

and

$$\operatorname{tr}(f(M)) = \sum_{i=1}^{n} f(\alpha_i), \qquad \det(f(M)) = \prod_{i=1}^{n} f(\alpha_i).$$

Proof. Assume first that k is a field. Then using the canonical decomposition in terms of matrices given in Theorem 2.4, we find that our assertion is immediately obvious. When k is a ring, we use a substitution argument. It is however necessary to know that if $X = (x_{ij})$ is a matrix with algebraically independent coefficients over $\mathbf{Z}$, then $P_X(t)$ has n distinct roots $y_1, \ldots, y_n$ [in an algebraic closure of $\mathbf{Q}(X)$] and that we have a homomorphism

$$\mathbf{Z}[x_{ij}, y_1, \ldots, y_n] \to k$$

mapping X on M and $y_1, \ldots, y_n$ on $\alpha_1, \ldots, \alpha_n$. This is obvious to the reader who read the chapter on integral ring extensions, and the reader who has not can forget about this part of the theorem.

EXERCISES

1. Let T be an upper triangular square matrix over a commutative ring (i.e. all the elements below and on the diagonal are 0). Show that T is nilpotent.
2. Carry out explicitly the proof that the determinant of a matrix

$$\begin{pmatrix} M_1 & & & * & * \\ 0 & M_2 & & & \\ 0 & 0 & \ddots & & * \\ \vdots & \vdots & & & \\ 0 & 0 & \cdots & 0 & M_s \end{pmatrix}$$

   where each M_i is a square matrix, is equal to the product of the determinants of the matrices $M_1, \ldots, M_s$.
3. Let k be a commutative ring, and let M, M' be square $n \times n$ matrices in k. Show that the characteristic polynomials of MM' and $M'M$ are equal.
4. Show that the eigenvalues of the matrix

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

   in the complex numbers are $\pm 1, \pm i$.

5. Let M, M' be square matrices over a field k. Let q, q' be their respective minimal polynomials. Show that the minimal polynomial of

$$\begin{pmatrix} M & 0 \\ 0 & M' \end{pmatrix}$$

is the least common multiple of q, q'.

6. Let A be a nilpotent endomorphism of a finite dimensional vector space E over the field k. Show that $\operatorname{tr}(A) = 0$.

7. Let R be a principal entire ring. Let E be a free module over R, and let $E^\vee = \operatorname{Hom}_R(E, R)$ be its dual module. Then $E^\vee$ is free of dimension n. Let F be a submodule of E. Show that $E^\vee/F^\perp$ can be viewed as a submodule of $F^\vee$, and that its invariants are the same as the invariants of F in E.

8. Let E be a finite-dimensional vector space over a field k. Let $A \in \operatorname{Aut}_k(E)$. Show that the following conditions are equivalent:
   (a) $A = I + N$, with N nilpotent.
   (b) There exists a basis of E such that the matrix of A with respect to this basis has all its diagonal elements equal to 1 and all elements above the diagonal equal to 0.
   (c) All roots of the characteristic polynomial of A (in the algebraic closure of k) are equal to 1.

9. Let k be a field of characteristic 0, and let M be an $n \times n$ matrix in k. Show that M is nilpotent if and only if $\operatorname{tr}(M^v) = 0$ for $1 \leqq v \leqq n$.

10. Generalize Theorem 3.10 to rational functions (instead of polynomials), assuming that k is a field.

11. Let E be a finite-dimensional space over the field k. Let $\alpha \in k$. Let E_α be the subspace of E generated by all eigenvectors of a given endomorphism A of E, having α as an eigenvalue. Show that every non-zero element of E_α is an eigenvector of A having α as an eigenvalue.

12. Let E be finite dimensional over the field k. Let $A \in \operatorname{End}_k(E)$. Let v be an eigenvector for A. Let $B \in \operatorname{End}_k(E)$ be such that $AB = BA$. Show that Bv is also an eigenvector for A (if $Bv \neq 0$), with the same eigenvalue.

Diagonalizable endomorphisms

Let E be a finite-dimensional vector space over a field k, and let $S \in \operatorname{End}_k(E)$. We say that S is **diagonalizable** if there exists a basis of E consisting of eigenvectors of S. The matrix of S with respect to this basis is then a diagonal matrix.

13. (a) If S is diagonalizable, then its minimal polynomial over k is of type

$$q(t) = \prod_{i=1}^{m} (t - \lambda_i),$$

where $\lambda_1, \ldots, \lambda_m$ are distinct elements of k.

(b) Conversely, if the minimal polynomial of S is of the preceding type, then S is diagonalizable. [*Hint*: The space can be decomposed as a direct sum of the subspaces E_{λ_i} annihilated by $S - \lambda_i$.]

(c) If S is diagonalizable, and if F is a subspace of E such that $SF \subset F$, show that S is diagonalizable as an endomorphism of F, i.e. that F has a basis consisting of eigenvectors of S.

(d) Let S, T be endomorphisms of E, and assume that S, T commute. Assume that both S, T are diagonalizable. Show that they are simultaneously diagonalizable, i.e. there exists a basis of E consisting of eigenvectors for both S and T. [*Hint*: If λ is an eigenvalue of S, and E_λ is the subspace of E consisting of all vectors v such that $Sv = \lambda v$, then $TE_\lambda \subset E_\lambda$.]

14. Let E be a finite-dimensional vector space over an algebraically closed field k. Let $A \in \mathrm{End}_k(E)$. Show that A can be written in a unique way as a sum

$$A = S + N$$

where S is diagonalizable, N is nilpotent, and $SN = NS$. Show that S, N can be expressed as polynomials in A. [*Hint*: Let $P_A(t) = \prod (t - \lambda_i)^{m_i}$ be the factorization of $P_A(t)$ with distinct λ_i. Let E_i be the kernel of $(A - \lambda_i)^{m_i}$. Then E is the direct sum of the E_i. Define S on E so that on E_i, $Sv = \lambda_i v$ for all $v \in E_i$. Let $N = A - S$. Show that S, N satisfy our requirements. To get S as a polynomial in A, let g be a polynomial such that $g(t) \equiv \lambda_i \bmod (t - \lambda_i)^{m_i}$ for all i, and $g(t) \equiv 0 \bmod t$. Then $S = g(A)$ and $N = A - g(A)$.]

15. After you have read the section on the tensor product of vector spaces, you can easily do the following exercise. Let E, F be finite-dimensional vector spaces over an algebraically closed field k, and let $A : E \to E$ and $B : F \to F$ be k-endomorphisms of E, F, respectively. Let

$$P_A(t) = \prod (t - \alpha_i)^{n_i} \quad \text{and} \quad P_B(t) = \prod (t - \beta_j)^{m_j}$$

be the factorizations of their respectively characteristic polynomials, into distinct linear factors. Then

$$P_{A \otimes B}(t) = \prod_{i,j} (t - \alpha_i \beta_j)^{n_i m_j}.$$

[*Hint*: Decompose E into the direct sum of subspaces E_i, where E_i is the subspace of E annihilated by some power of $A - \alpha_i$. Do the same for F, getting a decomposition into a direct sum of subspaces F_j. Then show that some power of $A \otimes B - \alpha_i \beta_j$ annihilates $E_i \otimes F_j$. Use the fact that $E \otimes F$ is the direct sum of the subspaces $E_i \otimes F_j$, and that $\dim_k(E_i \otimes F_j) = n_i m_j$.]

16. Let Γ be a free abelian group of dimension $n \geqq 1$. Let Γ' be a subgroup of dimension n also. Let $\{v_1, \ldots, v_n\}$ be a basis of Γ, and let $\{w_1, \ldots, w_n\}$ be a basis of Γ'. Write

$$w_i = \sum a_{ij} v_j.$$

Show that the index $(\Gamma : \Gamma')$ is equal to the absolute value of the determinant of the matrix (a_{ij}).

17. Prove the normal basis theorem for finite extensions of a finite field.

18. Let $A = (a_{ij})$ be a square $n \times n$ matrix over a commutative ring k. Let A_{ij} be the matrix obtained by deleting the i-th row and j-th column from A. Let $b_{ij} = (-1)^{i+j} \det(A_{ji})$, and let B be the matrix (b_{ij}). Show that $\det(B) = \det(A)^{n-1}$, by reducing the problem to the case when A is a matrix with variable coefficients over the integers. Use this same method to give an alternative proof of the Cayley-Hamilton theorem, that $P_A(A) = 0$.

19. Let (E, A) and (E', A') be pairs consisting of a finite-dimensional vector space over a field k, and a k-endomorphism. Show that these pairs are isomorphic if and only if their invariants are equal.

20. (a) How many non-conjugate elements of $GL_2(\mathbf{C})$ are there with characteristic polynomial $t^3(t+1)^2(t-1)$?
(b) How many with characteristic polynomial $t^3 - 1001t$?

21. Let V be a finite dimensional vector space over $\mathbf{Q}$ and let $A: V \to V$ be a $\mathbf{Q}$-linear map such that $A^5 = \mathrm{Id}$. Assume that if $v \in V$ is such that $Av = v$, then $v = 0$. Prove that $\dim V$ is divisible by 4.

22. Let V be a finite dimensional vector space over $\mathbf{R}$, and let $A: V \to V$ be an $\mathbf{R}$-linear map such that $A^2 = -\mathrm{Id}$. Show that $\dim V$ is even, and that V is a direct sum of 2-dimensional A-invariant subspaces.

23. Let E be a finite-dimensional vector space over an algebraically closed field k. Let A, B be k-endomorphisms of E which commute, i.e. $AB = BA$. Show that A and B have a common eigenvector. [*Hint*: Consider a subspace consisting of all vectors having a fixed element of k as eigenvalue.]

24. Let V be a finite dimensional vector space over a field k. Let A be an endomorphism of V. Let $\mathrm{Tr}(A^m)$ be the trace of A^m as an endomorphism of V. Show that the following power series in the variable t are equal:

$$\exp\left(\sum_{m=1}^{\infty} -\mathrm{Tr}(A^m)\frac{t^m}{m}\right) = \det(I - tA) \quad \text{or} \quad -\frac{d}{dt}\log\det(I - tA) = \sum_{m=1}^{\infty}\mathrm{Tr}(A^m)t^m.$$

Compare with Exercise 23 of Chapter XVIII.

25. Let V, W be finite dimensional vector spaces over k, of dimension n. Let $(v, w) \mapsto \langle v, w\rangle$ be a non-singular bilinear form on $V \times W$. Let $c \in k$, and let $A: V \to V$ and $V: W \to W$ be endomorphisms such that

$$\langle Av, Bw\rangle = c\langle v, w\rangle \text{ for all } v \in V \text{ and } w \in W.$$

Show that

$$\det(A)\det(tI - B) = (-1)^n\det(cI - tA)$$

and

$$\det(A)\det(B) = c^n.$$

For an application of Exercises 24 and 25 to a context of topology or algebraic geometry, see Hartshorne's *Algebraic Geometry*, Appendix C, §4.

26. Let $G = SL_n(\mathbf{C})$ and let K be the complex unitary group. Let A be the group of diagonal matrices with positive real components on the diagonal.
(a) Show that if $g \in \mathrm{Nor}_G(A)$ (normalizer of A in G), then $\mathbf{c}(g)$ (conjugation by g) permutes the diagonal components of A, thus giving rise to a homomorphism $\mathrm{Nor}_G(A) \to W$ to the group W of permutations of the diagonal coordinates.

By definition, the kernel of the above homomorphism is the centralizer $\mathrm{Cen}_G(A)$.

(b) Show that actually all permutations of the coordinates can be achieved by elements of K, so we get an isomorphism

$$W \approx \mathrm{Nor}_G(A)/\mathrm{Cen}_G(A) \approx \mathrm{Nor}_K(A)/\mathrm{Cen}_K(A).$$

In fact, the K on the right can be taken to be the real unitary group, because permutation matrices can be taken to have real components (0 or ± 1).