## § 4.3 LINEAR OPERATORS: DIAGONALIZATION AND NORMAL FORMS

The notions and results to be presented in this last section on the operators of a vector space are of a special interest from both theoretical and practical points of view. They are essential for other fields of MATHEMATICS like the *Differential Equations and Systems*, but they are also met in the *Theoretical Mechanics*, in *Physics*, *Theory of Structures* – hence in the analysis and design of CIVIL ENGINEERING (and other industrial) structures and facilities.

### Eigenvalues and Eigenvectors of Endomorphisms / Operators

The eigenvalues – eigenvectors associated to a square matrix were presented in **§ 2.3** (see *Definition 3.8* - page 103, with the subsequent results), as essential concepts for the diagonalization of quadratic forms by orthogonal transformations. However, we are going to define them again in the context of linear operators, under more general conditions.

***Definition 3.1.*** Let $L: V \longrightarrow V$ be an endomorphism of the vector space $V$ over the field $\mathbf{K}$ (= $\mathbb{R}$ / = $\mathbb{C}$). A scalar $\lambda \in \mathbf{K}$ is said to be an *eigenvalue* of $L$ and $x \in V$ an (associated) *eigenvector* of $L$ if $x \neq \mathbf{0}$ and

$$\boxed{L x = \lambda x.} \tag{3.1}$$

More explicitly, $x \in V^{\star} =_{\text{not}} V \setminus \{\mathbf{0}\}$ is an eigenvector (associated to the eigenvalue $\lambda$) if its image through $L$ is proportional to itself under the multiplication by scalars – the second operation in the definition of a vector (or linear) space : $L x$ is a $\lambda$ – multiple of $x$. ◊

The way an eigenvector is associated to an eigenvalue is not quite clear so far ; it will be stated precisely in the results that follow. Let us now remark that several eigenvectors correspond to the same eigenvalue $\lambda$, all of them being (naturally) characterized by Eq. (3.1). This property is stated in

**PROPOSITION 3.1.** *Let be an endomorphism of the vector space $V$ and $\lambda \in \mathbf{K}$, $x \in V$ an eigenvalue and an eigenvector of $L$. Then* $(\forall\, \alpha \in \mathbf{K}^{\star} =_{\text{not}} \mathbf{K} \setminus \{0\})$ $\alpha x$ *is also an eigenvector corresponding to* $\lambda$.

***Proof.*** Let us first remark that

$$x \in V^{\star} =_{\text{not}} V \setminus \{\mathbf{0}\} \ \& \ \alpha \in \mathbf{K}^{\star} =_{\text{not}} \mathbf{K} \setminus \{0\}) \ \Rightarrow \ \alpha x \in V^{\star}, \tag{3.2}$$

that is $\alpha x \neq \mathbf{0}$ ; this follows from some consequences of the *Definition 1.1* of a vector space, in **§ 1.1** . The property that $\alpha x$ satisfies Eq. (3.1) is immediate :

$$L(\alpha x) = \alpha L x \underset{(3.1)}{=} \alpha (\lambda x) = (\lambda \alpha) x = \lambda (\alpha x). \tag{3.3}$$

Thus (3.3) with (3.2) imply the assertion in the statement. ■

This PROPOSITION naturally leads to the definition of a whole set of eigenvector corresponding to the same eigenvalue $\lambda$ :

$$S^*(\lambda) = \{\alpha x : Lx = \lambda x \ \& \ \alpha \in \mathbf{K}^*\}. \tag{3.4}$$

It obviously follows from (3.4) that $S^*(\lambda) \subseteq V^*$. A slightly larger (sub)set may also be associated to an eigenvalue $\lambda$:

$$\boxed{W(\lambda) = \{\alpha x : Lx = \lambda x \ \& \ \alpha \in \mathbf{K}\}.} \tag{3.5}$$

The difference between (3.4) and (3.5) is obvious : since the scalars $\alpha$ that occur in (3.5) are not required to be nonzero and $\alpha = 0 \Rightarrow \alpha x = \mathbf{0}$, it follows that $W(\lambda) \ni \mathbf{0}$ and, moreover,

$$W(\lambda) = S^*(\lambda) \cup \{\mathbf{0}\} \quad \text{or} \quad S^*(\lambda) = W(\lambda) \setminus \{\mathbf{0}\}. \tag{3.6}$$

Some properties of the eigenvalues & eigenvectors are presented in the result that follows, but another notion should be previously introduced by

***Definition* 3.2.** Let $L : V \longrightarrow V$ be a linear operator and $W \subseteq V$ a subset (or subspace) of $V$. Then $W$ is said to be *invariant under* $L$ if $LW \subseteq W$, that is

$$(\forall x \in W) \ Lw \in W \underset{\text{not}}{\Longleftrightarrow} \boxed{LW \subseteq W.} \tag{3.7}$$

In other words, a subset or a subspace of space $V$ is invariant with respect to (or under) the operator $L$ if $L$ does not take any vector $x \in W$ outside of $W$. We can now state

**PROPOSITION 3.2.** *Let $L : V \longrightarrow V$ be a linear endomorphism of space $V$ and let $\lambda \in \mathbf{K}$, $x \in V$ be an eigenvalue and an eigenvector of $L$. Then*

$(i)$ $W(\lambda)$ (defined by (3.5)) *is a subspace of $V$ and it is invariant under $L$.*

$(ii)$ If $\lambda_1, \lambda_2 \in \mathbf{K}$ *are eigenvalues of $L$ then*

$$\boxed{\lambda_1 \neq \lambda_2 \ \Rightarrow \ S^*(\lambda_1) \cap S^*(\lambda_2) = \varnothing \ \& \ W(\lambda_1) \cap W(\lambda_2) = \{\mathbf{0}\}.} \tag{3.8}$$

***Proof.*** $(i)$ Let us take $u_1, u_2 \in W(\lambda)$ and $\mu_1, \mu_2 \in \mathbf{K}$ two arbitrary scalars. It follows from definition (3.5) of $W(\lambda)$ that there exist two scalars $\alpha_1, \alpha_2 \in \mathbf{K}$ such that

$$u_1 = \alpha_1 x, u_2 = \alpha_2 x \ \& \ Lx = \lambda x. \tag{3.9}$$

$$(3.8) \Rightarrow \ \mu_1 u_1 + \mu_2 u_2 = \mu_1 \alpha_1 x + \mu_2 \alpha_2 x = (\mu_1 \alpha_1 + \mu_2 \alpha_2) x \ \& \ Lx = \lambda x. \tag{3.10}$$

But we have, inside $(\ldots)$ of (3.9), a scalar $\alpha = \mu_1 \alpha_1 + \mu_2 \alpha_2 \in \mathbf{K}$ what shows that $W(\lambda)$ is closed under arbitrary linear combinations of (two) vectors, hence it is a subspace $V$. If we now consider a vector $u \in W(\lambda)$, it follows from (3.5) that $u = \alpha x$ for some $\alpha \in \mathbf{K}$ & $Lx = \lambda x$. Therefore

$$Lu = L(\alpha x) = \alpha L x \underset{(3.1)}{=} \alpha(\lambda x) = (\alpha\lambda) x = \beta x \in W(\lambda),$$

in view of the same definition (3.5). Thus $(i)$ is entirely proved.

$(ii)$ Let us assume that any of the equalities in (2.8) does not hold. In fact, they are equivalent in view of (3.6). Hence it would exist a nonzero vector

$$u \in S^{\star}(\lambda_1) \cap S^{\star}(\lambda_2) \quad \text{or} \quad u \in (W(\lambda_1) \cap W(\lambda_2) \setminus \{\mathbf{0}\}. \tag{3.11}$$

But it follows from definitions (3.4) / (3.5) that $u$ is a vector corresponding to both eigenvalues $\lambda_1, \lambda_2 \in \mathbf{K}$ $(\lambda_1 \neq \lambda_2)$:

$$L u = \lambda_1 u \ \ \& \ \ L u = \lambda_2 u.$$

If these two equations are side-by-side subtracted we get

$$L u - L u = \lambda_1 u - \lambda_2 u = (\lambda_1 - \lambda_2) u \ \Rightarrow \ \lambda_1 - \lambda_2 = 0. \tag{3.12}$$

This equation in (3.12) follows from $u \neq \mathbf{0}$ and from consequence $(\mathbf{L}_{15})$ of *Definition 1.1* in § **2.1**. Clearly, (3.12) $\Rightarrow$ $\lambda_1 = \lambda_2$, what contradicts the hypothesis in $(ii)$. This concludes the proof.

■

***Remark 3.1.*** It follows from PROPOSITIONS 3.1 & 3.2 that any eigenvalue of an operator is associated with several eigenvectors, or - more exactly - *several eigenvectors correspond to a single eigenvalue*. In fact, the whole set $S^{\star}(\lambda)$ of eigenvectors corresponds to $\lambda$. Instead, an eigenvector cannot correspond to two (or more) *distinct eigenvalues*. This follows from part $(ii)$ in the previous PROPOSITION, but it can also be directly proved. Indeed, if we suppose that $L x = \lambda_1 x$ and also $L x = \lambda_2 x$, it easily follows (through side-by-side subtraction) that $L x - L x = \lambda_1 x - \lambda_2 x = (\lambda_1 - \lambda_2) x = \mathbf{0} \Rightarrow \lambda_1 = \lambda_2$, as in (3.12), since $x \neq \mathbf{0}$ like any eigenvector.

The number of eigenvectors corresponding to the same eigenvalue is infinite ; it results from (3.4) that there exist at least "as many" eigenvectors in $S^{\star}(\lambda)$ as scalars in $\mathbf{K}^{\star}$. It will be seen that not all of them are "relevant" ; in fact, a (finite) basis of eigenvectors spanning $S^{\star}(\lambda)$ or $W(\lambda)$ will entirely determine these sets.

Another property concerning the distinct eigenvalues of a linear endomorphism (and corresponding eigenvectors) is stated in

**PROPOSITION 3.3.** *Let $L : V \longrightarrow V$ be a linear endomorphism of space $V$ with $p$ distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$. Then any $p$ eigenvectors respectively corresponding to these eigenvalues are linearly independent.*

***Proof.*** Before writing down the $p$ eigenvectors, let us notice that *they are necessarily distinct*, in view of the previous PROPOSITION 3.2. Certainly, for each eigenvalue $\lambda_i$ $(1 \leq i \leq p)$ infinitely many vectors correspond to it, but we consider (or choose) only one of them. The diagram below illustrates this correspondence.

$$\begin{array}{cccc} \lambda_1 & \lambda_2 & \cdots & \lambda_p \\ \downarrow & \downarrow & \cdots & \downarrow \\ x_1 & x_2 & \cdots & x_p \end{array} \quad \text{with} \quad (\forall\, i \in \{1, 2, \ldots, p\}) \ \ L x_i = \lambda_i x_i. \tag{3.13}$$

Certainly, the eigenvalues on the first line in (3.13) satisfy the condition in the statement :

$$(\forall\, i,j \in \{1,2,\ldots,p\})\ i \neq j \ \Rightarrow\ \lambda_i \neq \lambda_j. \tag{3.14}$$

We must prove the linear independence of the $p$ vectors in (3.13), what is equivalent to the implication

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_p x_p = \mathbf{0} \ \Rightarrow\ \alpha_1 = \alpha_2 = \ldots = \alpha_p = 0. \tag{3.15}$$

The proof proceeds by induction with respect to $p$. Let us denote the property to be proved as $\mathbf{P}(p)$: (3.13) & (3.14) $\Rightarrow$ (3.15). $\mathbf{P}(1)$ obviously holds, in view of *Def. 3.1*. If $x_1$ is an eigenvector corresponding to the eigenvalue $\lambda_1$ we have $x_1 \neq \mathbf{0} \Rightarrow [\alpha_1 x_1 \neq \mathbf{0} \Rightarrow \alpha_1 = 0]$ according to consequence $(\mathbf{L}_{15})$ of *Definition 1.1* of **§ 1.1** . Hence $x_1$ is a linearly independent vector – the simplest case of linear independence. Let us now assume that $\mathbf{P}(p)$ holds and try to show that $\mathbf{P}(p+1)$ holds, too. Certainly, hypotheses (3.13) and (3.14) should be extended to $p+1$ pairs of eigenvalues – eigenvectors : $p$ should be replaced by $p+1$ in (3.13) , (3.14) and in (3.15) – the property to be proved – too. The equation involved in (3.15), for $p+1$ vectors, is

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_p x_p + \alpha_{p+1} x_{p+1} = \mathbf{0}. \tag{3.16}$$

Next, the operator $L$ is applied to Eq. (3.16) and its extended linearity (see **§ 3.2** ) is used giving

$$\alpha_1 L x_1 + \alpha_2 L x_2 + \ldots + \alpha_p L x_p + \alpha_{p+1} L x_{p+1} = \mathbf{0}. \tag{3.17}$$

But $x_1, x_2, \ldots, x_p, x_{p+1}$ are eigenvectors corresponding to the respective $p+1$ eigenvalues. Hence $L x_i = \lambda_i x_i$ $(1 \leq i \leq p+1)$ and (3.17) becomes

$$\alpha_1(\lambda_1 x_1) + \alpha_2(\lambda_2 x_2) + \ldots + \alpha_p(\lambda_p x_p) + \alpha_{p+1}(\lambda_{p+1} x_{p+1}) = \mathbf{0} \ \Rightarrow$$

$$\Rightarrow\ \lambda_1(\alpha_1 x_1) + \lambda_2(\alpha_2 x_2) + \ldots + \lambda_p(\alpha_p x_p) + \lambda_{p+1}(\alpha_{p+1} x_{p+1}) = \mathbf{0}. \tag{3.18}$$

But $\alpha_{p+1} x_{p+1}$ in the last term of (3.18) can be expressed in terms of the other $p$ terms from Eq. (3.16) :

$$\alpha_{p+1} x_{p+1} = -\alpha_1 x_1 - \alpha_2 x_2 - \ldots - \alpha_p x_p. \tag{3.19}$$

Eqs. (3.18) & (3.19) $\Rightarrow \ldots$

$$\ldots \Rightarrow\ \alpha_1(\lambda_1 - \lambda_{p+1}) x_1 + \alpha_2(\lambda_2 - \lambda_{p+1}) x_2 + \ldots + \alpha_p(\lambda_p - \lambda_{p+1} x_p) = \mathbf{0}. \tag{3.20}$$

The linear independence of the $p$ vectors that occur in (3.20), assumed by $\mathbf{P}(p)$, implies

$$(\forall\, i \in \{1,2,\ldots,p\})\ \alpha_i(\lambda_i - \lambda_{p+1}) x_i = \mathbf{0}. \tag{3.21}$$

But condition (3.14), extended to the $p+1$ distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_{p+1}$, implies that

$$(\forall\, i \in \{1,2,\ldots,p\})\ \lambda_i - \lambda_{p+1} \neq 0 \ \underset{(3.21)}{\Rightarrow}\ \alpha_1 = \alpha_2 = \ldots = \alpha_p = 0 \ \underset{(3.16)}{\Rightarrow}$$

$$\Rightarrow\ \alpha_{p+1} x_{p+1} = \mathbf{0} \ \Rightarrow\ \alpha_{p+1} = 0$$

since $x_{p+1} \neq \mathbf{0}$ as any eigenvector. Hence $\alpha_1 = \alpha_2 = \ldots = \alpha_p = \alpha_{p+1} = 0$ and the $p+1$ eigenvectors are thus linearly independent, $\mathbf{P}(p+1)$ holds and the Proposition is proved.

■

The definitions and results, so far presented, have been rather theoretical. But it is also important to see how the eigenvalues and eigenvectors of an operator can be effectively found. It is necessary to rewrite definition (3.1) of an eigenvalue with a corresponding eigenvector $x$ under an equivalent form. Let us recall (from § **4.1** & § **4.2** ) that the identity endomorphism of a vector space $V$ is defined by

$$(\forall\, x \in V)\ \ \mathrm{id}_V\, x \underset{\text{not}}{=} \mathbf{1}_V\, x = x. \tag{3.22}$$

In other words, the identity map(ping) $\mathrm{id}_V$ or $\mathbf{1}_V$ leaves unchanged any vector $x$ in $V$. It is obvious that the matrix of $\mathrm{id}_V$ in any basis of a finitely generated space $V$ (with $\dim V = n$) equals the identity (or unit) matrix $I_n$. With these preliminaries, Eq. (3.1) can be equivalently rewritten as

$$L\,x = \lambda\, \mathrm{id}_V\, x \iff (L - \lambda\, \mathrm{id}_V)\, x = \mathbf{0}. \tag{3.23}$$

It follows, from (3.23), that any eigenvector $x$ corresponding to the eigenvalue $\lambda$ is taken to the zero vector $\mathbf{0}$ by the operator $L - \lambda\, \mathrm{id}_V$, that is

$$L\,x = \lambda\, x \iff (L - \lambda\, \mathrm{id}_V)\, x = \mathbf{0} \iff x \in \mathrm{Ker}\,(L - \lambda\, \mathrm{id}_V). \tag{3.24}$$

The last membership relation in (3.24) allows to conclude that

$$\boxed{L\,x = \lambda\, x \iff x \in \mathrm{Ker}\,(L - \lambda\, \mathrm{id}_V).} \tag{3.25}$$

Thus, *the kernel of the operator* $\mathrm{L} - \lambda\, \mathrm{id}_V$ *contains all the eigenvectors corresponding to the eigenvalue* $\lambda$. On another hand, it is easy to see that the matrix of operator $L - \lambda\, \mathrm{id}_V$ in any basis $A$ of $V$ is

$$(L - \lambda\, \mathrm{id}_V)_A = L_A - \lambda\, I_n. \tag{3.26}$$

If we now look for the coordinates $X_A$ of an eigenvector $x \in V$ corresponding to the eigenvalue $\lambda$ in basis $A$, it follows from the previous remarks and from PROPOSITION 2.3 of § **4.2** that its coordinates $X_A$ will have to satisfy the matrix equation

$$X_A^{\mathrm{T}} \cdot (L_A - \lambda\, I_n) = \mathbf{0}^{\mathrm{T}} = [0\ \ 0\ \ \ldots\ \ 0]. \tag{3.27}$$

The zero vector $\mathbf{0} \in \mathbf{K}^n$ is the vector of coordinates of $\mathbf{0} \in V$ in any basis $A$ of space $V$. The matrix equation (3.26) can be equivalently rewritten (by a simple transposition of the matrix and row vectors there involved) as

$$\boxed{(L_A^{\mathrm{T}} - \lambda\, I_n) \cdot X_A = \mathbf{0} \in \mathbf{K}^n.} \tag{3.27'}$$

This last matrix equation represents, in fact, a homogeneous system whose matrix is $L_A^{\mathrm{T}} - \lambda\, I_n$. But let us recall that we look for the coordinates of a <u>nonzero</u> (eigen)vector. Hence, only the notrivial solution of the H-system ( $\Longleftrightarrow$ (3.27)) are acceptable. As known (and stated in **§ 1.2** ), the necessary (and sufficient) condition for a homogeneous system to admit nontrivial solutions is that *the rank of its matrix be less than the number of its unknowns* (= the number columns of its matrix). And this latter condition is equivalent to the singularity of this square matrix, that is the determinant of the system should be $= 0$. We have thus arrived to a necessary condition for the coordinates $X_A$ of a vector $x$ in basis $A$ of space $V$ to be an eigenvector corresponding to the eigenvalue $\lambda$. Let us state it as

**PROPOSITION 3.4.** *Let* $L : V \longrightarrow V$ *be a linear endomorphism of space* $V$ *with its matrix* $L_A$ *in basis* $A$ *of space* $V$ *then the coordinates* $X_A$ *of any eigenvector* $\mathbf{x} \in V$ *corresponding to the eigenvalue* $\lambda$ *should satisfy the homogeneous system* (3.27) *with the necessary condition*

$$\det(L_A^{\mathrm{T}} - \lambda I_n) = 0 \iff \boxed{\det(L_A - \lambda I_n) = 0.} \tag{3.28}$$

But this condition (3.28) is also a condition on any eigenvalue $\lambda$ of $L$. Therefore any eigenvalue $\lambda$ of the operator is obliged to satisfy Eq. (3.28), which is - in fact - an algebraic equation of order $n$ over the field $\mathbf{K}$. This follows from the fact that

$$\det(L_A - \lambda I_n) \underset{\text{not}}{=} P_L(\lambda) \tag{3.29}$$

is a polynomial of order $n$ with its coefficients in the field $\mathbf{K}$. This $P_L(\lambda)$ is called the *characteristic polynomial* of the operator $L$. Obviously, its coefficients are products of entries of matrix $L_A$ and it is therefore basis-dependent. However, it can be proved that the roots of the equation (3.28) remain the same for any basis considered in the space $V$. The set of these roots is called the *spectrum* of $L$ and it is denoted as $\sigma(L)$.

An important question here arises :

Are all the roots of Eq. (3.28) in the field $\mathbf{K}$ ?

The answer is positive when the field is "algebraically closed", that is any algebraic equation with its coefficients in $\mathbf{K}$ has all its roots in $\mathbf{K}$. Such a field is the complex field $\mathbf{K} = \mathbb{C}$, according to the famous *Fundamental Theorem of Algebra* due to the French mathematician Evariste Galois (1811-1832). The answer can be negative for particular equations over a not algebraically closed field like the field of real numbers $\mathbb{R}$. Let us recall that we met this problem in **§ 3.3**, where the method of orthogonal transformations (or the EVV-based method) was presented for the diagonalization of quadratic forms. But the matrix of any such Q-form is symmetric and the *symmetric matrices over* $\mathbb{R}$ *have only real eigenvalues*. As regards the operators, let us consider a very simple example.

**Example 3.1.** Let $V$ be a vector space of dimension 2, spanned by the basis $A = [a_1 \ \ a_2]$ and let the operator $L : V \longrightarrow V$ be defined by

$$L\,a_1 = a_2 \ \ \& \ \ L\,a_2 = -a_2 \ \Rightarrow \ L_A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \ \Rightarrow \ P_L(\lambda) = \begin{vmatrix} \lambda & 1 \\ -1 & \lambda \end{vmatrix} = \lambda^2 + 1 . \tag{3.30}$$

The two roots of the polynomial in (3.30) are $\lambda_{1,2} = \pm i \ \ (i^2 = -1) \ \Rightarrow \ \lambda_{1,2} \in \mathbb{C} \setminus \mathbb{R}$. □

For the moment, let us assume that all the eigenvalues of $L$ are in $\mathbf{K}$ (that is, $\sigma(L) \subset \mathbf{K}$) or let us consider only those roots of Eq. (3.28) that meet this condition. The next problem regards the way to find the eigenvectors corresponding to an eigenvalue $\lambda_j \in \sigma(L) \cap \mathbf{K}$.

## Finding the eigenvalues and eigenvectors of an operator

The matrix $L_A \in \mathscr{M}_n(\mathbf{K})$ is assumed to be known / given.

**EVV 1** The characteristic polynomial $P_L(\lambda) = \det(L_A - \lambda I_n)$ is written.

**EVV 2** The characteristic equation $P_L(\lambda) = 0$ is solved (in $\mathbf{K}$) giving

$$\lambda_1, \lambda_2, \ldots, \lambda_m \in \sigma(L) \cap \mathbf{K} \tag{3.31}$$

as *distinct* eigenvalues.

**EVV 3** For each $\lambda_j \in \sigma(L) \cap \mathbf{K} \ (1 \le j \le m)$ the homogeneous system (3.27) is written and solved. The basis spanning the set of its nontrivial solutions $S^\star(\lambda_j)$ consists of the relevant eigenvectors corresponding to $\lambda_j$.

**EVV 4** A sub-basis $C_j$ spaning $S^\star(\lambda_j)$ has to be found. $C = [C_1 \ \ldots \ C_j \ \ldots \ C_m]$ will be a *canonical basis* in which the matrix of the operator is expected to be diagonal.

***Remark 3.1.*** Before giving an example (to illustrate how the eigenvalues and eigenvector of an operator can be (found), let us see that the homogeneous system whose (general) solution for $\lambda_j$ gives $S^\star(\lambda_j)$ can be solved on its matrix, obtained from the matrix of the H-system in (3.27') by simply taking $\lambda_j$ instead of $\lambda$. The resulting matrix will be singular since $\lambda_j$ is a root of Eq. (3.28). Hence the corresponding H-system will admit non-trivial solutions.

**Example 3.2.** The operator $L : V \longrightarrow V$ is specified by its matrix in a basis $A$,

$$L_A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 1 & 1 & 2 \end{bmatrix}. \tag{3.32}$$

It is required to find its eigenvalues and eigenvectors (eigen-subspaces).

The characteristic polynomial is

$$P_L(\lambda) = \det(L_A - \lambda I_3) = \det \begin{bmatrix} 2-\lambda & 1 & 1 \\ 2 & 3-\lambda & 2 \\ 1 & 1 & 2-\lambda \end{bmatrix} = \begin{vmatrix} 1-\lambda & 0 & \lambda-1 \\ 0 & 1-\lambda & 2\lambda-2 \\ 1 & 1 & 2-\lambda \end{vmatrix} =$$

$$= (\lambda - 1)^2 \begin{vmatrix} -1 & 0 & 1 \\ 0 & -1 & 2 \\ 1 & 1 & 2-\lambda \end{vmatrix} = (\lambda - 1)^2 (5 - \lambda). \tag{3.33}$$

(3.33) $\Rightarrow$ $\sigma(L) = \{1, 5\}$. $\lambda_1 = \lambda_2 = 1$ is a double root, hence a double eigenvalue, while $\lambda_3 = 5$ is a simple eigenvalue. The matrix of the first homogeneous system (for $\lambda_1 = \lambda_2 = 1$) is

$$L_A^{\mathrm{T}} - 1 I_3 = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \end{bmatrix} \sim [\,1 \;\; 2 \;\; 1\,]. \tag{3.34}$$

The general solution of this system will depend on two variables (parameters) that can be denoted as $\xi_2 = \alpha$ and $\xi_3 = \beta$; we thus obtain the first general eigenspace corresponding to

$$\lambda_1 = \lambda_2 = 1: \; u = U_1(\alpha, \beta) \cdot A^{\mathrm{T}} = \begin{bmatrix} -2\alpha - \beta \\ \alpha \\ \beta \end{bmatrix} \cdot A^{\mathrm{T}}. \tag{3.35}$$

Hence the eigen-subspace $W(\lambda_1) = W(\lambda_2) = W(1)$ is spanned by two eigenvectors that follow from (3.35); they can be obtained by giving particular values to the parameters $\alpha$ & $\beta$. Obviously, $\alpha = \beta = 0$ is excluded. Two pairs of values that ensure the linear independence of the eigenvectors thus obtained are

$$\alpha = 1 \;\&\; \beta = 0 \underset{(3.35)}{\Rightarrow} u_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \cdot A^{\mathrm{T}} = -2\,a_1 + a_2 \underset{\text{not}}{=} b_1, \tag{3.36}$$

$$\alpha = 0 \;\&\; \beta = 1 \underset{(3.35)}{\Rightarrow} u_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \cdot A^{\mathrm{T}} = -a_1 + a_3 \underset{\text{not}}{=} b_2. \tag{3.37}$$

Since the vectors in (3.36) & (3.37) are two eigenvectors corresponding to the first eigenvalue, the corresponding eigenspace can be effectively written as

$$\begin{aligned} W(\lambda_1) = W(\lambda_2) = W(1) &= \{\mu(-2\,a_1 + a_2) + \nu(-a_1 + a_3) : \mu, \nu \in \mathbb{R}\} = \\ &= \{(-2\mu - \nu)a_1 + \mu\, a_2 + \nu\, a_3 : \mu, \nu \in \mathbb{R}\}. \end{aligned}$$

In fact, this expression of a general vector in the first eigen-space is consistent with the general solution of the H-system, written in Eq. (3.35). For the third (or second distinct) eigenvalue we similarly have

$$L_A^{\mathrm{T}} - 5\, I_3 = \begin{bmatrix} -3 & 2 & 1 \\ 1 & -2 & 1 \\ 1 & 2 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 \\ 0 & -4 & 4 \\ 0 & 4 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix} \Rightarrow$$

$$\Rightarrow \lambda_3 = 5: \; v = U_3(\gamma) \cdot A^{\mathrm{T}} = \begin{bmatrix} \gamma \\ \gamma \\ \gamma \end{bmatrix} \cdot A^{\mathrm{T}} \;\Rightarrow\; u_3 = a_1 + a_2 + a_3. \tag{3.38}$$

$$(3.38) \quad \Rightarrow \quad W(\lambda_3) = W(5) = \{ \gamma (a_1 + a_2 + a_3) : \gamma \in \mathbb{R} \}.$$

□

***Remark* 3.2.** This example illustrates an interesting case, namely the one when the eigenspace $W(\lambda_j)$ is spanned by a basis consisting of more than one (eigen)vector. In the previous example, $W(\lambda_1) = W(\lambda_2) = W(1) = \mathscr{L}(\{b_1, b_2\})$. The dimension of an eigen-subspace $W(\lambda_j)$ is called the *geometric multiplicity* of this eigenvalue and it is denoted as

$$h_j = \dim W(\lambda_j). \tag{3.39}$$

On another hand, the characteristic polynomial $P_L(\lambda)$ of an operator has been assumed to admit $m$ distinct roots $\lambda_j \; (1 \le j \le m)$: see Eq. (3.31). It is known from the theory of the algebraic equations that such a polynomial can be factorized into $m$ factors corresponding to its distinct roots - the eigenvalues :

$$P_L(\lambda) = (-1)^n (\lambda - \lambda_1)^{k_1} \ldots (\lambda - \lambda_j)^{k_j} \ldots (\lambda - \lambda_m)^{k_m}, \tag{3.40}$$

where $k_j$ is said to be the *algebraic multiplicity* of $\lambda_j \; (1 \le j \le m)$. The next result states an inequality between these two multiplicities.

**PROPOSITION 3.5.** *Let $L: V \longrightarrow V$ be a linear operator of space $V$ with its matrix $L_A$ in basis $A$ of space $V$ and its characteristic polynomial $P_L(\lambda)$ admitting a factorization of the form* (3.40) *that corresponds to the spectrum*

$$\sigma(L) = \{ \lambda_1, \lambda_2, \ldots, \lambda_m \}. \tag{3.41}$$

*Then, for any eigenvalue $\lambda_j \; (1 \le j \le m)$, its geometric multiplicity is at most equal to its algebraic multiplicity* :

$$\boxed{h_j \le k_j \; (1 \le j \le m).} \tag{3.42}$$

We do not give a proof of this result. We only mention that it will be involved in an important theorem regarding the possibility to change the matrix of an endomorphism to a simpler form. Before giving the next definition, let us remark – on the algebraic multiplicities that occur in (3.40), if $\sigma(L) = \{ \lambda_1, \lambda_2, \ldots, \lambda_m \} \subset \mathbf{K}$ – that they satisfy the equation

$$\sum_{j=1}^{m} k_j = n = \dim V. \tag{3.43}$$

***Definition* 3.3.** Let $L: V \longrightarrow V$ be a linear operator of space $V$ with its matrix $L_A$ in basis $A$ of space $V$. The operator $L$ is said to be *diagonalizable* if there exists a basis $B$ of $V$ such that

$$L_B = \begin{bmatrix} \delta_1 & 0 & \dots & 0 \\ 0 & \delta_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \delta_n \end{bmatrix} \underset{\text{not}}{=} \lceil \delta_1 \ \delta_2 \ \dots \ \delta_n \rfloor. \tag{3.44}$$

***Remarks 3.3.*** Let us recall that we met this problem – of *bringing a matrix to a simpler form* – in some of the previous chapters and sections. For instance, the Gaussian elimination method for solving linear systems (in **§ 1.2**) implied the transformation of the matrix of a system $A \cdot X = \boldsymbol{b}$ by (linear) transformations on the rows of the augmented matrix $\tilde{A} = [A \mid \boldsymbol{b}]$ until the identity matrix was obtained instead of $A$ of instead of a square submatrix of $A$ (of order $r = \operatorname{rank} A$); the respective form was called a diagonal / quasi-diagonal form. Earlier in the same **§ 1.2** , the transformations with the rows / columns of a matrix were presented (and used in examples) for determining the rank of a matrix, until a quasi-triangular was obtained. But a more typical instance of turning a square matrix into a diagonal matrix of the form (3.44) was discussed and exemplified in **§ 2.3** - *The Diagonalization of Quadratic Forms*. All the three methods there presented (**GAUSS**, **JACOBI** and **OT-EVV** or **ORTHOGONAL TRANSFORMATIONS**) turned the (symmetric) matrix $[\alpha]$ or $[\varepsilon]$ of a Q-form into a diagonal matrix. These transformations involved the relationship defining *similar matrices*, and it is appropriate to recall a couple of equations / formulas in that **§ 2.3** :

Two matrices $A, B$ were said to be *similar* (denoted $A \sim B$) by Eq. (3.134) at page 105 :

$$\boxed{A \sim B \iff (\exists\, S:\ \det S \neq 0)\ \ B = S^{-1} \cdot A \cdot S.} \tag{3.45}$$

We proved that a square matrix $[\varepsilon]$ consisting of $n$ independent eigenvectors of matrix $A$ can be the similarity matrix $S$ of (3.45) - see Eq. (3.132) in **§ 2.3** , page 105 :

$$S = [U_1 \ U_2 \ \dots \ U_n] \text{ with } A \cdot U_i = \lambda_i U_i. \tag{3.46}$$

Characterization (3.45) of similar matrices was obtained from Eq. (3.133) in **§ 2.3** , that is

$$A \cdot S = S \cdot \lceil \lambda_1 \ \lambda_2 \ \dots \ \lambda_n \rfloor. \tag{3.47}$$

$$(3.47) \Rightarrow \quad \boxed{S^{-1} \cdot A \cdot S = \lceil \lambda_1 \ \lambda_2 \ \dots \ \lambda_n \rfloor.} \tag{3.48}$$

Obviously, the eigenvectors $U_1, U_2, \dots, U_n$ in (3.46) respectively correspond to the eigenvalues of (3.47) - (3.48) and their linear independence is essential : it ensures the non-singularity of the similarity matrix $S$. This property makes possible the implication

$$[SB = AS \text{ or } AS = SB] \Rightarrow B = S^{-1} \cdot A \cdot S \underset{\text{def}}{\Rightarrow} A \sim B. \tag{3.49}$$

***Remarks 3.4.*** As regards the operators, we have presented and proved the formula for the matrix change of an operator after a change of basis in **§ 4.2** : Eq. (2.73) in PROPOSITION 2.8 (page 179) :

$$\boxed{L_B = T \cdot L_A \cdot T^{-1}.} \tag{3.50}$$

We also remarked the equivalence between this transformation formula and the similarity relation of (3.49) : (3.50) $\Rightarrow$ $L_A \sim L_B$ with $S = T^{-1} \iff T = S^{-1}$.

## Diagonalization of Operators

The diagonalization of an operator is (as just mentioned) a problem of transforming its matrix its matrix $L_A$ in basis $A$ of space $V$ into a simpler matrix like the diagonal matrix of Eq. (3.43). There are cases when an effective diagonalization (by means of EVVs) is not possible, but a close-to-diagonal matrix can however be obtained. Such situations will be considered in the next subsection.

A diagonalization process essentially involves a change of basis $A \to B$. If such a change is achieved by means of a transformation matrix $T$, that is $B^{\mathrm{T}} = T \cdot A^{\mathrm{T}}$, then the matrix of the operator changes by formula (3.49), just recalled above. The relation among the matrices $L_A$ and $L_B$ is also a similarity relation, as earlier mentioned :

$$L_B = T \cdot L_A \cdot T^{-1} \iff L_B^{\mathrm{T}} = S^{-1} \cdot L_A^{\mathrm{T}} \cdot S \quad \text{with} \quad S = T^{-\mathrm{T}} \iff T = S^{-\mathrm{T}}. \tag{3.51}$$

In this equivalence we have replaced $L_A \to L_A^{\mathrm{T}}$ and (consequently) $L_B \to L_B^{\mathrm{T}}$ taking into account the equations (3.45) & (3.46) at page 206 : they imply that the matrix satisfying the EVV equation is *the transpose* of $L_A$ and not $L_A$ itself. This remark is even more relevant in the case of operators on an Euclidean space $\mathbf{K}^n$ or $\mathbb{R}^n$, when the transpose $M = L_E^{\mathrm{T}}$ is usually involved in determining the eigenvectors.

To conclude, the term of *diagonalization* of an endomorphism will be used in this (narrower) sense, the one of *Definition 3.3*. Although the transformation matrix is not there (explicitly) considered, it always exist when a change of basis $A \to B$ is taken into account.

A necessary and sufficient condition for an operator to be diagonalizable is given by

**THEOREM 3.1.** *Let $L : V \longrightarrow V$ be a linear operator of space $V$ with its matrix $L_A$ in the basis $A$ of space $V$ and its characteristic polynomial $P_L(\lambda)$. $L$ is diagonalizable if and only if there exists a basis of $V$ consisting of eigenvectors of $L$.*

***Proof.*** $(\Rightarrow)$ : According to *Definition 3.3.*, there exists a basis $B$ in which the matrix of $L$ is of the form (3.44). According to the definition of the matrix $L_B$ in a basis (see PROPOSITION 2.2 in § **4.2** - Eq. (2.10) at page 165, rewritten for basis $B$)

$$L\,B^{\mathrm{T}} = L_B \cdot B^{\mathrm{T}} \underset{(3.44)}{\Rightarrow} \begin{cases} L\,b_1 = \delta_1 b_1, \\ L\,b_2 = \delta_2 b_2, \\ \quad \vdots \qquad \vdots \\ L\,b_n = \delta_n b_n. \end{cases} \tag{3.52}$$

But it follows from (3.52) and Eq. (3.1) of *Definition 3.1* that $b_1, b_2, \ldots, b_n$ are eigenvectors of $L$, respectively corresponding to the eigenvalues $\delta_1, \delta_2, \ldots, \delta_n$. Hence the condition in the statement is necessary.

$(\Leftarrow)$: Let us now assume that $\mathcal{B} = \{b_1, b_2, \ldots, b_n\}$ or $B = [b_1\ b_2\ \ldots\ b_n]$ is a basis of $V$ consisting of the $n$ eigenvectors of $L$, respectively corresponding to $L$'s eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. According to *Definition 3.1,*

$$(\forall\, i \in \{1, 2, \ldots, n\})\ L\, b_i = \lambda_i\, b_i. \tag{3.53}$$

The $n$ equations (3.53) can be written one under the other resulting in the matrix equation

$$\begin{bmatrix} L\, b_1 \\ L\, b_2 \\ \vdots \\ L\, b_n \end{bmatrix} = L\, B^{\mathrm{T}} = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} \cdot B^{\mathrm{T}}. \tag{3.54}$$

It obviously follows from equation (3.54), with Eq. (2.10) at page 165, that the matrix $L_B$ of $L$ in the basis $B$ is just the diagonal matrix with the $n$ eigenvalues on its main diagonal :

$$L_B = \lceil \lambda_1\ \lambda_2\ \ldots\ \lambda_n \rfloor. \tag{3.55}$$

Hence the condition in the statement is sufficient, too. ■

***Remarks* 3.5.** The preceding THEOREM gives not only a necessary and sufficient condition for an operator to be diagonalizable. It also gives the structure of the matrix $L_B$ in the basis consisting of the $n$ eigenvectors : $n$ eigenvalues to which these vectors $b_1, b_2, \ldots, b_n$ correspond are the entries on the main diagonal of $L_B$ (while the other entries are $= 0$). But it is important to notice that these eigenvalues *should not be pairwise distinct* while the vectors $b_1, b_2, \ldots, b_n$ are necessarily distinct : otherwise they would be linearly dependent. A basis cannot contain two identical vectors ! In other words, an eigenvalue can appear several times on the main diagonal of $L_B$. Without now going into details, we state that such a situation occurs in the case when an eigenvalue $\lambda_j$ $(1 \le j \le n)$ has its algebraic multiplicity $k_j > 1$ : see factorization (3.40) of the characteristic polynomial (at page 205) and **Example 4.2**. If $h_j = k_j$ in inequality (3.42) (PROPOSITION 3.5 at page 205), $\lambda_j$ will appear exactly $k_j$ on the main diagonal of $L_B$ and $h_j = k_j$ vectors will occur in basis $B$. Although this is a technical detail, the $k_j$ identical values of $\lambda_j$ are usually written together on the diagonal of $L_B$: they form a diagonal submatrix (or block),

$$\lceil \lambda_j\ \lambda_j\ \ldots\ \lambda_j \rfloor \text{ of size } (k_j, k_j).$$

Similarly, the $h_j = k_j$ distinct vectors corresponding to this multiple eigenvalue can be written together, if the basis is written as an ordered $n$-tuple of vectors, $B = [b_1\ b_2\ \ldots\ b_n]$. The structure of the basis will be $B = [\ldots\ b_{j_1}\ b_{j_2}\ \ldots\ b_{j_{k_j}}\ \ldots]$. If we "take off" this sub-basis

$$B_j = [\,b_{j_1}\ b_{j_2}\ \ldots\ b_{j_{k_j}}\,], \tag{3.56}$$

we shall see that it is just the basis spanning the solution set $S_j$ to the homogeneous system

$$\boxed{(L_A^{\mathrm{T}} - \lambda_j\, I_n) \cdot U = \mathbf{0}.} \tag{3.57}$$

The next theorem gives another necessary and sufficient condition for the possibility to diagonalize an operator, in terms of the multiplicities of its eigenvalues and corresponding eigenvectors.

**THEOREM 3.2.** *Let $L : V \longrightarrow V$ be a linear operator of space $V$ with its matrix $L_A$ in a basis $A$ of space $V(\dim V = n)$ and the characteristic polynomial $P_L(\lambda)$. Let $\sigma(L) = \{\lambda_1, \lambda_2, \ldots, \lambda_m\} \subset \mathbf{K}$ be the distinct eigenvalues of $L$ with their respective algebraic multiplicities $k_1, k_2, \ldots, k_m$ and the corresponding eigen(sub)spaces*

$$W(\lambda_1),\ W(\lambda_2), \ldots,\ W(\lambda_m). \tag{3.58}$$

$(i)$ *The endomorphism $L$ is diagonalizable if and only if*

$$\boxed{(\forall\, j \in \{1, 2, \ldots, m\})\ \dim W(\lambda_j) \underset{\text{not}}{=} h_j = k_j\,.} \tag{3.59}$$

$(ii)$ *The sum of the eigen(sub)spaces in* (3.58) *is direct* :

$$W(\lambda_1) + W(\lambda_2) + \ldots + W(\lambda_m) = W(\lambda_1) \oplus W(\lambda_2) \oplus \ldots \oplus W(\lambda_m). \tag{3.60}$$

***Proof.*** $(i)$ $(\Leftarrow)$: By the hypothesis on the $m$ eigenvalues,

$$(\forall\, i, j \in \{1, 2, \ldots, m\})\ i \neq j \ \Rightarrow\ \lambda_i \neq \lambda_j\,. \tag{3.61}$$

In view of PROPOSITION 3.2 - Eq. (3.8), (3.61) $\Rightarrow$

$$(\forall\, i, j \in \{1, 2, \ldots, m\})\ i \neq j \ \Rightarrow\ W(\lambda_i) \cap W(\lambda_j) = \{\mathbf{0}\}. \tag{3.62}$$

According to the last mention in *Remarks 3.5*, the coordinates $X_A$ of any vector in the eigenspace corresponding to $\lambda_j$ should satisfy a homogeneous system of the form (3.57) :

$$x = A X_A \in W(\lambda_j) \iff (L_A^{\mathrm{T}} - \lambda_j I_n) \cdot X_A = \mathbf{0}. \tag{3.63}$$

But we have also noticed that the subspace of (3.63) is just the solution subspace $S_j$ of the homogeneous system (3.56) or (3.63). Hence

$$x = A X_A \in W(\lambda_j) \iff X_A \in S_j\,. \tag{3.64}$$

By the inequality (3.42) in PROPOSITION 3.5 and Eq. (3.43) at page 205,

$$\sum_{j=1}^{m} h_j \leq \sum_{j=1}^{m} k_j = n \underset{(3.56)}{\Rightarrow} \sum_{j=1}^{m} h_j = n \ \Rightarrow\ \bigcup_{j=1}^{m} W(\lambda_j) = V. \tag{3.65}$$

The union of subspaces in (3.64) is an "almost disjoint" union : the $m$ subspaces have the zero vector $\mathbf{0}$ as their single common element. Therefore,

$$(\forall\, x \in V^{*})(\exists!\, j \in \overline{1,m})\ x \in W(\lambda_j) \ \Rightarrow\ L x = \lambda_j x. \tag{3.65}$$

Each eigen-subspace $W(\lambda_j)$ with $\dim W(\lambda_j) = h_j = k_j$ is spanned by a specific basis

$$B_j = [\, b_{j_1} \; b_{j_2} \; \ldots \; b_{j_{k_j}} \,] \;\Rightarrow\; L\, B_j^{\mathrm{T}} = \lambda_j\, B_j^{\mathrm{T}}. \tag{3.66}$$

The $m$ equations of the form (3.65) can be assembled into a single matrix equation involving the basis of the whole space $V$, $B = [B_1 \; \ldots \; B_j \; \ldots \; B_m]$, so that

$$L\,B^{\mathrm{T}} = \begin{bmatrix} \lambda_1\, B_1^{\mathrm{T}} \\ \vdots \\ \lambda_j\, B_j^{\mathrm{T}} \\ \vdots \\ \lambda_m\, B_m^{\mathrm{T}} \end{bmatrix} = \lceil \lambda_1 \ldots \lambda_1 / \ldots / \lambda_j \ldots \lambda_j / \ldots / \lambda_m \ldots \lambda_m \rfloor \cdot B^{\mathrm{T}}. \tag{3.67}$$

In the rightmost side of (3.67) we have used the symbol "/ " for separating the $m$ diagonal submatrices, each of them consisting of exactly $k_j$ "copies" of the eigenvalue $\lambda_j$ $(1 \le j \le m)$ on its diagonal. Thus, this part of the proof is complete.

$(\Rightarrow)$: If the spectrum of $L$ is $\sigma(L) = \{\lambda_1, \lambda_2, \ldots, \lambda_m\} \subset \mathbf{K}$ and $L$ is diagonalizable, it satisfies *Definition 3.3* - Eq. (3.43). On another hand, the general structure of the characteristic polynomial $P_L(\lambda)$ of (3.40) implies the existence of the $m$ eigen-subspaces in (3.58), with

$$(\forall\, j \in \{1, 2, \ldots, m\})\ \dim W(\lambda_j) = h_j \le k_j. \tag{3.68}$$

But the subspaces in (2.58) are pairwise disjoint, up to the zero vector $\mathbf{0}$ (as earlier argued). Their disjoint union, as in Eq. (3.64), covers the whole space $V$. In view of *Definition 3.3* - Eq. (3.43), any vector $x \in V^{\star}$ admits a linear expression (in basis $B$) of the form

$$x = \sum_{i=1}^{n} \alpha_i\, b_i. \tag{3.69}$$

The sum in (3.69) can be split into $m$ subsums since each eigen-subspace of (3.58) admits its own sub-basis, as we have earlier written $B = [B_1 \; \ldots \; B_j \; \ldots \; B_m]$: $W(\lambda_j) = \mathcal{L}(B_j)$. Therefore

$$x = \sum_{i=1}^{h_1} \alpha_i\, b_i + \ldots + \sum_{i=1}^{h_j} \beta_i\, b_i + \ldots + \sum_{i=1}^{h_m} \gamma_i\, b_i. \tag{3.70}$$

This expression (3.70) shows that $V$ is *the sum* of the $m$ eigen-spaces of (3.58).

On another hand, the trivial intersections in (3.62) plus a consequence of GRASSMANN's THEOREM on the dimensions of two subspaces, of their sum and their intersection implies property $(ii)$ in the statement and also the equality

$$\sum_{j=1}^{m} h_j = \sum_{j=1}^{m} \dim W(\lambda_j) = \dim\, [(W(\lambda_1) + W(\lambda_2) + \ldots + W(\lambda_m)] = \dim V = n. \tag{3.71}$$

Eq. (3.71) with the inequality (3.42) leads to

$$n = \sum_{j=1}^{m} h_j \le \sum_{j=1}^{m} k_j = n \;\Rightarrow\; (\forall\, j \in \{1, 2, \ldots, m\})\ h_j = k_j$$

and $(i)$ is thus entirely proved.

As regards part $(ii)$, it holds in view of the intersection in (3.62) – a consequence of PROPOSITION 3.2 : *the sum of the eigen-spaces* (respectively) *corresponding to distinct eigenvalues is a direct sum.* In the textbook [C. Radu, 1986], this property appears as part $2^\circ$ in **Theorem 1.1** at page 57, but it is stated for $p$ distinct eigenvalues and not necessarily for all eigenvalues of the endomorphism $V$. Prof. C. Radu's proof (at page 58) is developed by induction with respect to $p$. It is possibly simpler than ours, but our proof has included the structure of the "canonical" basis $B$,

$$B = [B_1 \ \dots \ B_j \ \dots \ B_m] : \ W(\lambda_j) = \mathcal{L}(B_j),$$

as well as the structure of the diagonal matrix – in Eq. (3.64). Thus, the equation (3.64) can now be completed :

$$\sum_{j=1}^{m} W(\lambda_j) = W(\lambda_1) \oplus W(\lambda_2) \oplus \dots \oplus W(\lambda_m) = V. \qquad \blacksquare$$

**COROLLARY 3.1.** (**Sufficient Conditions for Diagonalization**). *If the spectrum of an endomorphism* $L : V \longrightarrow V$ $(\dim V = n)$ *is*

$$\sigma(L) = \{\lambda_1, \lambda_2, \dots, \lambda_n\} \subset \mathbf{K} \tag{3.72}$$

*then the endomorphism is diagonalizable.*

Indeed, this is an immediate consequence THEOREM 3.2. Let us only see that (3.72) implies

$$(\forall j \in \{1, 2, \dots, m\}) \dim W(\lambda_j) = h_j = k_j = 1 .$$

Each sub-basis $B_j$ consists of a single eigenvector : $B_j = \{u_j\}$. As regards the matrix $L_B$, it consists of the $n$ distinct eigenvalues on its diagonal :

$$L_B = \lceil \lambda_1 \ \ \lambda_2 \ \dots \ \lambda_n \rfloor . \tag{3.73}$$

Another sufficient condition for the possibility to diagonalize an operator can be expressed in terms of the geometric multiplicities of its eigenvalues, as discussed in the proof of the previous THEOREM :

$$\sum_{j=1}^{m} h_j = n \ \Rightarrow \ m = n, \ h_1 = h_2 = \dots = h_m = 1 \ \text{ and } \ L \ \text{ is diagonalizable.}$$

Another aspect regards the most often met type of operators met in practical applications, the linear maps of the form $L : \mathbf{K}^n \longrightarrow \mathbf{K}^n$.

A series of examples follow, able to illustrate various situations that can occur when the problem of operators' diagonalization is approached. All the theoretical issues so far presented are applicable, but – for instance – the homogeneous systems producing the eigenvectors $U_j \in \operatorname{Ker}(L - \lambda_j \,\mathrm{id})$ can be obtained from the solution(s) of homogeneous systems of the form

$$\boxed{(L_E^{\mathrm{T}} - \lambda_j I_n) \cdot U = \mathbf{0} \underset{\text{not}}{\Longleftrightarrow} (M - \lambda_j I_n) \cdot U = \mathbf{0} .} \tag{3.74}$$

**Example 3.3.** The endomorphism $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is an involution. Check that its eigenvalues are $\lambda_{1,2} = \pm 1$.

Let us recall, from *Definition 2.8* in § **4.2** (page 190), that an involution is defined by $L^2 = \mathrm{id}$, in this case $L^2 = \mathrm{id}_{\mathbb{R}^2}$ or $L \circ L = \mathbf{1}_{\mathbb{R}^2}$. For a vector $x \neq \mathbf{0}$,

$$Lx = \lambda x \;\Rightarrow\; x = L^2 x = L(Lx) = L(\lambda x) = \lambda^2 x \;\Rightarrow\; (\lambda^2 - 1)x = \mathbf{0} \;\Rightarrow$$

$$\Rightarrow\; \lambda^2 - 1 = \mathbf{0} \;\Rightarrow\; \lambda = \pm 1. \qquad \square$$

**Examples 3.4.** Find the EVVs (eigenvalues and eigenvectors / eigen-subspaces) of the next 4 operators $L_i : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ $(i = \overline{1,3})$ and $L_4 : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$, given by their matrices, and check which of them is / are diagonalizable.

$$L_1 :\; M_1 = \begin{bmatrix} -6 & 2 & -5 \\ -4 & 4 & -2 \\ 10 & -3 & 8 \end{bmatrix}; \qquad L_2 :\; M_2 = \begin{bmatrix} -8 & -13 & -14 \\ -6 & -5 & -8 \\ 14 & 17 & 21 \end{bmatrix};$$

$$L_3 :\; M_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 15 \\ 0 & 2 & 8 \end{bmatrix}; \qquad L_4 :\; M_4 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

**1°** $$P_{L_1}(\lambda) = \det(M_1 - \lambda I_3) = \begin{vmatrix} -6-\lambda & 2 & -5 \\ -4 & 4-\lambda & -2 \\ 10 & -3 & 8-\lambda \end{vmatrix} = \begin{vmatrix} 2-\lambda & 2\lambda-6 & -1 \\ -4 & 4-\lambda & -2 \\ 2 & 5-2\lambda & 4-\lambda \end{vmatrix} = \ldots$$

$$\ldots = -\begin{vmatrix} \lambda-2 & 6-2\lambda & 1 \\ 2\lambda-8 & 16-5\lambda & 0 \\ \lambda^2-6\lambda+10 & -2\lambda^2+12\lambda-19 & 0 \end{vmatrix} = \ldots = -(\lambda^3 - 6\lambda^2 + 12\lambda - 8) =$$

$$= -(\lambda-2)^3 \;\Rightarrow\; \lambda_1 = \lambda_2 = \lambda_3 = 2. \tag{3.75}$$

$$(3.71) \;\Rightarrow\; M_1 - 2I_3 = \begin{bmatrix} -8 & 2 & -5 \\ -4 & 2 & -2 \\ 10 & -3 & 6 \end{bmatrix} \sim \begin{bmatrix} -8 & 2 & -5 \\ -4 & 2 & -2 \\ 2 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & -3 & 0 \\ 0 & 0 & 0 \\ 2 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -3/2 & 0 \\ 0 & 2 & 1 \end{bmatrix} \Rightarrow$$

$$\Rightarrow\; U(\alpha) = \begin{bmatrix} 3\alpha \\ 2\alpha \\ -4\alpha \end{bmatrix} \;\Rightarrow\; \dim W(\lambda_{1,2,3}) = 1 = h < k = 3.$$

Therefore, $L_1$ is not diagonalizable. $\square$

**2°** For the second operator we proceed similarly but we offer less calculation details. The characteristic polynomial is

$$P_{L_2}(\lambda) = \det(M_2 - \lambda I_3) = \begin{vmatrix} -8-\lambda & -13 & -14 \\ -6 & -5-\lambda & -8 \\ 14 & 17 & 21-\lambda \end{vmatrix} = \begin{vmatrix} 5-\lambda & 1 & -14 \\ \lambda-1 & 3-\lambda & -8 \\ -3 & \lambda-4 & 21-\lambda \end{vmatrix} = \ldots$$

$$\ldots = \begin{vmatrix} 5-\lambda & 1 & -14 \\ 4 & 4-\lambda & -22 \\ 1 & 0 & -1-\lambda \end{vmatrix} = \ldots = -(\lambda-2)(\lambda-3)^2 \Rightarrow \lambda_1 = 2 \ \& \ \lambda_2 = \lambda_3 = 3. \tag{3.76}$$

For the double eigenvalue in (3.76), the matrix of the corresponding homogeneous system is

$$M_2 - 3I_3 = \begin{bmatrix} -11 & -13 & -14 \\ -6 & -8 & -8 \\ 14 & 17 & 18 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 2 \\ 0 & 5 & 2 \\ 0 & -25 & -10 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 5/2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 0 \\ 0 & 5/2 & 1 \end{bmatrix} \Rightarrow$$

$$\Rightarrow U(\alpha) = \begin{bmatrix} 4\alpha \\ 2\alpha \\ -5\alpha \end{bmatrix} \Rightarrow \dim W(\lambda_{2,3}) = 1 = h_2 < k_2 = 2.$$

Hence, $L_2$ is not diagonalizable, too. □

$$\textbf{3°} \quad P_{L_3}(\lambda) = \det(M_3 - \lambda I_3) = \begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & -3-\lambda & 15 \\ 0 & 2 & 8-\lambda \end{vmatrix} = \begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & -3-\lambda & 15 \\ 0 & 2 & 8-\lambda \end{vmatrix} =$$

$$= (1-\lambda)(\lambda^2 - 5\lambda - 54) \Rightarrow \lambda_1 = 1,\ \lambda_2 = (5 - \sqrt{241})/2,\ \lambda_3 = (5 + \sqrt{241})/2. \tag{3.77}$$

It follows from (3.77) and COROLLARY 3.1 that $L_3$ *is diagonalizable*. Since it is the first operator (so far met, in this section) that can be diagonalized, we are going to determine its eigenvectors and its diagonal matrix (in the basis consisting of the three eigenvectors). For each eigenvalue in (3.77) we have to solve the corresponding homogeneous system. We present these calculations (in detail) taking into account that it is a little more difficult to operate with *irrational eigenvalues*, hence with irrational entries in the matrices of the systems to be solved.

$$M_3 - \lambda_1 I_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -4 & 15 \\ 0 & 2 & 7 \end{bmatrix} \sim \begin{bmatrix} 0 & 2 & 7 \\ 0 & -4 & 15 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 7/2 \\ 0 & 0 & 29 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow U_1(\alpha) = \begin{bmatrix} \alpha \\ 0 \\ 0 \end{bmatrix};$$

$$M_3 - \lambda_2 I_3 = \begin{bmatrix} -(3+\sqrt{241})/2 & 0 & 0 \\ 0 & (\sqrt{241}-11)/2 & 15 \\ 0 & 2 & (\sqrt{241}+11)/2 \end{bmatrix}. \tag{3.78}$$

It can be seen that the second and the third rows in the matrix of (3.78) are proportional, hence one of them may be deleted. It suffices to check that the south-eastern minor of order 2 in (3.78) is = 0. Hence, this matrix is equivalent to

$$\frac{1}{2}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 11+\sqrt{241} \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & (11+\sqrt{241})/4 \end{bmatrix} \Rightarrow$$

$$\Rightarrow\ U_2(\beta) = \begin{bmatrix} 0 \\ -(\sqrt{241}+11)\,\beta \\ 4\beta \end{bmatrix} = -\beta \begin{bmatrix} 0 \\ \sqrt{241}+11 \\ -4 \end{bmatrix}. \tag{3.79}$$

Finally, the matrix of the third H-system can be obtained from the matrix in (3.78) by taking the conjugates of the irrational entries :

$$M_3 - \lambda_3 I_3 = \begin{bmatrix} -(3+\sqrt{241})/2 & 0 & 0 \\ 0 & -(11+\sqrt{241})/2 & 15 \\ 0 & 2 & (11-\sqrt{241})/2 \end{bmatrix} \sim \ldots$$

$$\ldots \sim \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 11-\sqrt{241} \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & (11-\sqrt{241})/4 \end{bmatrix} \Rightarrow$$

$$\Rightarrow\ U_3(\gamma) = \begin{bmatrix} 0 \\ (\sqrt{241}-11)\,\gamma \\ 4\gamma \end{bmatrix} = \gamma \begin{bmatrix} 0 \\ \sqrt{241}-11 \\ 4 \end{bmatrix}. \tag{3.80}$$

Taking $\alpha = 1,\ \beta = -1,\ \gamma = 1$ in the three eigenvectors just found we arrive to the (particular) eigenvectors

$$U_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\ U_2 = \begin{bmatrix} 0 \\ \sqrt{241}+11 \\ -4 \end{bmatrix},\ U_3 = \begin{bmatrix} 0 \\ \sqrt{241}-11 \\ 4 \end{bmatrix}. \tag{3.81}$$

According to the theoretical discussion preceding these examples, the three vectors in (3.81) should form the basis $B$ in which the matrix of $L_3$ is expected to be diagonal, namely $\lceil \lambda_1\ \ \lambda_2\ \ \lambda_3 \rfloor$. It is possible to check this conjecture as follows : The three eigenvectors of (3.77) should be assembled to form the similarity matrix $S = [U_1\ \ U_2\ \ U_3]$ of Eq. (3.48) at page 206, for the similarity of two arbitrary matrices (not necessarily connected with linear operators), respectively of Eq. (3.50) at page 206. In fact, Eq. (3.47) is more convenient to check since it does not require to invert the matrix $S$: it comes to equation $A \cdot S = S \cdot \lceil \lambda_1\ \ \lambda_2\ \ \lambda_3 \rfloor$ with $A \rightarrow M_3 \Rightarrow$

$$\boxed{M_3 \cdot S = S \cdot \lceil \lambda_1\ \ \lambda_2\ \ \lambda_3 \rfloor.} \tag{3.82}$$

Hence we get, from (3.81),

$$S = [U_1 \ U_2 \ U_3] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{241}+11 & \sqrt{241}-11 \\ 0 & -4 & 4 \end{bmatrix}. \tag{3.83}$$

$$M_3 \cdot S \underset{(3.83)}{=} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 15 \\ 0 & 2 & 8 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{241}+11 & \sqrt{241}-11 \\ 0 & -4 & 4 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3(\sqrt{241}+11)-60 & -3(\sqrt{241}-11)+60 \\ 0 & 2(\sqrt{241}+11)-32 & 2(\sqrt{241}-11)+32 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3\sqrt{241}-93 & -3\sqrt{241}+93 \\ 0 & 2\sqrt{241}-10 & 2\sqrt{241}+10 \end{bmatrix}. \tag{3.84}$$

$$S \cdot \lceil \lambda_1 \ \lambda_2 \ \lambda_3 \rfloor \underset{(3.77,83)}{=} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{241}+11 & \sqrt{241}-11 \\ 0 & -4 & 4 \end{bmatrix} \cdot \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5-\sqrt{241} & 0 \\ 0 & 0 & 5+\sqrt{241} \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -93-3\sqrt{241} & 93-3\sqrt{241} \\ 0 & -10+2\sqrt{241} & 10+2\sqrt{241} \end{bmatrix}. \tag{3.85}$$

It follows from (3.84) and (3.85) that equation (3.82) is satisfied. Therefore, this operator $L_3$ has the diagonal matrix with the eigenvalues of (3.77), in the basis whose vectors are the columns of the matrix in (3.83). □

$$\textbf{4°} \quad L_4 : M_4 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \Rightarrow P_{L_4}(\lambda) = \begin{vmatrix} -\lambda & 1 & 0 & 0 \\ 0 & -\lambda & 1 & 0 \\ 0 & 0 & -\lambda & 1 \\ 1 & 0 & 0 & -\lambda \end{vmatrix} = \lambda^4 - 1. \tag{3.86}$$

The characteristic equation is, by (3.86),

$$\lambda^4 - 1 = 0 \ \Rightarrow \ (\lambda^2 - 1)(\lambda^2 + 1) = 0 \ \Rightarrow$$

$$\Rightarrow \ \lambda_1 = -1, \ \lambda_2 = 1, \ \lambda_3 = -i, \ \lambda_4 = i \ \ (i^2 = -1). \tag{3.87}$$

Therefore, the spectrum of this endomorphism is

$$\sigma(L_4) = \{-1, 1, -i, i\} \subset \mathbb{C} \ \text{ with } \ \{-i, i\} \subset \mathbb{C} \setminus \mathbb{R}.$$

Hence $\sigma(L_4) \not\subset \mathbb{R}$ and one of the conditions in the statement of THEOREMS 3.1 & 3.2, namely $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbf{K}$, is not satisfied and $L_4$ is not diagonalizable. □

We shall see, in what follows, that an endomorphism over the real field $\mathbb{R}$ with some complex eigenvalues admits a matrix, in a certain basis consisting of eigenvectors, which is close to a diagonal matrix. But this falls in what we call the **Normal Forms** of operators, to be approached in the next subsection.. We offer one more example of an endomorphism *defined over* $\mathbb{C}$, with its eigenvectors and diagonal matrix.

**Example 3.5.** The endomorphism $L : \mathbb{C}^3 \longrightarrow \mathbb{C}^3$ is defined by (the transpose of) its matrix in the standard basis $E_3$

$$L_E^{\mathrm{T}} = M = \begin{bmatrix} -1 & -1 & 1 \\ 2 & 1 & 2i \\ 1+i & 0 & 0 \end{bmatrix}. \tag{3.88}$$

It is required to find its matrix in the basis $B = [(i, 1, -1)\ (-2, i, 0)\ (2i, 1, i)]$ and to determine

$$L(i\, b_1 + (2-i)\, b_2 + 4\, b_3) \tag{3.89}$$

using both the standard basis and basis $B$.

The standard basis of the complex Euclidean space $\mathbb{C}^n$ is the same with the well-known $E_n$ of the space $\mathbb{R}^n$. Hence the transformation matrix for $E_n \rightarrow B$ is

$$T = B^{\mathrm{T}} = \begin{bmatrix} i & 1 & -1 \\ -2 & i & 0 \\ 2i & 1 & i \end{bmatrix}. \tag{3.90}$$

In order to apply the matrix change formula (2.81) in **§ 4.2**, this matrix in (3.90) has to be inverted. The Gaussian elimination technique is the most convenient to be applied.

$$[T \mid I_3] = \left[\begin{array}{ccc|ccc} i & 1 & -1 & 1 & 0 & 0 \\ -2 & i & 0 & 0 & 1 & 0 \\ 2i & 1 & i & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} -1 & 0 & i & -i & 1 & 0 \\ i & 1 & -1 & 1 & 0 & 0 \\ i & 0 & 1+i & -1 & 0 & 1 \end{array}\right] \sim$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & -i & i & -1 & 0 \\ 0 & 1 & -2 & 2 & i & 0 \\ 0 & 0 & i & 0 & i & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & -i & i & -1 & 0 \\ 0 & 1 & -2 & 2 & i & 0 \\ 0 & 0 & 1 & 0 & 1 & -i \end{array}\right] \sim$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & i & i-1 & 1 \\ 0 & 1 & 0 & 2 & 2+i & -2i \\ 0 & 0 & 1 & 0 & 1 & -i \end{array}\right] \Rightarrow T^{-1} = \begin{bmatrix} i & i-1 & 1 \\ 2 & 2+i & -2i \\ 0 & 1 & -i \end{bmatrix}. \tag{3.91}$$

The reader can check that the inverse in (3.91) is correct : $T \cdot T^{-1} = I_3$.

$$(3.88) \;\Rightarrow\; L_E = \begin{bmatrix} -1 & 2 & 1+i \\ -1 & 1 & 0 \\ 1 & 2i & 0 \end{bmatrix}. \tag{3.92}$$

(3.88) , (3.92) & (3.91) $\Rightarrow$

$$\Rightarrow\; T \cdot L_E \cdot T^{-1} = \begin{bmatrix} i & 1 & -1 \\ -2 & i & 0 \\ 2i & 1 & i \end{bmatrix} \cdot \begin{bmatrix} -1 & 2 & 1+i \\ -1 & 1 & 0 \\ 1 & 2i & 0 \end{bmatrix} \cdot \begin{bmatrix} i & i-1 & 1 \\ 2 & 2+i & -2i \\ 0 & 1 & -i \end{bmatrix} =$$

$$= \begin{bmatrix} -2-i & 1 & i-1 \\ 2-i & i-4 & -2-2i \\ -1-i & 4i-1 & 2i-2 \end{bmatrix} \cdot \begin{bmatrix} i & i-1 & 1 \\ 2 & 2+i & -2i \\ 0 & 1 & -i \end{bmatrix} = \begin{bmatrix} 3-2i & 4+i & -1-2i \\ -7+4i & -12-i & 2+9i \\ -1+7i & -6+9i & 9+3i \end{bmatrix} = L_B. \tag{3.91}$$

With the matrix in (3.91) we can find the images of the vectors in basis $B$, for instance $Lb_1$.

$$Lb_1 = M \cdot b_1 = \begin{bmatrix} -1 & -1 & 1 \\ 2 & 1 & 2i \\ 1+i & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} i \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -2-i \\ 1 \\ -1+i \end{bmatrix}. \tag{3.92}$$

If we want to work in basis $B$, we can apply formula (2.15) in § **4.2** (page 166), written for basis $B$: $Lx = X_B^{\mathrm{T}} \cdot L_B \cdot B^{\mathrm{T}}$. But it is simpler to apply the definition of the matrix of an endomorphism in a basis, that is $L\,B^{\mathrm{T}} = L_B \cdot B^{\mathrm{T}}$. Let us check the image for the first vector, $b_1 = [i\;\;1\;\;-1]^{\mathrm{T}}$:

$$L\,B^{\mathrm{T}} = L_B \cdot B^{\mathrm{T}} \;\Rightarrow\; L\,b_1 = (L_B)_1 \cdot B^{\mathrm{T}} = [3-2i\quad 4+i\quad -1-2i\,] \cdot B^{\mathrm{T}} =$$

$$= (3-2i)\begin{bmatrix} i \\ 1 \\ -1 \end{bmatrix} + (4+i)\begin{bmatrix} -2 \\ i \\ 0 \end{bmatrix} + (-1-2i)\begin{bmatrix} 2i \\ 1 \\ i \end{bmatrix} = \begin{bmatrix} 3i+2-8-2i+4-2i \\ 3-2i+4i-1-2i-1 \\ -3+2i+2-i \end{bmatrix} = \begin{bmatrix} -2-i \\ 1 \\ -1+i \end{bmatrix}.$$

Hence, the image of $b_1$, found in (3.92), has been retrieved. The images $Lb_2, Lb_3$ can be similarly checked .

The image of the vector in (3.89) can be found (using basis $B$) with matrix $L_B$ of (3.91) and the coordinates resulting from (3.89) :

$$X_B = \begin{bmatrix} i \\ 2-i \\ 4 \end{bmatrix} \;\Rightarrow\; LX = [\,i\quad 2-i\quad 4\,] \cdot \begin{bmatrix} 3-2i & 4+i & -1-2i \\ -7+4i & -12-i & 2+9i \\ -1+7i & -6+9i & 9+3i \end{bmatrix} \cdot B^{\mathrm{T}} = \ldots$$

$$\ldots = [-12+46i\quad -50+50i\quad 51+27i\,] \cdot B^{\mathrm{T}} =$$

$$= (-12+46i)\,b_1 + (-50+50i)\,b_2 + (51+27i)\,b_3. \tag{3.93}$$

This expression of $Lx$ allows to write it as a (column) vector in $\mathbb{C}^3$ ; but it is appropriate to write the vectors of $B$ as column vectors, for an easier calculation :

$$B: b_1 = \begin{bmatrix} i \\ 1 \\ -1 \end{bmatrix},\ b_2 = \begin{bmatrix} -2 \\ i \\ 0 \end{bmatrix},\ b_3 = \begin{bmatrix} 2i \\ 1 \\ i \end{bmatrix} \Rightarrow \tag{3.94}$$

$$\Rightarrow LX = (-12+46i)\begin{bmatrix} i \\ 1 \\ -1 \end{bmatrix} + (-50+50i)\begin{bmatrix} -2 \\ i \\ 0 \end{bmatrix} + (51+27i)\begin{bmatrix} 2i \\ 1 \\ i \end{bmatrix} =$$

$$= \begin{bmatrix} -12i - 46 + 100 - 100i + 102i - 54 \\ -12 + 46i - 50i - 50 + 51 + 27i \\ 12 - 46i + 51i - 27 \end{bmatrix} = \begin{bmatrix} -10i \\ -11 + 23i \\ -15 + 5i \end{bmatrix}. \tag{3.95}$$

The same image $LX$ can be found, as a vector in $\mathbb{C}^3$, using the (transpose) of $L$'s matrix in the standard basis, that is the matrix in (3.88). But the vector $X \in \mathbb{C}^3$ should be previously found, with the basis in (3.94) and its coordinates in (3.89) :

$$X = i\,b_1 + (2-i)\,b_2 + 4\,b_3 = i\begin{bmatrix} i \\ 1 \\ -1 \end{bmatrix} + (2-i)\begin{bmatrix} -2 \\ i \\ 0 \end{bmatrix} + 4\begin{bmatrix} 2i \\ 1 \\ i \end{bmatrix} =$$

$$= \begin{bmatrix} -1 - 4 + 2i + 8i \\ i + 2i + 1 + 4 \\ -i + 4i \end{bmatrix} = \begin{bmatrix} -5 + 10i \\ 5 + 3i \\ 3i \end{bmatrix}. \tag{3.96}$$

Eqs. (3.88) & (3.96) $\Rightarrow$

$$\Rightarrow LX = \begin{bmatrix} -1 & -1 & 1 \\ 2 & 1 & 2i \\ 1+i & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} -5 + 10i \\ 5 + 3i \\ 3i \end{bmatrix} = \begin{bmatrix} -10i \\ -11 + 23i \\ -15 + 5i \end{bmatrix}.$$

Thus, the image of (3.95) has been retrieved and the example is completed. □

**Note.** Two cases when an operator cannot be diagonalized were earlier presented. It comes to the situation when $\sigma(L) = \{\lambda_1, \lambda_2, \dots, \lambda_m\} \not\subset \mathbf{K} : (\exists\, j \in \overline{1,m})\ \lambda_j \notin \mathbf{K}$. A typical case is the one when the operator is defined over a *real* vector space but the characteristic polynomial admits at least one complex eigenvalue, $\lambda_j \in \mathbb{C} \setminus \mathbb{R}$. The other case regards the condition in point $(i)$ of THEOREM 3.2, that is $\dim W(\lambda_j) = h_j = k_j$ for all $\lambda_j \in \sigma(L)$. In the former case, an "almost diagonal" form can be obtained, with 2-by-2 cells on the diagonal of the canonical matrix, corresponding to each pair of complex-conjugate roots $\lambda_j = \alpha_j \pm i\,\beta_j$. In the latter case, a so-called JORDAN NORMAL FORM of operator's matrix can be found. Theoretical issues and examples on these cases can be found in ⟦𝒜.𝒞., 1999⟧ , pages 166-180, in ⟦𝒜.𝒞., 2014⟧ and in other textbooks of LINEAR ALGEBRA.

## § 4.3-A APPLICATIONS TO ENDOMORPHISMS: EVVS - NORMAL FORMS

**3 – A.1** The endomorphism $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ is given by its matrix in the standard basis $E$,

$$L_E = \begin{bmatrix} 5 & 6 & -3 \\ -1 & 0 & -1 \\ 1 & 2 & -1 \end{bmatrix}.$$

It is required to find its (three) eigenvalues and the corresponding eigenvectors ; then write its diagonal matrix in the basis consisting of these vectors and check it.

**3 – A.2** Find the eigenvalues and the corresponding eigenvectors for the operators of the form $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ given by their (transposed) matrices in the corresponding standard bases :

$$a)\ \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}; \qquad b)\ \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}; \qquad c)\ \begin{bmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{bmatrix};$$

$$d)\ \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

**3 – A.3** Study the possibility for the following matrices (representing linear endomorphisms $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$) to be diagonalized :

$$a)\ \begin{bmatrix} 5 & 2 & 3 \\ 2 & -1 & 0 \\ 3 & 0 & 1 \end{bmatrix}; \quad b)\ \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}; \quad c)\ \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}; \quad d)\ \begin{bmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{bmatrix}.$$

**3 – A.4** Check that the matrix given below satisfies its characteristic equation, that is $P_A(A) = \mathbf{0}$, where $P_A(\lambda) = \det(A - \lambda I_4)$ :

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

*Hint* : The result follows from the Cayley-Hamilton Theorem. But the equation $P_A(A) = \mathbf{0}$ can be easier checked after the factorization of the characteristic polynomial into linear factors.