We now begin in earnest our prolongued excursion into the life of a single linear transformation. This is where we will finally see the pay-off for all of our hard work studying the structure of modules over a PID. First let us set up some notation and terminology.

Fix an $\mathbb{F}$-vector space V of dimension n. For vectors $v_1, \ldots, v_r \in V$, put

$$\mathrm{sp}(v_1, \ldots, v_r) := \mathbb{F}\{v_1, \ldots, v_r\} = \{\alpha_1 v_1 + \ldots + \alpha_r v_r \mid \alpha_i \in \mathbb{F}\},$$

the $\underline{\mathbb{F}\text{-LINEAR SPAN}}$ of $v_1, \ldots, v_r$. For $1 \le i \le k$, let A_i be an $n_i \times n_i$ matrix for some positive integer n_i where $n_1 + n_2 + \ldots + n_k = n$. Then define

$$\mathrm{diag}(A_1, A_2, \ldots, A_k) \;:=\; \begin{pmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & A_k \end{pmatrix}$$

where the block entry entry "0" denotes the zero matrix of the appropriate dimensions (for example, the $(2,1)$ entry is the $n_1 \times n_2$ zero matrix). Suppose that, for some $T \in \mathrm{End}_{\mathbb{F}}(V)$ and some basis $\mathcal{B}$ of V, we have ${}_{\mathcal{B}}A_T = \mathrm{diag}(A_1, \ldots, A_k)$ for some square matrices A_i, as above. Cluster the vectors in $\mathcal{B}$ together into groups as follows:

$$\mathcal{B} = v_{11}, \ldots, v_{1n_1}, v_{21}, \ldots, v_{2n_2}, \ldots, v_{k1}, \ldots, v_{kn_k},$$

and, for $1 \le i \le k$, put $V_i := \mathrm{sp}(v_{i1}, \ldots, v_{in_i})$, of $v_{i1}, \ldots, v_{in_i}$. Then clearly we have $V = V_1 \oplus \ldots \oplus V_k$. In addition, T moves the vectors of V_i only within V_i; that is, $V_i T \le V_i$. We say that a subspace $W \le V$ is $\underline{T\text{-INVARIANT}}$ in case $WT \le W$. Hence, for our T, each V_i is T-invariant and $V = V_1 \oplus \ldots \oplus V_k$ is called a $\underline{T\text{-INVARIANT DIRECT SUM DECOMPOSITION OF } V}$.

Next suppose that W is any T-invariant subspace of V and let $\mathcal{B}_W = w_1, \ldots w_l$ be any basis of W. Let T_W denote the restriction of T to W. Since W is T-invariant, we have $T_W \in \mathrm{End}_{\mathbb{F}}(W)$; let A_W denote ${}_{\mathcal{B}_W}A_{T_W}$, the $l \times l$ matrix of T_W relative to $\mathcal{B}_W$. Now extend $\mathcal{B}_W$ to a basis of V, say $\mathcal{B} = w_1, \ldots, w_l, u_1, \ldots, u_{n-l}$. Then we have

$${}_{\mathcal{B}}A_T = \begin{pmatrix} A_W & 0 \\ B & C \end{pmatrix},$$

where B is $(n-l) \times l$ and C is $(n-l) \times (n-l)$. Let $\mathcal{B}_U = u_1, \ldots, u_{n-l}$ and $U = \mathrm{sp}(u_1, \ldots, u_{n-l})$; then we have $V = W \oplus U$. Now suppose that U is also T-invariant (so that $V = W \oplus U$ is a T-invariant direct sum decomposition). Then $B = 0$ and $C = A_U$ is the $(n-l) \times (n-l)$ matrix ${}_{\mathcal{B}_U}A_{T_U}$. A simple induction argument now gives us the following nice fact.

(Dec1) Theorem. *Let V be an $\mathbb{F}$-vector space of dimension n and let $T \in \mathrm{End}_{\mathbb{F}}(V)$. Then there exists a basis $\mathcal{B}$ of V and square matrices $A_1, \ldots, A_k$ such that ${}_{\mathcal{B}}A_T = \mathrm{diag}(A_1, \ldots, A_k)$ if and only if there exists a T-invariant decomposition $V = V_1 \oplus \ldots \oplus V_k$ of V. Moreover $\mathcal{B}$ is the concatenation $\mathcal{B}_1, \ldots, \mathcal{B}_k$, where $\mathcal{B}_i$ is a basis of V_i such that $A_i = {}_{\mathcal{B}_i}A_{T_{V_i}}$ for $1 \leq i \leq k$.* $\square$

> **Remark:** It is clear from **(Dec1)** that it is in our interest to investigate ways of finding T-invariant direct sum decompositions of the vector space V. For, if we can, then it suffices to study the restriction of T to each of the direct summands.

Let us revisit Exercise 6 of the **(Mod)** lecture. Fix $T \in \mathrm{End}_{\mathbb{F}}(V)$ and let $\langle T \rangle$ denote the subring of $\mathrm{End}_{\mathbb{F}}(V)$ defined by

$$\langle T \rangle \;=\; \{f(T) \mid f(x) \in \mathbb{F}[x]\};$$

then $\varphi_T \colon \mathbb{F}[x] \to \langle T \rangle$ sending $f(x) \mapsto f(T)$ is an epimorphism of rings. Recall that, as an $\mathbb{F}$-vector space, $\mathrm{End}_{\mathbb{F}}(V)$ has dimension n^2. It follows that the n^2+1 endomorphisms

$$1,\ T,\ T^2,\ \ldots,\ T^{n^2}$$

are linearly dependant. That is, there exist $\alpha_0, \alpha_1, \ldots, \alpha_{n^2} \in \mathbb{F}$, not all zero, such that

$$\alpha_0 + \alpha_1 T + \ldots + \alpha_{n^2} T^{n^2} = 0,$$

the zero transformation of V. Defining $k(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_{n^2} x^{n^2}$, we have $k(x) \in \ker \varphi_T$. In particular, the kernel is nonzero. Since $\mathbb{F}[x]$ is a PID, it follows that $\ker \varphi_T = (m_T(x))$, where $m_T(x)$ is the **unique** monic polynomial generating this principal ideal. The polynomial $m_T(x)$ is absolutely central to the study of the linear transformation T; it is called the MINIMAL POLYNOMIAL OF T. Observe that we can factorize $m_T(x)$ uniquely as

$$m_T(x) = p_1(x)^{n_1} p_2(x)^{n_2} \ldots p_k(x)^{n_k},$$

where each $p_i(x)$ is a monic irreducible polynomial in $\mathbb{F}[x]$.

Next, we use φ_T to define an action of the ring $\mathbb{F}[x]$ on the vector space V. For $v \in V$ and $f(x) \in \mathbb{F}[x]$, define

$$vf(x) \;:=\; v\varphi_T(f(x)) \;=\; vf(T).$$

This turns the (left) $\mathbb{F}$-vector space ${}_{\mathbb{F}}V$ into a (right) $\mathbb{F}[x]$-module $V_{\mathbb{F}[x]}$. Now, since $\mathbb{F}[x]$ is a PID, we can use our powerful machinery to pin down the structure of the $\mathbb{F}[x]$-module V. Before doing so, however, we make a crucial observation concerning T-invariant subspaces of ${}_{\mathbb{F}}V$ and $\mathbb{F}[x]$-submodules of $V_{\mathbb{F}[x]}$.

(Dec2) Lemma. *A subspace W of V is T-invariant iff it is a $\mathbb{F}[x]$-submodule of $V_{\mathbb{F}[x]}$.*

Proof. $(\Leftarrow)$ Let W be an $\mathbb{F}[x]$-submodule. Then W is stable under the action of **any** element of $f(x) \in \mathbb{F}[x]$ (i.e. $Wf(x) \subseteq W$); in particular W is stable under the action of $x \in \mathbb{F}[x]$. But x acts as T ($Wx = WT \subseteq W$) so that W is T-invariant.

$(\Rightarrow)$ Let W be T-invariant. We show, by induction on $\deg(f)$, that W is also $f(T)$-invariant for any $f(x) \in \mathbb{F}[x]$. The case $\deg(f) = 0$ is trivial, so assume that $\deg(f) > 0$ and that W is $g(T)$-invariant whenever $\deg(g) < \deg(f)$. Write $f(x) = \alpha x^n + g(x)$, where $\deg(g) < n = \deg(f)$. Then, by definition of $f(T)$, we have $Wf(T) = WT^n + Wg(T) \leq WT^n$. By induction we have $Wg(T) \leq W$. Furthermore, since W is T-invariant, we have $WT^n = (WT)T^{n-1} \leq WT^{n-1}$. Now, by induction again, we have $WT^n \leq W$, and hence $Wf(T) \leq W$, as required. $\square$

A moment's thought should convince you that this little Lemma will be very useful: we are interested in finding T-invariant subspaces of V; we know now that these are just the $\mathbb{F}[x]$-submodules of $V_{\mathbb{F}[x]}$; and we know quite a good deal about the latter. Before stating our first two decomposition theorems for the linear transformation T, let us first translate an important notion from module theory into our present setting. We call a subspace W of V T-CYCLIC if it is cyclic as $\mathbb{F}[x]$-module under the action of φ_T. That is, for some $w \in W$,

$$W \;=\; w\mathbb{F}[x] \;=\; \mathrm{sp}\{wf(T) \mid f(x) \in \mathbb{F}[x]\} \;=\; w\langle T\rangle.$$

Examples.

1. Fix a positive integer n, and let $W = \{\alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n \mid \alpha_i \in \mathbb{Q}\} \leq \mathbb{Q}[x]$. Let $D \in \mathrm{End}_{\mathbb{Q}}(W)$ denote the formal derivative. Then W is D-cyclic *[let w be any polynomial of degree n and verify that $W = w\langle D\rangle$.]* Is it true that **all** D-invariant subspaces of W are D-cyclic?

2. Consider the linear transformation of $\mathbb{Q}^3$ represented, relative to the elementary basis $\mathcal{B}_e = e_1, e_2, e_3$, by the matrix
$$A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then $\mathbb{Q}^3$ is not itself A-cyclic, but $\mathbb{Q}^3 = \mathrm{sp}(e_1, e_3) \oplus \mathrm{sp}(e_2)$ is an A-invariant decomposition of $\mathbb{Q}^3$ into A-cyclic subspaces of dimensions 1 and 2 respectively.

(Dec3) Theorem. *Let $T \in \mathrm{End}_{\mathbb{F}}(V)$ and let $m_T(x) = p_1(x)^{n_1} \ldots p_k(x)^{n_k}$ be the unique factorization of the minimal polynomial $m_T(x)$ of T into monic irreducibles. Then, for each $1 \leq i \leq k$, there exists a unique sequence*

$$n_i = n_{i1} \geq n_{i2} \geq \ldots \geq n_{im_i} \geq 1$$

of natural numbers and a set $V_{i1}, V_{i2}, \ldots, V_{im_i}$ of T-cyclic subspaces of V such that

$$V = \bigoplus_{i=1}^{k} \bigoplus_{j=1}^{m_i} V_{ij},$$

and the minimal polynomial of $T_{V_{ij}}$ on V_{ij} is $p(x)^{n_{ij}}$.

Proof. By the definition of $m_T(x)$, we have $V m_T(x) = 0$. Hence, as $\mathbb{F}[x]$-module, V is torsion. By **(PID6)**, there exists a unique set $q_1(x), \ldots, q_k(x)$ of monic irreducible polynomials in $\mathbb{F}[x]$ and, for each $1 \leq i \leq k$, a unique sequence $h_i = h_{i1} \geq \ldots \geq h_{ik_i} \geq 1$ of integers such that, if $V_{ij} = \mathbb{F}[x]/(q_i(x)^{h_{ij}})$, then

$$V = \bigoplus_{i=1}^{k} \bigoplus_{j=1}^{m_i} V_{ij}.$$

Since $V_{ij} = \mathbb{F}[x]/(q_i(x)^{h_{ij}})$, it is immediate that $q_i(x)^{h_{ij}}$ is the minimal polynomial of T restricted to the T-invariant subspace V_{ij} of V.

To complete the proof, we need only show that $q(x) = q_1(x)^{h_1} \ldots q_k(x)^{h_k}$ is the minimal polynomial $m_T(x)$ of T. But, for $1 \leq i \leq k$, $V_{ij} q_i(x)^{h_i} = 0$, so that $V q(x) = 0$; it follows that $m_T(x) | q(x)$. On the other hand, $V m_T(x) = 0$ so that $V_{ij} m_T(x) = 0$ for all i, j. It follows that $q_i(x)^{h_{ij}} | m_T(x)$ for all i, j, so that $q(x) | m_T(x)$. $\square$

As suggested by **(PID6)**, the polynomials $p_i(x)^{n_{ij}}$ are called the <u>ELEMENTARY DIVISORS OF T</u>. A similar strategy leads us to the following striking analogue of **(PID7)**.

(Dec4) Theorem. *There exist unique monic polynomials $q_1(x), \ldots, q_k(x) \in \mathbb{F}[x]$ such that*

$$(q_1(x)) \leq (q_2(x)) \leq \ldots \leq (q_k(x))$$

and there exist unique T-cyclic subspaces $V_1, \ldots, V_k$ of V such that

$$V = V_1 \oplus V_2 \oplus \ldots \oplus V_k$$

where $q_i(x)$ is the minimal polynomial of T_{V_i} on V_i for $1 \leq i \leq k$, and $q_1(x) = m_T(x)$ is the minimal polynomial of T on V.

As in **(PID7)**, the polynomials $q_1(x), \ldots, q_n(x)$, are called the <u>INVARIANT FACTORS OF T</u>.

Examples. We continue with our previous examples.

1. If $f(x) \in W$ is any polynomial of degree n, then $f(x)D^{n+1} = 0$ but $f(x)D^i \neq 0$ if $i < n+1$. It follows that $m_D(x) = x^{n+1}$. Furthermore, we have seen that W is D-cyclic so that, as $\mathbb{Q}[x]$-modules under the action of D, $W \cong \mathbb{Q}[x]/(x^{n+1})$ *[it would be a good idea for you to verify this directly again to help you get a feel for what's going on.]* In this case, the elementary divisors and the invariant factors are the same, namely they are both the single polynomial x^{n+1}.

2. Here, $m_A(x) = (x-1)^2$ and $\mathbb{Q}^3 = V_1 \oplus V_2$, where $V_1 = \mathrm{sp}(e_1, e_3) \cong \mathbb{Q}[x]/(x-1)^2$ and $V_2 = \mathrm{sp}(e_2) \cong \mathbb{Q}[x]/(x-1)$. Once again the elementary divisors and the invariant factors coincide: they are $(x-1)^2$ and $(x-1)$.

Exercises.

1. For each of the following transformations of $\mathbb{Q}^3$, find the minimal polynomial, the elementary divisors, the invariant factors, and a decomposition of $\mathbb{Q}^3$ into cyclic subspaces.

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

2. If V is an $\mathbb{F}$-space of dimension n, then $T \in \mathrm{End}_{\mathbb{F}}(V)$ is *nilpotent* in case $T^m = 0$ for some $m \geq 0$. Say as much as you can about the minimal polynomial $m_T(x)$ of a nilpotent transformation T.

3. Prove that, if T is a linear transformation of rank 1, then there exists $\alpha \in \mathbb{F}$ such that $T^2 = \alpha T$.