Following the successful conclusion of the theoretical part of the course (the canonical form theory of a single linear transformation) we now face the not insignificant task of finding canonical forms in practice. If we could get our hands on the minimal polynomial, $m_T(x)$, of a given transformation $T \in \mathrm{End}_{\mathbb{F}}(V)$ then we would have made good progress. The key tool which will help us is yet another polynomial which has close relationship with $m_T(x)$; the so-called characteristic polynomial.

Before we begin, let us review some elementary facts about determinants (see Curtis Chapter 5, for a more detailed discussion). Let $A = [[\alpha_{ij}]] \in \mathbb{M}_m(\mathbb{F})$ and denote the ith row of A by $a_i = (\alpha_{i1}, \dots, \alpha_{in})$ $(1 \leq i \leq n)$. A <u>DETERMINANT FUNCTION</u> is a function $\det\colon \mathbb{M}_n(\mathbb{F}) \to \mathbb{F}$, satisfying the following conditions:

1. $\det(a_1, \dots, a_{i-1}, a_i + a_j, a_{i+1}, \dots, a_n) = \det(A)$ for $1 \leq i \neq j \leq n$;
2. $\det(a_1, \dots, a_{i-1}, \lambda a_i, a_{i+1}, \dots, a_n) = \lambda \det(A)$ for $1 \leq i \leq n$; and
3. $\det(I_n) = 1$.

It turns out, of course, that there is such a function, and it is unique; we therefore refer to det as *the* determinant function. We now give a concrete, recursive definition of det. If $n = 1$, put $\det(A) = \det([[\alpha]]) := \alpha$. For $n > 1$ and $1 \leq i, j \leq n$, let A_{ij} denote the $(n-1) \times (n-1)$ matrix obtained by deleting the ith row and jth column of $A \in \mathbb{M}_n(\mathbb{F})$. Now, for any $1 \leq i \leq n$, define

$$\det(A) := \sum_{j=1}^{n} (-1)^{i+j} \alpha_{ij} \det(A_{ij}).$$

Among the many useful properties of the determinant, we highlight two for immediate use.

(i) For $A, B \in \mathbb{M}_n(\mathbb{F})$, $\det(AB) = \det(BA) = \det(A)\det(B)$.

(ii) If $A \in \mathbb{M}_n(\mathbb{F})$ has an inverse, then $\det(A^{-1}) = 1/\det(A)$.

We next wish to define the determinant of a linear transformation. Let $T \in \mathrm{End}_{\mathbb{F}}(V)$, choose a basis $\mathcal{B}$ of V, and put $\det(T) := \det({}_{\mathcal{B}}A_T)$. Apparently, this definition depends upon the choice of $\mathcal{B}$, but the next result states that this is not the case.

(Ch1) Lemma. *If* $A' \sim A$ *then* $\det(A') = \det(A)$.

Proof. $A' \sim A$ iff there exists invertible P such that $A' = PAP^{-1}$. But then

$$\det(A') = \det(PAP^{-1}) = \det(P)\det(A)\det(P^{-1}) = \det(P)\det(P)^{-1}\det(A) = \det(A). \qquad \square$$

Fix $T \in \mathrm{End}_{\mathbb{F}}(V)$. We will call a scalar $\lambda \in \mathbb{F}$ an EIGENVALUE of T if there exists $0 \neq v \in V$ such that $vT = \lambda v$; such a vector v is called an EIGENVECTOR of T corresponding to λ. For an eigenvalue λ of T, define the λ-EIGENSPACE of T to be

$$V_T(\lambda) := \{v \in V \mid v \text{ is an eigenvector of } T \text{ corresponding to } \lambda\}.$$

We leave it as an easy exercise to show that $V_T(\lambda)$ is a subspace of V.

Let $\lambda \in \mathbb{F}$ and let 1_V denote the identity transformation on V. Then λ is an eigenvalue of T iff $\exists 0 \neq v \in \mathrm{NS}(\lambda 1_V - T)$ iff $\lambda 1_V - T$ is not invertible iff $\det(\lambda 1_V - T) = 0$. Define the CHARACTERISTIC POLYNOMIAL of T, denoted $c_T(x)$, to be

$$c_T(x) := \det(x1_V - T),$$

a monic polynomial of degree $n = \dim(V)$; the roots of $c_T(x)$ are called the CHARACTERISTIC VALUES of T. Hence, we have proved:

(Ch2) Lemma. *The eigenvalues of T are the characteristic values of T.* □

We define $c_A(x)$, for $A \in \mathbb{M}_n(\mathbb{F})$ in the obvious way. Note that, while computing determinants is not particularly nice for large matrices, at least we have a formula for $c_T(x)$; the minimal polynomial $m_T(x)$, on the other hand, is somewhat more elusive. The main goal of this lecture is to establish a connection between those two polynomials.

Here are some elementary properties of $c_T(x)$ that we will need (recall that $C(f)$ denotes the companion matrix of the monic polynomial $f(x) \in \mathbb{F}[x]$ and, for a T-invariant subspace W, T_W denotes the restriction of T to W).

(Ch3) Lemma. *Let $T \in \mathrm{End}_{\mathbb{F}}(V)$.*

(i) If $V = W_1 \oplus W_2$ is a T-invariant decomposition of V, then $c_T(x) = c_{T_{W_1}}(x) \cdot c_{T_{W_2}}(x)$.

(ii) If $f \in \mathbb{F}[x]$, then $c_{C(f)}(x) = f(x)$.

(iii) If V is T-cyclic then $m_T(x) = c_T(x)$.

Proof. For (i), if $W_1 \oplus W_2$ is a T-invariant decomposition, then select a basis $\mathcal{B} = \mathcal{B}_1, \mathcal{B}_2$ of V with $\mathcal{B}_i$ a basis for W_i $(i = 1, 2)$. It follows that

$$A = {}_{\mathcal{B}}A_T = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

where, if T_i denotes $T|_{V_i}$, we have $A_i = {}_{\mathcal{B}_i}A_{T_i}$. Hence

$$\det(xI_n - A) = \det \begin{pmatrix} xI_{n_1} - A_1 & 0 \\ 0 & xI_{n_2} - A_2 \end{pmatrix} = \det(xI_{n_1} - A_1) \cdot \det(xI_{n_2} - A_2) = c_{T_1}(x) \cdot c_{T_2}(x).$$

Part (ii) is **(Can)**, Exercise 2. Part (iii) now follows easily. For, if V is T-cyclic, then there exists a basis of V relative to which T is represented by the matrix $C(m_T)$. Now, by part (ii) of the lemma, we have $c_T(x) = c_{C(m_T)}(x) = m_T(x)$. □

We can now prove the main result of this lecture which will be of great value in trying to compute the minimal polynomial of a linear transformation.

(Ch4) Theorem. [Generalised Cayley-Hamilton Theorem] *Let $T \in \mathrm{End}_{\mathbb{F}}(V)$.*

(a) $m_T(x)$ divides $c_T(x)$.

(b) $c_T(x)$ is the product of the invariant factors of T (and hence also the product of the elementary divisors of T).

(c) $m_T(x)$ and $c_T(x)$ have the same irreducible factors except for multiplicities.

(d) If $m_T(x) = p_1(x)^{n_1} \dots p_k(x)^{n_k}$ is the unique factorisation of $m_T(x)$ into monic irreducibles, then $c_T(x) = p_1(x)^{d_1} \dots p_k(x)^{d_k}$, where $d_i = n(p_i(T)^{n_i})/\deg(p_i)$.

Proof. By **(Dec4)**, there exist unique monic polynomials

$$m_T(x) = q_1(x), q_2(x), \dots, q_m(x)$$

and a T-cyclic decomposition $V = V_1 \oplus V_2 \oplus \dots \oplus V_m$ such that $q_i(x)$ is the minimal polynomial of T_{V_i} for $1 \leq i \leq m$. By **(Ch3)**(iii), $c_{T_{V_i}}(x) = m_{T_{V_i}}(x) = q_i(x)$ and, by **(Ch3)**(i),

$$c_T(x) = \prod_{i=1}^{m} c_{T_{V_i}}(x) = \prod_{i=1}^{m} q_i(x).$$

This proves (a) and (b). For (c) note that, if an irreducible divides $c_T(x) = q_1(x) \dots q_m(x)$, then it must divide $q_i(x)$ for some $1 \leq i \leq m$; since $q_i(x)|q_1(x)$, it must also divide $m_T(x)$. The converse is trivial in view of (a). To prove (d) we turn to our other big decomposition theorem, **(Dec3)**. For $1 \leq i \leq k$, if $W_i = \mathrm{NS}(p_i(T)^{n_i})$, we have a T-invariant decomposition $V = W_1 \oplus W_2 \oplus \dots \oplus W_k$ (the subspace W_i is well known to us; it is just the sum of all of the V_{ij}s in **(Dec3)**). Let T_i denote the restriction of T to W_i. Then $m_{T_i}(x) = p_i(x)^{n_i}$ so, by part (c) of this theorem, $c_{T_i}(x) = p_i(x)^{d_i}$ for some $d_i \geq n_i$. Hence

$$n(p_i(T)^{n_i}) = \dim(W_i) = \deg(c_{T_{W_i}}(x)) = d_i \deg(p_i),$$

and the result follows. □

Remark: Let's summarise the situation for a linear transformation T: the invariant factors of T determine T up to similarity (or, alternatively, $\mathrm{GL}_n(\mathbb{F})$-conjugacy); $m_T(x)$ is the first invariant factor; and $c_T(x)$ is the product of the invariant factors. Hence knowledge of both the minimal, and characteristic polynomials of T gives a great deal of information about T, but they alone do not completely determine T.

We have used eigenvalues to introduce the characteristic polynomial as a computational means of obtaining information about the minimal polynomial. However, the determination and study of eigenvalues is of independent interest. We therefore conclude this lecture with some elementary properties of, and comments concerning, eigenvalues and eigenspaces.

Suppose that $(x-\lambda)|c_T(x)$ and suppose that m is the largest integer such that $(x-\lambda)^m|c_T(x)$; then we call $a(\lambda) := m$ the ALGEBRAIC MULTIPLICITY of λ. Next we define the GEOMETRIC MULTIPLICITY of λ to be the integer $g(\lambda) := \dim(V_T(\lambda))$. It follows from **(Dec3)** and **(Ch4)**(c) that the T-invariant subspace $W := NS((x-\lambda)^m)$ has dimension m. It is clear also that $V_T(\lambda) \leq W$, whence $g(\lambda) \leq a(\lambda)$. Moreover, $V_T(\lambda) = W$ if and only if the minimal polynomial of T restricted to W is $x - \lambda$, in which case T induces the scalar transformation $\lambda 1_W$ on W. The following alternate characterisation of diagonalisability now follows easily

(Ch5) Lemma. *$T \in \mathrm{End}_{\mathbb{F}}(V)$ is diagonalisable if and only if $c_T(x)$ is a product of linear factors (not necessarily distinct) and the geometric and algebraic multiplicities coincide.* □

Exercises.

1. Consider the 4×4 real matrix
$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ -2 & -2 & 2 & 1 \\ 1 & 1 & -1 & 0 \end{pmatrix}.$$
Show that $c_A(x) = x^2(x-1)^2 = m_A(x)$.
2. Find a 3×3 matrix A such that $c_A(x) = x^3$ and $m_A(x) = x^2$.
3. Let $V = \mathbb{M}_n(\mathbb{F})$ be the vector space of $n \times n$ matrices over $\mathbb{F}$ and let $A \in \mathbb{M}_n(\mathbb{F})$ be fixed. Let T_A be the element of $\mathrm{End}_{\mathbb{F}}(V)$ defined by
$$T_A \colon M \mapsto MA.$$

Show that $m_{T_A}(x) = m_A(x)$.

4. Let $T \in \mathrm{End}_{\mathbb{F}}(V)$ and let $\lambda \neq \mu$ be elements of $\mathbb{F}$. Show directly that

$$V_T(\lambda) + V_T(\mu) = V_T(\lambda) \oplus V_T(\mu).$$

5. Let $N \in \mathbb{M}_n(\mathbb{F})$ be nilpotent. Show that $N^n = 0$.

6. Let A be any 3×3 matrix whose nullspace is 2-dimensional. For each assertion below, provide either a proof or counterexample.

   (a) x^2 divides $c_A(x)$.

   (b) The trace of A (the sum of its diagonal entries) is an eigenvalue of A.

   (c) A is diagonalisable.

7. Let V be a $\mathbb{Q}$-space of dimension n. Let φ be an automorphism of V which fixes no nonzero vector. Suppose that φ^p is the identity map on V, where p is a prime number. Show that $p-1$ divides n.