2 Jordan Canonical Form

In this chapter we are going to classify all $\mathbb{C}$-linear endomorphisms of a n-dimensional $\mathbb{C}$-vector space V. This means that we are going to be primarily studying $\text{End}_{\mathbb{C}}(V)$, the $\mathbb{C}$-vector space of $\mathbb{C}$-endomorphisms of V (up to conjugation). For those of you that know about such things, we are going to identify the orbits of the group $\text{GL}_{\mathbb{C}}(V)$ acting on the set $\text{End}_{\mathbb{C}}(V)$ by conjugation. Since there exists an isomorphism

$$\text{End}_{\mathbb{C}}(V) \rightarrow Mat_n(\mathbb{C}) \ ; \ f \mapsto [f]_{\mathcal{B}},$$

(once we choose an ordered basis $\mathcal{B}$ of V) this is the same thing as trying to classify all $n \times n$ matrices with $\mathbb{C}$-entries up to similarity.

You may recall that given any square matrix A with $\mathbb{C}$-entries we can ask whether A is *diagonalisable* and that there exists matrices that are not diagonalisable. For example, the matrix

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

is not diagonalisable.[36]

In fact, this example is typical, in the following sense: let $A \in Mat_2(\mathbb{C})$. Then, A is similar to one of the following types of matrices

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \ a, b \in \mathbb{C}, \quad \text{or} \quad \begin{bmatrix} c & 1 \\ 0 & c \end{bmatrix}, \ c \in \mathbb{C}.$$

In general, we have the following

Theorem (Jordan Canonical Form)**.** *Let $A \in Mat_n(\mathbb{C})$. Then, there exists $P \in \text{GL}_n(\mathbb{C})$ such that*

$$P^{-1}AP = \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & J_k \end{bmatrix},$$

where, for each $i = 1, \ldots, k$, we have an $n_i \times n_i$ matrix

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_i & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_i \end{bmatrix}, \quad \lambda_i \in \mathbb{C}.$$

Hence, every $n \times n$ matrix with $\mathbb{C}$-entries is similar to an *almost-diagonal* matrix.

We assume throughout this chapter that we are working with $\mathbb{C}$-vector spaces and $\mathbb{C}$-linear morphisms. Furthermore, we assume that all matrices have $\mathbb{C}$-entries.

2.1 Eigenthings

([1], p.108-113)

This section should be a refresher on the notions of *eigenvectors, eigenvalues* and *eigenspaces* of an $n \times n$ matrix A (equivalently, of a $\mathbb{C}$-linear morphism $f \in \text{End}_{\mathbb{C}}(V)$).

Definition 2.1.1. Let $f \in \text{End}_{\mathbb{C}}(V)$ be a $\mathbb{C}$-linear endomorphism of the $\mathbb{C}$-vector space V. Let $\lambda \in \mathbb{C}$.

[36] Try and recall why this was true.

- The λ-*eigenspace of* f is the set

$$E_\lambda \stackrel{def}{=} \{v \in V \mid f(v) = \lambda v\}.$$

  This is a vector subspace of V (possibly the zero subspace).

  If $E_\lambda \neq \{0_V\}$ and $v \in E_\lambda$ is a nonzero vector, then we say that v is an *eigenvector of* f *with associated eigenvalue* λ.

- If A is an $n \times n$ matrix with $\mathbb{C}$-entries then we define the λ-*eigenspace of* A to be the λ-eigenspace of the linear morphism T_A. Similarly, we say that $v \in \mathbb{C}^n$ is an *eigenvector of* A *with associated eigenvalue* λ if v is an eigenvector of T_A with associated eigenvalue λ.

Lemma 2.1.2. *Let* $f \in \mathrm{End}_{\mathbb{C}}(V)$, $v_1, \dots, v_k \in V$ *be eigenvectors of* f *with associated eigenvalues* $\lambda_1, \dots, \lambda_k \in \mathbb{C}$. *Assume that* $\lambda_i \neq \lambda_j$ *whenever* $i \neq j$. *Then,* $\{v_1, \dots, v_k\}$ *is linearly independent.*

Proof: Let $S = \{v_1, \dots, v_k\}$. Let $T \subset S$ denote a maximal linearly independent subset (we know that a linearly independent subset exists, just take $\{v_1\}$; then choose a linearly independent subset of largest size). We want to show that $T = S$. Suppose that $T \neq S$, we aim to provide a contradiction. As $T \neq S$, then we can assume, without loss of generality, that $v_k \notin T$.

We are going to show that $v_k \notin \mathrm{span}_{\mathbb{C}} T$, and then use Corollary 1.3.5 to deduce that $T \cup \{v_k\}$ is linearly independent, contradicting the maximality of T.

Suppose that $v_k \in \mathrm{span}_{\mathbb{C}} T$, we aim to provide a contradiction. So, as $v_k \in \mathrm{span}_{\mathbb{C}} T$ then

$$v_k = c_1 v_{i_1} + \dots + c_s v_{i_s},$$

where $c_1, \dots, c_s \in \mathbb{C}$, $v_{i_1}, \dots, v_{i_s} \in T$. Apply f to both sides of this equation to obtain

$$\lambda_k v_k = c_1 \lambda_{i_1} v_{i_1} + \dots + c_s \lambda_{i_s} v_{i_s}.$$

Taking this equation away from λ_k times the previous equation gives

$$0_V = c_1(\lambda_{i_1} - \lambda_k)v_{i_1} + \dots + c_s(\lambda_{i_s} - \lambda_k)v_{i_s}.$$

This is a linear relation among vectors in T so must be the trivial linear relation since T is linearly independent. Hence, we have, for each $j = 1, \dots, s$,

$$c_j(\lambda_{i_j} - \lambda_k) = 0,$$

and as $v_k \notin T$ (by assumption) we have $\lambda_{i_j} \neq \lambda_k$. Hence, we must have that $c_j = 0$, for every j. Then, we have $v_k = 0_V$, which is absurd as v_k is an eigenvector, hence nonzero by definition.

Therefore, our initial assumption that $v_k \in \mathrm{span}_{\mathbb{C}} T$ must be false, so that $v_k \notin \mathrm{span}_{\mathbb{C}} T$. As indicated above, this implies that $T \cup \{v_k\}$ is linearly independent, which contradicts the maximality of T. Therefore, T must be equal to S (otherwise $T \neq S$ and we run into the previous 'maximality' contradiction) so that S is linearly independent. □

Corollary 2.1.3. *Let* $\lambda_1, \dots, \lambda_k$ *denote all eigenvalues of* $f \in \mathrm{End}_{\mathbb{C}}(V)$. *Then,*

$$E_{\lambda_1} + \dots + E_{\lambda_k} = E_{\lambda_1} \oplus \dots \oplus E_{\lambda_k},$$

that is, the sum of all eigenspaces is a direct sum.

Proof: Left to the reader. □

Consider the case when

$$E_{\lambda_1} \oplus \dots \oplus E_{\lambda_k} = V;$$

what does this tell us? In this case, we can find a basis of V consisting of eigenvectors of f (each λ_i-eigenspace E_{λ_i} is a subspace we can find a basis of it $\mathcal{B}_i$ say. Then, since we have in this case

$$\dim V = \dim E_{\lambda_1} + \dots + \dim E_{\lambda_k},$$

we see that

$$\mathcal{B} = \mathcal{B}_1 \cup ... \cup \mathcal{B}_k,$$

is a basis of V.[37]) If we write $\mathcal{B} = (b_1, ..., b_n)$ (where $n = \dim V$) then we see that

$$[f]_{\mathcal{B}} = [[f(b_1)]_{\mathcal{B}} \cdots [f(b_n)]_{\mathcal{B}}],$$

and since $f(b_i)$ is a scalar multiple of b_i we see that $[f]_{\mathcal{B}}$ is a diagonal matrix.

Theorem 2.1.4. *Let $f \in \mathrm{End}_{\mathbb{C}}(V)$ be such that*

$$E_{\lambda_1} \oplus ... \oplus E_{\lambda_k} = V.$$

Then, there exists a basis of V such that $[f]_{\mathcal{B}}$ is a diagonal matrix.

Corollary 2.1.5 (Diagonalisation)**.** *Let $A \in Mat_n(\mathbb{C})$ be such that there exists a basis of $\mathbb{C}^n$ consisting of eigenvectors of A. Then, there exists a matrix $P \in \mathrm{GL}_n(\mathbb{C})$ such that*

$$P^{-1}AP = D,$$

where D is a diagonal matrix. In fact, the entries on the diagonal of D are the eigenvalues of A.

Proof: Let $\mathcal{B} = (b_1, ..., b_n)$ be an ordered basis of $\mathbb{C}^n$ consisting of eigenvectors of A. Then, if $P = P_{\mathcal{S}^{(n)} \leftarrow \mathcal{B}}$ we have

$$P^{-1}AP = D,$$

by applying Corollary 1.7.7 to the morphism $T_A \in \mathrm{End}_{\mathbb{C}}(\mathbb{C}^n)$. Here we note that $[T_A]_{\mathcal{B}} = D$ is a diagonal matrix by Theorem 2.1.4. □

Definition 2.1.6. We say that an endomorphism $f \in \mathrm{End}_{\mathbb{C}}(V)$ is *diagonalisable* if there exists a basis $\mathcal{B} \subset V$ of V such that $[f]_{\mathcal{B}}$ is a diagonal matrix

We say that an $n \times n$ matrix $A \in Mat_n(\mathbb{C})$ is *diagonalisable* if T_A is diagonalisable. This is equivalent to: A is diagonalisable if and only if A is similar to a diagonal matrix (this is discussed in the following Remark).

Remark 2.1.7. Corollary 2.1.5 implies that if there exists a basis of $\mathbb{C}^n$ consisting of eigenvectors of A then A is diagonalisable. In fact, the converse is true: if A is diagonalisable and $P^{-1}AP = D$ then there is a basis of $\mathbb{C}^n$ consisting of eigenvectors of A. Indeed, if we let $\mathcal{B} = (b_1, ..., b_n)$ where b_i is the i^{th} column of P, then b_i is an eigenvector of A. Why does this hold? Since we have

$$P^{-1}AP = D = \mathrm{diag}(d_1, ..., d_n),$$

where $\mathrm{diag}(d_1, ..., d_n)$ denotes the diagonal $n \times n$ matrix with $d_1, ..., d_n$ on the diagonal, then we have

$$AP = PD.$$

Then, the i^{th} column of the matrix AP is Ab_i, so that $AP = PD$ implies that $Ab_i = d_i b_i$ (equate the columns of AP and PD). Therefore, each column of P is an eigenvector of A.

2.1.1 Characteristic polynomial, diagonalising matrices

Corollary 2.1.5 tells us conditions concerning when we can we can diagonalise a given matrix $A \in Mat_n(\mathbb{C})$ - we must find a basis of $\mathbb{C}^n$ consisting of eigenvectors of A. In order to this we need to determine how we can find *any* eigenvectors, let alone a basis consisting of eigenvectors.

Suppose that $v \in \mathbb{C}^n$ is an eigenvector of A with associated eigenvalue $\lambda \in \mathbb{C}$. This means we have

$$Av = \lambda v \implies (A - \lambda I_n)v = 0_{\mathbb{C}^n},$$

[37] Why must this be a basis?

that is, $v \in \ker T_{A-\lambda I_n}$. Conversely, any nonzero $v \in \ker T_{A-\lambda I_n}$ is an eigenvector of A with associated eigenvalue λ. Note that

$$E_\lambda = \ker T_{A-\lambda I_n}.$$

Since $T_{A-\lambda I_n} \in \mathrm{End}_{\mathbb{C}}(\mathbb{C}^n)$ we know, by Proposition 1.7.4, that injectivity of $T_{A-\lambda I_n}$ is the same thing as bijectivity. Now, bijectivity of $T_{A-\lambda I_n}$ is the same thing as determining whether the matrix $A - \lambda I_n$ is invertible (using Theorem 1.7.4). Hence,

$$\ker T_{A-\lambda I_n} \neq \{0_{\mathbb{C}^n}\} \Leftrightarrow T_{A-\lambda I_n} \text{ not bijective } \Leftrightarrow \det(A - \lambda I_n) = 0.$$

Therefore, if v is an eigenvector of A with associated eigenvalue λ then we must have that $\det(A-\lambda I_n) = 0$ and $v \in \ker T_{A-\lambda I_n}$. Moreover, if $\lambda \in \mathbb{C}$ is such that $\det(A - \lambda I_n) = 0$ then $\ker T_{A-\lambda I_n} \neq \{0_{\mathbb{C}^n}\}$ and any nonzero $v \in \ker T_{A-\lambda I_n}$ is an eigenvector of A with associated eigenvalue λ.

Definition 2.1.8 (Characteristic polynomial)**.** Let $f \in \mathrm{End}_{\mathbb{C}}(V)$. Define the *characteristic polynomial of f*, denoted $\chi_f(\lambda)$, to be the polynomial in λ with complex coefficients

$$\chi_f(\lambda) = \det([f - \lambda \mathrm{id}_V]_{\mathcal{B}}),$$

where $\mathcal{B}$ is any ordered basis of V.[38]

If $A \in Mat_n(\mathbb{C})$ then we define the *characteristic polynomial of A*, denoted $\chi_A(\lambda)$, to be $\chi_{T_A}(\lambda)$. In this case, we have (using the standard basis $\mathcal{S}^{(n)}$ of $\mathbb{C}^n$)

$$\chi_A(\lambda) = \det(A - \lambda I_n).$$

Note that we are only considering λ as a 'variable' in the determinants, not an actual number. Also, note that the degree of $\chi_f(\lambda) = \dim V$[39] and the degree of $\chi_A(\lambda) = n$.

The *characteristic equation of f (resp. A)* is the equation

$$\chi_f(\lambda) = 0, \quad (\text{resp. } \chi_A(\lambda) = 0.)$$

Example 2.1.9. Let

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

Then,

$$A - \lambda I_2 = \begin{bmatrix} 1-\lambda & -3 \\ 2 & -1-\lambda \end{bmatrix}.$$

Hence, we have

$$\chi_A(\lambda) = (1-\lambda)(-1-\lambda) - 2.(-3) = \lambda^2 + 5.$$

Remark 2.1.10.

1. It should be apparent from the discussion above that the eigenvalues of a given linear morphism $f \in \mathrm{End}_{\mathbb{C}}(V)$ (or matrix $A \in Mat_n(\mathbb{C})$) are precisely the zeros of the characteristic equation $\chi_f(\lambda) = 0$ (or $\chi_A(\lambda) = 0$).

2. Example 2.1.9 highlights an issue that can arise when we are trying to find eigenvalues of a linear morphism (or matrix). You'll notice that in this example there are **no $\mathbb{R}$-eigenvalues**: the eigenvalues are $\pm\sqrt{-5} \in \mathbb{C} \setminus \mathbb{R}$. Hence, we have complex eigenvalues that are not real. In general, given a matrix A with $\mathbb{C}$-entries (or a $\mathbb{C}$-linear morphism $f \in \mathrm{End}_{\mathbb{C}}(V)$) we will always be able to find eigenvalues - this follows from the **Fundamental Theorem of Algebra**:

[38] If $\mathcal{C}$ is any other basis of V then there is an invertible matrix P such that

$$[f - \lambda \mathrm{id}_V]_{\mathcal{C}} = P^{-1}[f - \lambda \mathrm{id}_V]_{\mathcal{B}} P.$$

Then, since $\det(AB) = \det A \det B$, for any matrices A, B, we see that $\det([f - \lambda \mathrm{id}_V]_{\mathcal{C}}) = \det([f - \lambda \mathrm{id}_V]_{\mathcal{B}})$ (where we have also used $\det P^{-1} = (\det P)^{-1}$).

[39] This will be shown in homework.

Theorem (Fundamental Theorem of Algebra)**.** *Let $p(T)$ be a nonconstant polynomial with $\mathbb{C}$-coefficients. Then, there exists $\lambda_0 \in \mathbb{C}$ such that $p(\lambda_0) = 0$. Hence, every such polynomial can be written as a product of linear factors*

$$p(T) = (T - \lambda_1)^{n_1}(T - \lambda_2)^{n_2} \cdots (T - \lambda_k)^{n_k}.$$

Note that this result is false if we wish to find a real root: for $p(T) = T^2 + 1$ there are no real roots (ie, no $\lambda_0 \in \mathbb{R}$ such that $p(\lambda_0) = 0$).

It is a consequence of this Theorem that we are considering in this section only $\mathbb{K} = \mathbb{C}$ as this guarantees that eigenvalues exist.

We are now in a position to find eigenvectors/eigenvalues of a given linear morphism $f \in \text{End}_{\mathbb{C}}(V)$ (or matrix $A \in Mat_n(\mathbb{C})$):

0. Find an ordered basis $\mathcal{B} = (b_1, \dots, b_n)$ of V to obtain $[f]_{\mathcal{B}}$. Let $A = [f]_{\mathcal{B}}$. **This step is not required if you are asked to find eigenthings for a given** $A \in Mat_n(\mathbb{C})$.
1. Determine the characteristic polynomial $\chi_A(\lambda)$ and solve the equation $\chi_A(\lambda) = 0$. The roots of this equation are the eigenvalues of A (and f), denote them $\lambda_1, \dots, \lambda_k$.
2. $v \in V$ is an eigenvector with associated eigenvalue λ_i if and only if $v \in \ker(f - \lambda_i \text{id}_V)$ if and only if $[v]_{\mathcal{B}}$ is a solution to the matrix equation

$$(A - \lambda_i I_n)\underline{x} = \underline{0}.$$

Example 2.1.11. This follows on from Example 2.1.9 and we have already determined Step 1. above, we have

$$\lambda_1 = \sqrt{-5}, \ \lambda_2 = -\sqrt{-5}.$$

If we wish to find eigenvectors with associated eigenvalue λ_1 then we consider the matrix

$$A - \lambda_1 I_2 = \begin{bmatrix} 1 - \sqrt{-5} & -3 \\ 2 & -1 - \sqrt{-5} \end{bmatrix} \sim \begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix},$$

and so obtain that

$$\ker T_{A - \lambda_1 I_2} = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{C}^2 \mid x_1 - 3x_2 = 0 \right\} = \left\{ \begin{bmatrix} 3x \\ x \end{bmatrix} \mid x \in \mathbb{C} \right\}.$$

In particular, if we choose $x = 1$, we see that $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ is an eigenvector of A with associated eigenvalue $\sqrt{-5}$. Any eigenvector of A with associated eigenvalue $\sqrt{-5}$ is a nonzero vector in $\ker T_{A - \sqrt{-5} I_2}$.

Definition 2.1.12. Let $f \in \text{End}_{\mathbb{C}}(V)$ (or $A \in Mat_n(\mathbb{C})$). Supoose that

$$\chi_f(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_k)^{n_k}$$

so that $\lambda_1, \dots, \lambda_k$ are the eigenvalues of f.

- define the *algebraic multplicity of* λ_i to be n_i,
- define the *geometric multiplicity of* λ_i to be $\dim E_{\lambda_i}$.

Lemma 2.1.13. *Let $f \in \text{End}_{\mathbb{C}}(V)$ and λ be an eigenvalue of f. Then,*

'alg. multplicity of λ' $\geq$ 'geom. multiplicity of λ'

Proof: This will be proved later after we have introduced the polynomial algebra $\mathbb{C}[t]$ and the notion of a representation of $\mathbb{C}[t]$ (Definition 2.4.2) □

Proposition 2.1.14. *Let $A \in Mat_n(\mathbb{C})$. Denote the eigenvalues of A by $\lambda_1, \dots, \lambda_k$. Then, A is diagonalisable if and only if, for every i, the algebraic multplicity of λ_i is equal to the geometric multiplicity of λ_i.*

Proof: ($\Rightarrow$) Suppose that A is diagonalisable and that

$$\chi_A(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_k)^{n_k}.$$

Then, by Remark 2.1.7, we can find a basis of eigenvectors of $\mathbb{C}^n$. Hence, we must have

$$\mathbb{C}^n = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}.$$

Then, by Lemma 2.1.13 and Corollary 1.5.19, we have

$$n = \dim E_{\lambda_1} + \dots + \dim E_{\lambda_k} \leq n_1 + \dots + n_k = n,$$

where we have used that the degree of the characteristic polynomial is n.
This implies that we must have $\dim E_{\lambda_i} = n_i$, for every i: indeed, we have

$$\dim E_{\lambda_1} + \dots \dim E_{\lambda_k} = n_1 + \dots + n_k,$$

with $\dim E_{\lambda_i} \leq n_i$, for each i. If $\dim E_{\lambda_i} < n_i$, for some i, then we would coontradict this previous equality. The result follows.

($\Leftarrow$) Assume that $\dim E_{\lambda_i} = n_i$, for every i. Then, we know that

$$V \supset E_{\lambda_1} + \dots + E_{\lambda_k} = E_{\lambda_1} \oplus \dots \oplus E_{\lambda_k}.$$

Then, since

$$\dim(E_{\lambda_1} \oplus \dots \oplus E_{\lambda_k}) = \dim E_{\lambda_1} + \dots + \dim E_{\lambda_k} = n_1 + \dots + n_k = n,$$

we see that $V = E_{\lambda_1} \oplus \dots \oplus E_{\lambda_k}$, by Corollary 1.5.17. Hence, there is a basis of V consisting of eigenvectors of A so that A is diagonalisable. □

As a consequence of Proposition 2.1.14 we are now in a position to determine (in practice) when a matrix A is diagonalisable. Following on from the above list to find eigenvectors we have

3. For each eigenvalue λ_i determine a basis of $\ker T_{A - \lambda_i I_n}$ (by row-reducing the matrix $A - \lambda_i I_n$ to reduced echelon form, for example). Denote this basis $\mathcal{B}_i = (b_1^{(i)}, \dots, b_{m_i}^{(i)})$.

4. If $|\mathcal{B}_i| = m_i = n_i$, for every i, then A is diagonalisable. Otherwise, A is not diagonalisable. Recall that in Step 1. above you will have determined $\chi_A(\lambda)$, and therefore n_i.

5. If A is diagonalisable then define the matrix P to be the $n \times n$ matrix

   $$P = [b_1^{(1)} \cdots b_{n_1}^{(1)} b_1^{(2)} \cdots b_{n_2}^{(2)} \cdots b_1^{(k)} \cdots b_{n_k}^{(k)}].$$

   Then, Remark 2.1.7 implies that

   $$P^{-1}AP = \operatorname{diag}(\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_2, \dots, \lambda_k, \dots, \lambda_k),$$

   with each eigenvalue λ_i appearing n_i times on the diagonal.

   Note that the order of the eigenvalues appearing on the diagonal depends on the ordering we put on $\mathcal{B}$.

Corollary 2.1.15. *Let $A \in Mat_n(\mathbb{C})$. Then, if A has n distinct eigenvalues $\lambda_1, \dots, \lambda_n$, then A is diagonalisable.*

Proof: Saying that A has n distinct eigenvalues is equivalent to saying that

$$\chi_A(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_n),$$

so that the algebraic multiplicity n_i of each eigenvalue is 1. Furthermore, λ_i is an eigenvalue if and only if there exists a nonzero $v \in \mathbb{C}^n$ such that $Av = \lambda_i v$. Hence, we have

$$1 \leq \dim E_{\lambda_i} \leq n_i = 1,$$

by Lemma 2.1.13, so that $\dim E_{\lambda_i} = 1 = n_i$, for every i. Hence, A is diagonalisable by the previous Proposition. □

Example 2.1.16. Consider the matrix

$$A = \begin{bmatrix} 1 & -3 \\ 2 & -1 \end{bmatrix},$$

from the previous examples. Then, we have $\chi_A(\lambda) = (\lambda - \sqrt{-5})(\lambda - (-\sqrt{-5}))$, so that Corollary 2.1.15 implies that A is diagonalisable.

In this section we have managed to obtain a useful criterion for when a given matrix A is diagonalisable. Moreover, this criterion is practically useful in that we have obtained a procedure that allows us to determine the diagonalisability of A by hand (or, at least, a criterion we could program a computer to undertake).

2.2 Invariant subspaces

([1], p.106-108)

In the proceeding sections we will be considering endomorphisms f of a $\mathbb{C}$-vector space V and some natural subspaces of V that we can associate to f. You may have seen some of these concepts before but perhaps not the terminology that we will adopt.

Definition 2.2.1 (Invariant subspace)**.** Let $f \in \text{End}_{\mathbb{C}}(V)$ be a linear endomorphism of V, $U \subset V$ a vector subspace of V. We say that U is *f-invariant* or *invariant with respect to f* if, for every $u \in U$ we have $f(u) \in U$.

If $A \in Mat_n(\mathbb{C})$, $U \subset \mathbb{C}^n$ a subspace, then we say that U is *A-invariant* or *invariant with respect to A* if U is T_A-invariant.

Example 2.2.2.

1. Any subspace $U \subset V$ is invariant with respect to $\text{id}_V \in \text{End}_{\mathbb{C}}(V)$. In fact, any subspace $U \subset V$ is invariant with respect to the endomorphism $c \cdot \text{id}_V \in \text{End}_{\mathbb{C}}(V)$, where

$$(c \cdot \text{id}_V)(v) = cv, \quad \text{for every } v \in V.$$

   In particular, every subspace is invariant with respect to the zero morphism of V.

2. Suppose that $V = U \oplus W$ and p_U, p_W are the projection morphisms introduced in Example 1.4.8. Then, U is p_U-invariant: let $u \in U$, we must show that $p_U(u) \in U$. Recall that if $v = u + w$ is the unique way of writing $v \in V$ as a linear combination of vectors in U and W (since $V = U \oplus W$), then

$$p_U(v) = u, \quad p_W(v) = w.$$

   Hence, since $u \in V$ can be written as $u = u + 0_V$, then $p_U(u) = u \in U$, so that U is p_U-invariant. Also, if $w \in W$ then $w = 0_V + w$ (with $0_V \in U$), so that $p_U(w) = 0_V \in W$. Hence, W is also p_U-invariant. Similarly, we have U and W are both p_W-invariant.

   In general, if $V = U_1 \oplus \cdots \oplus U_k$, with $U_i \subset V$ a subspace, then each U_i is p_{U_j}-invariant, for any i, j.

3. Let $f \in \text{End}_{\mathbb{C}}(V)$ and suppose that λ is an eigenvalue of f. Then, E_λ is f-invariant: let $v \in E_\lambda$. Then, we have $f(v) = \lambda v \in E_\lambda$, since E_λ is a vector subspace of V.

Lemma 2.2.3. *Let $f \in \mathrm{End}_{\mathbb{C}}(V)$ and $U \subset V$ an f-invariant subspace of V.*

- *Denote $f^k = f \circ f \circ \cdots \circ f$ (the k-fold composition of f on V) then U is also f^k-invariant.*
- *If U is also g-invariant, for some $g \in \mathrm{End}_{\mathbb{C}}(V)$, then U is $(f+g)$-invariant.*
- *If $\lambda \in \mathbb{C}$ then U is a λf-invariant subspace.*

Proof: Left to reader. □

Remark 2.2.4. It is important to note that the converse of the above statements in Lemma 2.2.3 do not hold.

For example, consider the matrix

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

and the associated endomorphism $T_A \in \mathrm{End}_{\mathbb{C}}(\mathbb{C}^2)$. Then, $T_A^2 = T_{A^2} = T_{I_2} = \mathrm{id}_{\mathbb{C}^2}$ (because $A^2 = I_2$), so that every subspace of $\mathbb{C}^2$ is A^2-invariant. However, the subspace $U = \mathrm{span}_{\mathbb{C}}(e_1)$ is not A-invariant since $Ae_1 = e_2$.

We can also see that $A + (-A) = 0_2$ so that every subspace of $\mathbb{C}^2$ is $(A + (-A))$-invariant, while $U = \mathrm{span}_{\mathbb{C}}(e_1)$ is neither A-invariant nor $(-A)$-invariant.

Let $f \in \mathrm{End}_{\mathbb{C}}(V)$ and U be an f-invariant subspace. Suppose that $\mathcal{B}' = (b_1, \dots, b_k)$ is an ordered basis of U and extend to an ordered basis $\mathcal{B} = (b_1, \dots, b_k, b_{k+1}, \dots, b_n)$ of V. Then, the matrix of f relative to $\mathcal{B}$ is

$$[f]_{\mathcal{B}} = [[f(b_1)]_{\mathcal{B}} \ \cdots \ [f(b_k)]_{\mathcal{B}} \cdots \ [f(b_n)]_{\mathcal{B}}] = \begin{bmatrix} A & B \\ 0_{n-k,k} & C \end{bmatrix},$$

where $A \in Mat_k(\mathbb{C})$, $B \in Mat_{k,n-k}(\mathbb{C})$, $C \in Mat_{n-k,n-k}(\mathbb{C})$. This follows because $f(b_i) \in \mathrm{span}_{\mathbb{C}}\{b_1, \dots, b_k\}$, for each $i = 1, \dots, k$.

Moreover, we can see that if $V = U \oplus W$ with U *and* W *both f-invariant*, and if $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ is an ordered basis of V, where $\mathcal{B}_1$ is an ordered basis of U, $\mathcal{B}_2$ is an ordered basis of W, then the matrix of f relative to $\mathcal{B}$ is

$$[f]_{\mathcal{B}} = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix},$$

where $A \in Mat_{\dim U}(\mathbb{C})$, $B \in Mat_{\dim W}(\mathbb{C})$.

Definition 2.2.5. Let $A \in Mat_n(\mathbb{C})$. We say that A is *block diagonal* if there are matrices $A_i \in Mat_{n_i}(\mathbb{C})$, $i = 1, \dots, k$, such that

$$A = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & A_k \end{bmatrix}.$$

So, our previous discussion implies the following

Lemma 2.2.6. *Let $f \in \mathrm{End}_{\mathbb{C}}(V)$, $U_1, \dots, U_k \subset V$ subspaces of V that are all f-invariant and suppose that*

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

Then, there exists an ordered basis $\mathcal{B}$ of V such that

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