Following the brief interlude to study diagonalisable transformations and matrices, we must now get back to the serious business of the general case. In this lecture we will reach the first major goal of the course. For a linear transformation $T \in \mathrm{End}_{\mathbb{F}}(V)$ we are trying to choose a "canonical" representative of the set

$$\mathrm{Mat}(T) \;=\; \{{}_{\mathcal{B}}A_T \mid \mathcal{B} \text{ a basis of } V\} \subset \mathbb{M}_n(\mathbb{F}).$$

Equivalently, for $A \in \mathbb{M}_n(\mathbb{F})$, we seek a representative of the similarity class

$$[A] \;=\; \{A' \in \mathbb{M}_n(\mathbb{F}) \mid A \sim A'\}$$

which has the "simplest" form. For a diagonalisible matrix A, this choice will turn out to be the obvious one, namely a diagonal matrix similar to A. We already have a good headstart on the general case, thanks largely to **(Dec3)** and **(Diag2)**. The latter tells us that each matrix is similar to a block diagonal matrix, and the former suggests that if we want to further refine these individual blocks, then we should study the action of a linear transformation T on its T-cyclic subspaces.

First a definition. Let $f(x) = x^n - \alpha_{n-1}x^{n-1} - \ldots - \alpha_1 x - \alpha_0 \in \mathbb{F}[x]$. Then the <u>COMPANION MATRIX</u> of $f(x)$, denoted $C(f)$, is defined to be

$$C(f) \;=\; \begin{pmatrix} 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 0 \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & & & & 0 & 1 \\ \alpha_0 & \alpha_1 & \alpha_2 & \alpha_3 & \ldots & \alpha_{n-1} \end{pmatrix} \;\in\; \mathbb{M}_n(\mathbb{F}).$$

Examples.

1. $C(x-\lambda) = (\lambda) \in \mathbb{M}_1(\mathbb{F})$ (since $n = 1$, we have $\alpha_0 = \alpha_{n-1} = \lambda$.

2. $C((x-\lambda)^2) = \begin{pmatrix} 0 & 1 \\ -\lambda^2 & 2\lambda \end{pmatrix}$, and $C((x-\lambda)^3) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \lambda^3 & -3\lambda^2 & 3\lambda \end{pmatrix}$.

Let $T \in \mathrm{End}_{\mathbb{F}}(V)$ and suppose that V is T-cyclic of dimension n with cyclic generator $v_0 \in V$. That is,

$$V \;=\; v_0\mathbb{F}[x] \;=\; \mathrm{sp}(v_0, v_0T, v_0T^2, \ldots).$$

Let $m_T(x) = x^k - \alpha_{k-1}x^{k-1} - \ldots - \alpha_1 x - \alpha_0$ be the minimal polynomial of T, and consider the list

$$\mathcal{B}(T; v_0) = v_0, v_0T, \ldots, v_0T^{k-1}.$$

Fix $i > 0$, and use the Division Theorem to write

$$x^i = q_i(x)m_T(x) + r_i(x)$$

with $\deg(r_i) < \deg(m_T) = n$; then $v_0T^i = v_0r_i(T) \in \mathrm{sp}(\mathcal{B}(T; v_0))$. It follows that V is spanned by $\mathcal{B}(T; v_0)$. Next suppose that there exist $\beta_0, \ldots, \beta_{n-1} \in \mathbb{F}$ such that $\beta_0v_0 + \beta_1v_0T + \ldots \beta_{n-1}v_0T^{n-1} = 0$. Put $f(x) = \beta_0 + \beta_1x + \ldots + \beta_{k-1}x^{k-1}$. Then, for each $1 \leq i \leq k$, $(v_0T^{i-1})f(x) = (v_0f(T))T^{i-1} = 0$. It follows that $Vf(x) = 0$, and hence that $m_T(x)|f(x)$. But $\deg(f) \leq k - 1 < k = \deg(m_T)$, so we must have $f(x) \equiv 0$, whence $\beta_0 = \ldots = \beta_{k-1} = 0$. We have shown that $\mathcal{B}(T; v_0)$ is, in fact, a basis for V; in particular $n = \dim(V) = k = \deg(m_T)$.

For $1 \leq i \leq n-1$, let $v_i = v_0T^i$, so that $\mathcal{B}(T; v_0) = v_0, v_1, \ldots, v_{n-1}$. Then $v_iT = v_{i+1}$ for each $0 \leq i \leq n-2$. Furthermore, since $m_T(T) = T^n - \alpha_{n-1}T^{n-1} - \ldots - \alpha_1T - \alpha_0$ is the zero transformation, it follows that

$$v_{n-1}T = v_0T^{n-1}T = v_0T^n = \sum_{i=0}^{n-1} \alpha_i v_0 T^i = \sum_{i=0}^{n-1} \alpha_i v_i.$$

We have now proved:

(Can1) Lemma. *If V is T-cyclic with generator v_0 and T has minimal polynomial m_T, then* $\deg(m_T) = \dim_{\mathbb{F}}(V)$, $\mathcal{B} = \mathcal{B}(T; v_0) = v_0, v_0T, \ldots, v_0T^{n-1}$ *is an ordered basis for V and*

$${}_{\mathcal{B}}A_T = C(m_T).$$

□

Example.

Consider the matrix A from **(Dec)**, Exercise 1. Since $m_A(x) = (x-1)^3$, by the first Example, A is similar to $C((x-1)^3) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{pmatrix}$. Can you find an invertible matrix P such that $PAP^{-1} = C((x-1)^3)$?

The previous Lemma paves the way for our first two "canonical form" theorems for linear transformations. They follow easily from **(Dec3)** and **(Dec4)** respectively.

(Can2) Theorem. [Primary Rational Canonical Form]
Let $T \in \mathrm{End}_{\mathbb{F}}(V)$ *have elementary divisors*

$$\{p_i(x)^{n_{ij}} \in \mathbb{F}[x] \mid 1 \le i \le k,\ 1 \le j \le m_i\}.$$

Then there exists an ordered basis $\mathcal{B}$ *for* V *such that*

$${}_{\mathcal{B}}A_T = \mathrm{diag}(C(p_{ij}^{n_{ij}}) \mid 1 \le i \le k,\ 1 \le j \le m_i).$$

□

The matrix ${}_{\mathcal{B}}A_T = \mathrm{diag}(C(p_{ij}^{n_{ij}}) \mid 1 \le i \le k,\ 1 \le j \le m_i)$ appearing in **(Can2)** is called the <u>PRIMARY RATIONAL CANONICAL FORM</u> (PRCF) of T. The use of the article "the" in this definition is a little clumsy; while the companion matrices $C(p_{ij}^{n_{ij}})$ that appear on the diagonal are unique, their order depends upon the choice of $\mathcal{B}$. The next canonical form does not suffer from this weakness.

(Can3) Theorem. [Rational Canonical Form]
Let $T \in \mathrm{End}_{\mathbb{F}}(V)$ *have invariant factors* $q_1(x), \ldots, q_m(x)$*. Then there exists an ordered basis* $\mathcal{B}$ *for* V *such that*

$${}_{\mathcal{B}}A_T = \mathrm{diag}(C(q_1), \ldots, C(q_m)).$$

□

The matrix ${}_{\mathcal{B}}A_T = \mathrm{diag}(C(q_1), \ldots, C(q_m))$ is called the <u>RATIONAL CANONICAL FORM</u> (RCF) of T. This time the matrix is unique because of the ordering of the blocks induced by the natural ordering of the invariant factors. It is not the case, however, that the choice of basis, $\mathcal{B}$, is unique.

Example.

Let us determine the PRCF and RCF for the matrix

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \in \mathbb{M}_3(\mathbb{Q}).$$

An easy calculation shows that the minimal polynomial of A is $m_A(x) = (x-1)^2(x-2)$ and that the elementary divisors are $(x-1)^2$ and $x-2$. There is, therefore, only one invariant factor, namely $m_A(x)$ itself. Hence,

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -5 & 2 \end{pmatrix}$$

are, respectively, the PRCF and RCF of A.

This is about is much as we can say in general but we can derive a much neater result if we specialise slightly to the case where $m_T(x)$ factors as a product of (not necessarily distinct) linear polynomials in $\mathbb{F}[x]$. While this may sound like quite a leap, you should note that this *always* the case when $\mathbb{F} = \mathbb{C}$ (or, more generally, when $\mathbb{F}$ is algebraically closed).

If $m_T(x) = (x-\lambda_1)^{n_1} \dots (x-\lambda_k)^{n_k}$ then we know from the Elementary Divisor Theorem that V decomposes as a direct sum of T-invariant subspaces V_{ij} such that the minimal polynomial of $T_{V_{ij}}$ is $(x-\lambda_i)^{n_{ij}}$. Hence, we now restrict our attention to the case when V is T-cyclic with minimal polynomial $m_T(x) = (x-\lambda)^n$.

Recall that a trnaformation N is <u>NILPOTENT</u> is there exists an integer k such that $N^k = 0$. Suppose that V, of dimension n, is N-cyclic for some nilpotent transformation N. Since $Vx^k = 0$, it follows that $m_N(x)$ divides x^k. Moreover, since $n = \dim(V) = \deg(m_T)$, it follows that n is the smallest integer k such that $N^k = 0$. That is, $m_T(x) = x^n$. We see that a nilpotent transformation acting cyclically on V is the special case $\lambda = 0$ of the problem we are considering.

Let us now consider the general case. Since V is T-cyclic, there exists $v \in V$ such that $\mathcal{B}(v;T) = v, vT, \dots, vT^{n-1}$ is a basis for V. Denoting the identity transformation of V by 1_V, set $N := T - \lambda 1_V$. Since $N^n = (T-\lambda 1_V)^n = 0$, it follows that N is nilpotent. Furthermore, by a direct computation,

$$vT^i = v(N + \lambda 1_V)^i = vN^i + \sum_{j=0}^{i-1} \gamma_j v N^j.$$

Since the vT^i are linearly independent, it follows that the vN^i are linearly independent. Hence $\mathcal{B} = \mathcal{B}(v, N) = v, vN, \ldots, vN^{n-1}$ is also a basis for V, and V is N-cyclic. We have just seen that ${}_{\mathcal{B}}A_N = C(x^n)$ so that

$$
{}_{\mathcal{B}}A_T = {}_{\mathcal{B}}A_{N+\lambda 1_V} = {}_{\mathcal{B}}A_N + {}_{\mathcal{B}}A_{\lambda 1_V} = C(x^n) + \lambda I_n = \begin{pmatrix} \lambda & 1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & 0 & \ldots & 0 \\ 0 & 0 & \lambda & 1 & \ldots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & & & 0 & \lambda & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda \end{pmatrix}.
$$

We call a matrix of this sort a <u>Jordan block</u>; this particular matrix will be denoted $J_n(\lambda)$. We can now state the result we have been after.

(Can5) Theorem. [Jordan Canonical Form] $T \in \mathrm{End}_{\mathbb{F}}(V)$ *has minimal polynomial*

$$
m_T(x) = (x - \lambda_1)^{n_1} \ldots (x - \lambda_k)^{n_k}
$$

if and only if, for each $1 \le i \le k$*, there exists a unique sequence* $n_i = n_{i1} \ge n_{i2} \ge \ldots \ge n_{im_i} \ge 1$*, and a basis* $\mathcal{B}$ *such that*

$$
{}_{\mathcal{B}}A_T = \mathrm{diag}(J_{n_{ij}}(\lambda_i) \mid 1 \le i \le k, 1 \le j \le m_i).
$$

□

The matrix ${}_{\mathcal{B}}A_T$ in **(Can5)** is called the <u>Jordan Canonical Form</u> (JCF) or <u>Jordan normal form</u> (JNF) of the linear transformation T. We reiterate that not every linear transformation has such a form unless we are dealing with a special sort of field.

(Can6) Corollary. *If* $\mathbb{F}$ *is algebraically closed (for example, if* $\mathbb{F} = \mathbb{C}$*) then every linear transformation of an* $\mathbb{F}$*-vector space* V *has a Jordan Canonical Form.* □

Example.

Let us compute the JCF for the matrix A of the previous example. Since $m_A(x) = (x-1)^2(x-2)$ and $\mathrm{sp}(e_1, e_2)$ and $\mathrm{sp}(e_3)$ are A-cyclic subspaces of dimensions 1 and 2, the JCF of A is

$$\mathrm{diag}(J_2(1), J_1(2)) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Note that A was almost in JCF already!

Exercises.

1. Let V be a $\mathbb{Q}$-space of dimension 14 and consider the $f(x) = (x^4 - x^2 - 2)^2 \in \mathbb{Q}[x]$.
   (a) Write a PRCF for each $T \in \mathrm{End}_{\mathbb{Q}}(V)$ having $m_T(x) = f(x)$.
   (b) Write an RCF for each $T \in \mathrm{End}_{\mathbb{Q}}(V)$ having $m_T(x) = f(x)$.
2. Prove that $\det(xI_n - C(f)) = f(x)$ for any monic polynomial $f(x) \in \mathbb{F}[x]$. *[Hint. Expand* $\det(xI_n - C(f))$ *along the first column and induct]*
3. Is it true that for every choice of n and every $A \in \mathbb{M}_n(\mathbb{F})$, A is similar to its transpose A^{tr}? Explain.
4. Classify up to similarity all $A \in \mathbb{M}_3(\mathbb{C})$ such that $A^3 = I_3$.