Remark 2.4.13. The mathematical reason that $\mathbb{Z}$ and $\mathbb{C}[t]$ obey the same algebraic properties is that they are both examples of *Euclidean domains*: these are commutative rings for which there exists a division algorithm, Euclidean algorithm and the notion of prime elements.

More specifically, a Euclidean domain is a commutative ring without zerodivisors for which there exists a well defined 'degree' function. As a consequence of the existence of the degree function the division algorithm and Euclidean algorithm hold. Moreover, it can be show that such commutative rings are principal ideal domains and are therefore unique factorisation domains: this means that the 'unique factorisation' property holds.

2.5 Canonical form of an endomorphism

([1], p.142-146)

Throughout this section we fix a linear endomorphism $L \in \mathrm{End}_{\mathbb{C}}(V)$, for some finite dimensional $\mathbb{C}$-vector space V. We denote $n = \dim_{\mathbb{C}} V$.

We recall the notation from Corollary 2.4.7: for $L \in \mathrm{End}_{\mathbb{C}}(V)$, $f = a_0 + a_1 t + ... + a_k t^k \in \mathbb{C}[t]$, we define the endomorphism

$$f(L) = \rho_L(f) = a_0 \mathrm{id}_V + a_1 L + a_2 L^2 + ... + a_k L^k \in \mathrm{End}_{\mathbb{C}}(V),$$

where $L^i = L \circ \cdots \circ L$ is the i-fold composition of the endomorphism L.

Definition 2.5.1. Any nonzero $f \in \ker \rho_L$ is called an *annihilating polynomial of L*.

In particular, the minimal polynomial μ_L of L is an annihilating polynomial of L.

The following theorem is the culmination of our discussion regarding polynomials and representations of the polynomial algebra. It allows us to use the minimal polynomial of L to decompose V into a direct sum of L-invariant subspaces. Hence, we can find a basis of V for which the matrix of L with respect to this basis is block diagonal. We will then see that we can use our results on nilpotent endomorphisms to find a basis of V for which the matrix of L is 'almost diagonal' - this is the **Jordan canonical form** (Theorem 2.5.12).

Theorem 2.5.2. *Suppose that $f \in \ker \rho$ is an annihilating polynomial of L and that $f = f_1 f_2$, with f_1 and f_2 relatively prime. Then, we can write*

$$V = U_1 \oplus U_2,$$

with U_1 and U_2 both L-invariant (Definition 2.2.1), and such that

$$f_1(L)(u_2) = 0_V, \quad f_2(L)(u_1) = 0_V,$$

for every $u_1 \in U_1, u_2 \in U_2$.

Moreover,

$$U_1 = \ker f_2(L), \quad U_2 = \ker f_1(L).$$

Proof: As f_1 and f_2 are relatively prime we know that there exists $g_1, g_2 \in \mathbb{C}[t]$ such that

$$f_1 g_1 + f_2 g_2 = 1 \in \mathbb{C}[t].$$

This follows from Lemma 2.4.11. Hence, we have

$$\mathrm{id}_v = \rho_L(1) = \rho_L(f_1 g_1 + f_2 g_2) = \rho_L(f_1)\rho_L(g_1) + \rho_L(f_2)\rho_L(g_2) = f_1(L)g_1(L) + f_2(L)g_2(L).$$

Define

$$U_1 = \mathrm{im} f_1(L), \quad U_2 = \mathrm{im} f_2(L).$$

Then, since

$$f_1(L) \circ L = L \circ f_1(L), \quad f_2(L) \circ L = L \circ f_2(L),$$

(you should check this) we have that, if $u_1 = f_1(L)(x_1) \in U_1, u_2 = f_2(L)(x_2) \in U_2$, then

$$L(u_1) = L \circ f_1(L)(x_1) = f_1(L) \circ L(x_1) \in \mathrm{im} f_1(L) = U_1, \ L(u_2) = L \circ f_2(L)(x_2) = f_2(L) \circ L(x_2) \in \mathrm{im} f_2(L) = U_2.$$

Hence, U_1, U_2 are L-invariant.

Now, let $u_1 \in U_1 = \mathrm{im} f_1(L)$ so that $u_1 = f_1(L)(x_1)$, for some $x_1 \in V$. Then,

$$f_2(L)(u_1) = f_2(L)(f_1(L)(x_1)) = f(L)(x_1),$$

since $f_1 f_2 = f$ and $\rho_L(f) = \rho_L(f_1 f_2) = \rho_L(f_1)\rho_L(f_2)$ (ρ_L is a representation of $\mathbb{C}[t]$). Our assumption is that f is an annhiliating polynomial of L so that $f(L) = 0_{\mathrm{End}_{\mathbb{C}}(V)}$ and we obtain

$$f_2(L)(u_1) = f(L)(x_1) = 0_V.$$

Similarly we obtain that

$$f_1(L)(u_2) = 0_V, \quad \text{for every } u_2 \in U_2.$$

Let $v \in V$. Then,

$$v = \mathrm{id}_V(v) = (f_1(L)g_1(L) + f_2(L)g_2(L))(v) = f_1(L)(g_1(L)(v)) + f_2(L)(g_2(L)(v)) \in U_1 + U_2.$$

Hence, $V = U_1 + U_2$.

Now, let $x \in U_1 \cap U_2$. Therefore, we have $f_1(L)(x) = 0_V = f_2(L)(x)$ by what we showed above. Hence,

$$x = f_1(L)(g_1(L)(x)) + f_2(L)(g_2(L)(x)) = g_1(L)(f_1(L)(x)) + g_2(L)(f_2(L)(x)) = g_1(L)(0_V) + g_2(L)(0_V) = 0_V.$$

Here we have used that $h(L) \circ g(L) = g(L) \circ h(L)$, for any $g, h \in \mathbb{C}[t]$, which can be easily verified.

Hence, we have

$$V = U_1 \oplus U_2,$$

Finally, suppose that $f_2(L)(w) = 0_V$, for some $w \in V$. Then, we want to show that $w \in U_1$. Since $V = U_1 \oplus U_2$ then we have

$$w = u_1 + u_2,$$

where $u_1 \in U_1, u_2 \in U_2$. Thus, we have $x_1, x_2 \in V$ such that

$$u_1 = f_1(L)(x_1), \quad x_2 = f_2(L)(x_2).$$

Thus,

$$0_V = f_2(L)(w) = f_2(L)(u_1 + u_2) = f_2(L)(u_1) + f_2(L)(u_2) = 0_V + f_2(L)(u_2),$$

and

$$f_1(L)(u_2) = 0,$$

as $u_2 \in U_2$. Therefore,

$$u_2 = g_1(L)(f_1(L)(u_2)) + g_2(L)(f_2(L)(u_2)) = 0_V + 0_V = 0_V,$$

so that $w = u_1 \in U_1$. We obtain that $\ker f_1(L) = U_2$ similarly. □

Corollary 2.5.3 (Primary Decomposition Theorem)**.** *Let $f \in \mathbb{C}[t]$ be an annihilating polynomial of $L \in \mathrm{End}_{\mathbb{C}}(V)$. Suppose that f is decomposed into the following linear factors:*[50]

$$f = a(t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2} \cdots (t - \lambda_k)^{n_k}.$$

Then, there are L-invariant subspaces $U_1, \dots, U_k \subset V$ such that

$$V = U_1 \oplus \dots \oplus U_k,$$

and such that each U_i is annihilated by the endomorphism

$$(L - \lambda_i \mathrm{id}_V)^{n_i} = (L - \lambda_i \mathrm{id}_V) \circ \cdots \circ (L - \lambda_i \mathrm{id}_V).$$

[50] This is always possible by the Fundamental Theorem of Algebra.

Proof: This is a direct consequence of Theorem 2.5.2: apply Theorem 2.5.2 to

$$f_1 = (t-\lambda_1)^{n_1}, \; g_1 = (t-\lambda_2)^{n_2}\cdots(t-\lambda_k)^{n_k},$$

which are obviously relatively prime polynomials, to obtain

$$V = U_1 \oplus V_1,$$

where $U_1 = \ker f_1(L), V_1 = \ker g_1(L)$. Then, V_1 is L-invariant so that L restricts to a well-defined endomorphism of V_1, denoted $L_1 \in \text{End}_{\mathbb{C}}(V_1)$. Then, g_1 is an annihilating polynomial of L_1.

Now, we can write

$$g_1 = f_2 g_2,$$

where

$$f_2 = (t-\lambda_2)^{n_2}, \; g_2 = (t-\lambda_3)^{n_3}\cdots(t-\lambda_k)^{n_k}.$$

Then, f_2 and g_2 are relatively prime so we can apply Theorem 2.5.2 to V_1 to obtain

$$V_1 = U_2 \oplus V_2.$$

with $U_2 = \ker f_2(L), V_2 = \ker g_2(L)$. Then, V_2 is L_1-invariant (and also L-invariant, when we consider V_2 as a subspace of V) so that L_1 restricts to a well-defined endomorphism of V_2, denoted $L_2 \in \text{End}_{\mathbb{C}}(V_2)$. Then, g_2 is an annihilating polynomial of L_2.

Proceeding in this way we see that we can write

$$V = U_1 \oplus \cdots \oplus U_k,$$

where $U_i = \ker(L - \lambda_i \text{id}_v)^{n_i}$. $\square$

Remark 2.5.4. Theorem 2.5.2 and the Primary Decomposition Theorem (Corollary 2.5.3) form the theoretical basis for the study of endomorphisms of a finite dimensional $\mathbb{C}$-vector space. These results allow us to deduce many properties of an endomorphism L if we know its minimal polynomial (or its characteristic polynomial). The next few Corollaries demonstrate this.

Corollary 2.5.5. *Let $L \in \text{End}_{\mathbb{C}}(V)$. Then, L is diagonalisable if and only if μ_L is a product of* <u>*distinct*</u> *linear factors, ie,*

$$\mu_L = (t-c_1)(t-c_2)\cdots(t-c_k),$$

with $c_i \neq c_j$ for $i \neq j$.

Proof: ($\Rightarrow$) Suppose that L is diagonalisable so that we have

$$E^L_{\lambda_1} \oplus \cdots \oplus E^L_{\lambda_k} = V,$$

with $E^L_{\lambda_i}$ the λ_i-eigenspace of L. Consider the polynomial

$$f = (t-\lambda_1)\cdots(t-\lambda_k) \in \mathbb{C}[t].$$

Then, we claim that $\rho_L(f) = 0 \in \text{End}_{\mathbb{C}}(V)$: indeed, let $v \in V$ and write $v = e_1 + ... + e_k$ with $e_i \in E^L_{\lambda_i}$. Then, for each i, we have

$$\rho_L(f)(e_i) = (L - \lambda_1 \text{id}_V)\cdots(L-\lambda_k \text{id}_V)(e_i) = 0_V,$$

because $(L-\lambda_s \text{id}_V)(L-\lambda_t \text{id}_V) = (L-\lambda_t \text{id}_V)(L-\lambda_s \text{id}_V)$, for every s, t.[51] Hence, we must have $\rho_L(f)(v) = 0_V$, for every $v \in V$, so that $\rho_L(f) = 0 \in \text{End}_{\mathbb{C}}(V)$. Hence, by Proposition 2.4.4, there is some $g \in \mathbb{C}[t]$ such that

$$f = \mu_L g.$$

As f is a product of distinct linear factors the same must be true of μ_L.

[51]We can move $(L - \lambda_i \text{id}_V)$ to the front of $\rho_L(f)$ and, since $L(e_i) = \lambda_i e_i$, we obtain $(L-\lambda_i \text{id}_V)(e_i) = 0_V$.

($\Leftarrow$) Suppose that
$$\mu_L = (t - c_1) \cdots (t - c_k) \in \mathbb{C}[t].$$
Then, by Corollary 2.5.3, we obtain a direct sum decomposition
$$V = U_1 \oplus \cdots \oplus U_k,$$
where $U_i = \ker(L - c_i \mathrm{id}_V)$. Hence,
$$U_i = \{v \in V \mid (L - c_i \mathrm{id}_V)(v) = 0_V\} = \{v \in V \mid L(v) = c_i v\} = E^L_{c_i}$$
is precisely the c_i-eigenspace of L. Thus, as we have written V as a direct sum of eigenspaces of L we must have that L is diagonalisable. $\square$

Example 2.5.6. 1. Let $A \in Mat_n(\mathbb{C})$ be such that
$$A^k - I_n = 0_n,$$
for some $k \in \mathbb{N}$. Then, we see that
$$f = t^k - 1 \in \ker \rho_A,$$
where $\rho_A = \rho_{T_A}$ is the representation of $\mathbb{C}[t]$ defined by the endomorphism $T_A \in \mathrm{End}_{\mathbb{C}}(\mathbb{C}^n)$. Therefore, the minimal polynomial of A, μ_A, must divide f so that there is $g \in \mathbb{C}[t]$ such that
$$f = \mu_A g.$$
Now, we have
$$f = (t - 1)(t - \omega) \cdots (t - \omega^{k-1}),$$
where $\omega = \cos(2\pi/k) + \sin(2\pi/k)\sqrt{-1}$; in particular, f has distinct linear factors. Thus, the same must be true of μ_A. Hence, by Corollary 2.5.5 we have that A is diagonalisable.

For those of you that are taking Math 113 this has an important consequence:

'every commutative finite group can be realised as a subgroup of D_n, for some n'

where D_n is the group of diagonal $n \times n$ complex matrices. This uses Cayley's theorem (for groups) and the fact that a family of commuting diagonalisable matrices can be simultaneously diagonalised (mentioned as a footnote on LH3).

2. More generally, $A \in Mat_n(\mathbb{C})$ is such that there exists a polynomial relation
$$0 = f(A) = \rho_A(f),$$
for some $f \in \mathbb{C}[t]$ with distinct linear factors, then A is diagonalisable. For example, if
$$A^2 - 3A + 2I_n = 0_n,$$
then A is diagonalisable.

The previous Corollary shows that the zeros of the minimal polynomial are eigenvalues of L, for L diagonalisable. In fact, **this is true for any $L \in \mathrm{End}_{\mathbb{C}}(V)$.**

Corollary 2.5.7. *Let $L \in \mathrm{End}_{\mathbb{C}}(V)$ and $\mu_L \in \mathbb{C}[t]$ the minimal polynomial of L. Then, $\mu_L(c) = 0$ if and only if $c \in \mathbb{C}$ is an eigenvalue of L.*

Proof: Suppose that

$$\mu_L = (t - c_1)^{n_1} \cdots (t - c_k)^{n_k}.$$

Then, $\mu_L(c) = 0$ if and only if $c = c_i$, for some $i \in \{1, \dots, k\}$. We will show that each c_i is an eigenvalue of L and, conversely, if λ is an eigenvalue of L then $\lambda = c_i$, for some i. This shows that the set of eigenvalues of L is precisely $\{c_1, \dots, c_k\}$.

Let $U_1, \dots, U_k \subset V$ be the L-invariant subspaces such that

$$V = U_1 \oplus \cdots \oplus U_k,$$

from Corollary 2.5.3. Then, the proof of Corollary 2.5.3 shows that $U_i = \ker(L - c_i \mathrm{id}_V)^{n_i}$. As $n_i \geq 1$ we can find nonzero $w \in V$ such that $(L - c_i \mathrm{id}_V)(v) = 0_V$: namely, we take

$$w = (L - c_i \mathrm{id}_V)^{r-1}(v),$$

where $r = \mathrm{ht}(v)$ is equal to the height of any nonzero $v \in U_i$ with respect to the nilpotent endomorphism $(L_{|U_i} - c_i \mathrm{id}_{U_i}) \in \mathrm{End}_{\mathbb{C}}(U_i)$.[52] Hence,

$$(L - c_i \mathrm{id}_V)(w) = (L - c_i \mathrm{id}_V)^r(v) = 0_V,$$

so that w is eigenvector of L with associated eigenvalue c_i. In particular, c_i is an eigenvalue of L.

Conversely, suppose that $c \in \mathbb{C}$ is an eigenvalue of L and that v is an eigenvector such that $L(v) = cv$; in particular, $v \neq 0_V$. Then, since

$$V = U_1 \oplus \cdots \oplus U_k,$$

we have a unique expression

$$v = u_1 + \dots + u_k, \ u_i \in U_i.$$

Then,

$$L(u_1) + \dots + L(u_k) = L(v) = cv = cu_1 + \dots + cu_k,$$

and since $L(u_i) \in U_i$ (each U_i is L-invariant) we must have $L(u_i) = cu_i$, for each i: this follows because every $z \in V$ can be written as a <u>unique</u> linear combination of vectors in $U_1, \dots, U_k$.

Let $\Gamma_1 = \{i \in \{1, \dots, k\} \mid u_i = 0_V\}$ and $\Gamma_2 = \{1, \dots, k\} \setminus \Gamma_1$: as $v \neq 0_V$ we must have $\Gamma_2 \neq \varnothing$. Thus, for every $i \in \Gamma_2$ we have that $u_i \in U_i$ is also an eigenvector of L with associated eigenvalue c. As

$$U_i = \ker(L - c_i \mathrm{id}_V)^{n_i},$$

we have, for each $i \in \Gamma_2$,

$$0_V = (L - c_i \mathrm{id}_V)^{n_i}(u_i) = \left(\sum_{p=0}^{n_i} \binom{n_i}{p} (-c_i)^p L^{n-p} \right)(u_i) = \sum_{p=0}^{n_i} \binom{n_i}{p} (-c_i)^p c^{n-p} u_i = (c - c_i)^{n_i} u_i.$$

Hence, we see that $c = c_i$, for each $i \in \Gamma_2$. Since $c_i \neq c_j$, if $i \neq j$, then we must have that $c = c_j$, for some j, so that any eigenvalue of L is equal to some c_j.

We have just shown that the set of eigenvalues of L is precisely $\{c_1, \dots, c_k\}$. Moreover, the set of roots of μ_L is also equal to this set and the result follows. $\square$

Corollary 2.5.8. *Let $L \in \mathrm{End}_{\mathbb{C}}(V)$ and $\mu_L \in \mathbb{C}[t]$ the minimal polynomial of L. Suppose that*

$$V = U_1 \oplus \cdots \oplus U_k,$$

is the direct sum decomposition from Corollary 2.5.3. Then, if c is an eigenvalue of L we must have that the c-eigenspace of L satisfies

$$E_c^L \subset U_j,$$

for some j. Furthermore, if c, c' are eigenvalues of L and $E_c^L, E_{c'}^L \subset U_j$, then $c = c'$.

[52] This is an endomorphism of U_i since U_i is L-invariant.

Proof: This follows from the latter part of the the previous proof of Corollary 2.5.7: if $v \in E^L_c$ is nonzero, so that $L(v) = cv$, then we have

$$v = u_1 + ... + u_k,\ u_i \in U_i,$$

as above. Moreover, if we define Γ_2 as before, then the latter part of the previous proof shows that $\Gamma_2 = \{j\}$, for some j. Thus,

$$v = u_j \in U_j.$$

Hence, $E^L_c \subset U_j$, for some j. The last statement follow from the proof of Corollary 2.5.7. □

Corollary 2.5.9 (Cayley-Hamilton Theorem). *Let $L \in \mathrm{End}_{\mathbb{C}}(V)$ and $\chi_L \in \mathbb{C}[t]$ the characteristic polynomial of L. Then,*

$$\chi_L(L) = \rho_L(\chi_L) = 0_{\mathrm{End}_{\mathbb{C}}(V)} \in \mathrm{End}_{\mathbb{C}}(V).$$

Proof: This is a consequence of Corollary 2.5.7. The roots of the minimal polynomial of L, μ_L, are precisely the eigenvalues of L. The roots of χ_L are also the eigenvalues of L. Therefore, we see that

$$\mu_L = (t-\lambda_1)^{m_1}\cdots(t-\lambda_k)^{m_k}, \text{ and } \chi_L = (t-\lambda_1)^{n_1}\cdots(t-\lambda_k)^{n_k}.$$

We are going to show that $m_i \leq n_i$, for each i. First we need the following Lemma (which can be easily proved by induction on k and expanding the determinant across the top row)

Lemma 2.5.10. *Let $A \in Mat_n(\mathbb{C})$ and suppose that*

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix},$$

with $A_i \in Mat_k(\mathbb{C}), A_2 \in Mat_{n-k}(\mathbb{C})$. Then, $\chi_A(\lambda) = \chi_{A_1}(\lambda)\chi_{A_2}(\lambda)$

If $\mathcal{B} = \mathcal{B}_1 \cup ... \cup \mathcal{B}_k$ is a basis of V, with each $\mathcal{B}_i \subset U_i$, then the matrix $[L]_{\mathcal{B}}$ is block diagonal

$$[L]_{\mathcal{B}} = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{bmatrix}.$$

As a consequence of Lemma 2.5.10 we have that

$$\chi_L = \chi_{A_1}\chi_{A_2}\cdots\chi_{A_k}.$$

Moreover, it follows from the proof of Corollary 2.5.7 and Corollary 2.5.8 that the only eigenvalue of A_i is λ_i. Hence, using Lemma 2.5.10 we must have that

$$\chi_{A_i} = (t-\lambda_i)^{n_i}.$$

It is a further consequence of Lemma 2.5.10 that $\dim U_i = n_i$.

Since the endomorphism $N_i = L_{|U_i} - \lambda_i \mathrm{id}_{U_i} \in \mathrm{End}_{\mathbb{C}}(U_i)$ is nilpotent (Corollary 2.5.3) the structure theorem for nilpotent endomorphisms (Theorem 2.3.4) shows that $\eta(N_i) \leq n_i$, where $\eta(N_i)$ is the exponent of N_i.

By construction, we have that

$$U_i = \ker(L - \lambda_i \mathrm{id}_V)^{m_i},$$

which implies that $\eta(N_i) \leq m_i$. In fact, $\eta(N_i) = m_i$, for every i: otherwise, we must have $\eta(N_i) < m_i$, for some i, so that for every $u \in U_i$,

$$(L - \lambda \mathrm{id}_V)^{\eta(N_i)}(u) = 0_V.$$

Consider the polynomial

$$g = (t-\lambda_1)^{m_1}\cdots(t-\lambda_{i-1})^{m_{i-1}}(t-\lambda_i)^{\eta(N_i)}(t-\lambda_{i+1})^{m_{i+1}}\cdots(t-\lambda_k)^{m_k} \in \mathbb{C}[t].$$

We have that $\deg g < \deg \mu_L$ as $\eta(N_i) < m_i$. Then, for any $v \in V$, if we write $v = u_1 + ... + u_k$, then we see that

$$\begin{aligned}\rho_L(g)(v) &= \rho_L(g)(u_1 + ... + u_k)\\ &= \rho_L(g)(u_1) + ... + \rho_L(g)(u_k)\\ &= 0_V + ... + 0_V = 0_V,\end{aligned}$$

because

$$(L - \lambda_j \mathrm{id}_V)^{m_j}(u_j) = 0_V, \quad \text{for } j \neq i, \quad \text{and} \quad (L - \lambda_i \mathrm{id}_V)^{\eta(N_i)}(u_i) = 0_V.$$

But then this contradicts the definition of μ_L being a nonzero element of $\ker \rho_L$ of minimal degree. Hence, our initial assumption the $\eta(N_i) < m_i$, for some i, cannot hold so that $\eta(N_i) = m_i$, for every i.

Therefore, $m_i \leq n_i$, for every i, so that μ_L divides χ_L: there exists $f \in \mathbb{C}[t]$ such that

$$\chi_L = \mu_L f \in \mathbb{C}[t].$$

Hence, we obtain

$$\rho_L(\chi_L) = \rho_L(\mu_L f) = \rho_L(\mu_L)\rho_L(f) = 0_{\mathrm{End}_{\mathbb{C}}(V)} \in \mathrm{End}_{\mathbb{C}}(V),$$

where we use that $\mu_L \in \ker \rho_L$. $\square$

Remark 2.5.11. The Cayley-Hamilton theorem is important as it gives us an upper bound on the degree of the minimal polynomial: we know that the minimal polynomial of L must have degree at most n^2 (because the set $\{\mathrm{id}_v, L, ..., L^{n^2}\} \subset \mathrm{End}_{\mathbb{C}}(V)$ must be linearly dependent), so that $\deg \mu_L \leq n^2$. However, the Cayley-Hamilton theorem says that we actually have $\deg \mu_L \leq n$ thereby limiting the possibilities for μ_L.

2.5.1 The Jordan canonical form

Let us denote

$$N_i = L_{|U_i} - \lambda_i \mathrm{id}_{U_i} \in \mathrm{End}_{\mathbb{C}}(U_i).$$

Since each U_i is L-invariant it is also N_i-invariant (Lemma 2.2.3). Moreover, Corollary 2.5.3 implies that the restriction of N_i to U_i is a nilpotent endomorphism of U_i. Hence, by Theorem 2.3.4, we can find a basis $\mathcal{B}_i \subset U_i$ of U_i such that the matrix of the restriction of N_i with respect to $\mathcal{B}_i$ has the canonical form

$$\begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & J_{p_i} \end{bmatrix},$$

with each J_a a 0-Jordan block and such that the size of J_i is at least as large as the size of J_{i+1}. Let $\mathcal{B} = \mathcal{B}_1 \cup ... \cup \mathcal{B}_k$ be the subsequent ordered basis of V we obtain.

As we have

$$V = U_1 \oplus \cdots \oplus U_k,$$

then for each $v \in V$, we have

$$v = u_1 + ... + u_k, \; u_i \in U_i.$$

Thus, applying L to v gives

$$L(v) = L(u_1) + ... + L(u_k) = \lambda_1 u_1 + N_1(u_1) + ... + \lambda_k u_k + N_k(u_k).$$

Hence, the matrix of L with respect to the basis $\mathcal{B}$ takes the form

$$[L]_{\mathcal{B}} = \begin{bmatrix} A_1 & 0 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & A_k \end{bmatrix},$$

where, for each $i = 1, \ldots, k$, we have

$$A_i = \lambda_i I_{\dim U_i} + \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & J_{p_i} \end{bmatrix}$$

$$= \begin{bmatrix} \begin{matrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_i & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_i \end{matrix} & & & \\ & \ddots & & \\ & & \begin{matrix} \lambda_i & 1 & \cdots & 0 \\ 0 & \lambda_i & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & 1 \\ 0 & \cdots & 0 & \lambda_i \end{matrix} & \\ & & & \ddots \\ & & & & \lambda_i \end{bmatrix} \tag{2.5.1}$$

Theorem 2.5.12 (Jordan Canonical Form)**.** *Let $L \in \mathrm{End}_{\mathbb{C}}(V)$, V a finite dimensional $\mathbb{C}$-vector space. Then, there exists an ordered basis $\mathcal{B} \subset V$ such that $[L]_{\mathcal{B}}$ is a matrix of the form 2.5.1 above. We call $\mathcal{B}$ a* Jordan basis *of L.*

Proof: Since the minimal polynomial μ_L of L is an annihilating polynomial of L we can use Primary Decomposition (Corollary 2.5.3) to obtain a direct sum decomposition of V,

$$V = U_1 \oplus \ldots \oplus U_k.$$

Now, the previous discussion implies the existence of $\mathcal{B}$ so that $[L]_{\mathcal{B}}$ takes the desired form. □

Corollary 2.5.13. *Let $A \in Mat_n(\mathbb{C})$. Then, A is similar to a matrix of the form 2.5.1 above.*

Proof: Consider the endomorphism $T_A \in \mathrm{End}_{\mathbb{C}^n}$. Then, there is an ordered basis $\mathcal{B}$ of $\mathbb{C}^n$ such that $[T_A]_{\mathcal{B}}$ takes the desired form, by Theorem 2.5.12. Since $[T_A]_{\mathcal{S}^{(n)}} = A$, we have that A and $[T_A]_{\mathcal{B}}$ are similar (Corollary 1.7.7). □

Remark 2.5.14. 1. The Jordan canonical form is a remarkable result. However, practically it is quite difficult to determine the Jordan basis of L. The use of the Jordan canonical form is mostly in theoretical applications where you are (perhaps) only concerned with knowing what the matrix of an endomorphism looks like with respect to some basis of V. The fact that a Jordan basis exists allows us to consider only 'almost diagonal' matrices, for which it can be quite easy to show that certain properties hold true.

2. The Jordan canonical form allows us to classify *similarity classes* of matrices: a similarity class is the set of all matrices which are similar to a particular matrix. Since similiarity is an equivalence relation we can partition $Mat_n(\mathbb{C})$ into disjoint similarity classes. Then, the Jordan canonical form tells us that each similarity class is labelled by a set of eigenvalues (the entries on the diagonal of the Jordan form lying in that similarity class) and the partitions of each block. Two matrices are similar if and only if these pieces of data are equal.

3. In group-theoretic language, we see that the Jordan canonical form allows us to classify the orbits of $\mathrm{GL}_n(\mathbb{C})$ acting on the set $Mat_n(\mathbb{C})$. Furthermore, this is actually the same thing as classifying the Ad-orbits of the algebraic group $\mathrm{GL}_n(\mathbb{C})$ acting on its Lie algebra $\mathfrak{gl}_n(\mathbb{C})$ via the Adjoint representation.

Example 2.5.15. Consider the following matrix

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 3 \\ -5 & -1 & -4 \end{bmatrix}.$$

Then, you can check that

$$\chi_A(t) = -(t-2)^2(t+3).$$

Since

$$A^2 + A - 6I_3 \neq 0_3,$$

it is not possible for A to be diagonalisable as this is the only possibility for the minimal polynomial μ_A with distinct linear factors.

Therfore, it must be the case that there exists $P \in \mathrm{GL}_3(\mathbb{C})$ such that

$$P^{-1}AP = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix},$$

as this is the only possibility for the Jordan canonical form of A. Let's determine a basis $\mathcal{B} \subset \mathbb{C}^3$ such that

$$P_{\mathcal{B}\leftarrow\mathcal{S}^{(3)}}[T_A]_{\mathcal{S}^{(3)}}P_{\mathcal{S}^{(3)}\leftarrow\mathcal{B}} = [T_A]_{\mathcal{B}} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix}.$$

As

$$\mu_A = (t-2)^2(t+3),$$

is an annihilating polynomial of A and $f_1 = (t-2)^2, f_2 = (t+3)$ are relatively prime, then we can find A-invariant subspaces $U_1, U_2 \subset \mathbb{C}^3$ such that

$$\mathbb{C}^3 = U_1 \oplus U_2,$$

and where

$$U_1 = \ker T_{(A-2I_3)^2},\ U_2 = \ker T_{A+3I_3}.$$

You can check that

$$U_2 = E_{-3} = \mathrm{span}_{\mathbb{C}}\left\{\begin{bmatrix} -5/28 \\ -13/28 \\ 1 \end{bmatrix}\right\},$$

so that A defines an endomorphism $T_2 : U_2 \to U_2\ ;\ \underline{x} \mapsto A\underline{x}$ of U_2 and if $\mathcal{B}_2 = \left(\begin{bmatrix} -5/28 \\ -13/28 \\ 1 \end{bmatrix}\right) \subset U_2$ then

$$[T_2]_{\mathcal{B}_2} = [-3].$$

We also know that A defines and endomorphism $T_1 : U_1 \to U_1\ ;\ \underline{x} \mapsto A\underline{x}$. Now, since

$$(A-2I_3)^2 = \begin{bmatrix} 0 & 1 & 1 \\ 2 & 1 & 3 \\ -5 & -1 & -6 \end{bmatrix},$$

we find that

$$U_1 = \ker T_{(A-2I_3)^2} = \mathrm{span}_{\mathbb{C}}\left\{\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}.$$

So, if we let

$$\mathcal{C}_1 = \left(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) (= (c_1, c_2)),$$

then

$$[T_2]_{\mathcal{C}_1} = \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}.$$

If we set

$$N_1 = [T_2]_{\mathcal{C}_1} - 2I_2 = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix},$$

then we see that $N_1^2 = 0_2$, so that N_1 is nilpotent. Moreover, using our results on nilpotent matrices, if we set $P = [N_1 e_2\ e_2]$ then we have

$$P^{-1} N_1 P = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Hence, we have

$$[T_1]_{\mathcal{B}} = N_1 + 2I_2 = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$

Therefore, if we let

$$\mathcal{B}_1 = (c_1 + c_2, c_2) = \left(\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right),$$

and $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ then we have

$$[T_A]_{\mathcal{B}} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix}.$$

In particular, if we set

$$P = \begin{bmatrix} 1 & 0 & -5/28 \\ -1 & 1 & -13/28 \\ -1 & 0 & 1 \end{bmatrix},$$

then

$$P^{-1} A P = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix}.$$