NOTES ON LINEAR ALGEBRA

LIVIU I. NICOLAESCU

CONTENTS

Date: Started January 7, 2013. Completed on . Last modified on **April 28, 2014**.
These are notes for the Honors Algebra Course, Spring 2013.

1. MULTILINEAR FORMS AND DETERMINANTS

In this section, we will deal exclusively with *finite dimensional* vector spaces over the field $\mathbb{F} = \mathbb{R}, \mathbb{C}$. If $\boldsymbol{U}_1, \boldsymbol{U}_2$ are two $\mathbb{F}$-vector spaces, we will denote by $\operatorname{Hom}(\boldsymbol{U}_1, \boldsymbol{U}_2)$ the space of $\mathbb{F}$-linear maps $\boldsymbol{U}_1 \to \boldsymbol{U}_2$.

1.1. **Mutilinear maps.**

Definition 1.1. Suppose that $\boldsymbol{U}_1, \dots, \boldsymbol{U}_k, \boldsymbol{V}$ are $\mathbb{F}$-vector spaces. A map

$$\Phi : \boldsymbol{U}_1 \times \cdots \times \boldsymbol{U}_k \to \boldsymbol{V}$$

is called k-linear if for any $1 \leq i \leq k$, any vectors $\boldsymbol{u}_i, \boldsymbol{v}_i \in \boldsymbol{U}_i$, vectors $\boldsymbol{u}_j \in \boldsymbol{U}_j$, $j \neq i$, and any scalar $\lambda \in \mathbb{F}$ we have

$$\begin{gathered}
\Phi(\boldsymbol{u}_1, \dots, \boldsymbol{u}_{i-1}, \boldsymbol{u}_i + \boldsymbol{v}_i, \boldsymbol{u}_{i+1}, \dots, \boldsymbol{u}_k) \\
= \Phi(\boldsymbol{u}_1, \dots, \boldsymbol{u}_{i-1}, \boldsymbol{u}_i, \boldsymbol{u}_{i+1}, \dots, \boldsymbol{u}_k) + \Phi(\boldsymbol{u}_1, \dots, \boldsymbol{u}_{i-1}, \boldsymbol{v}_i, \boldsymbol{u}_{i+1}, \dots, \boldsymbol{u}_k), \\
\Phi(\boldsymbol{u}_1, \dots, \boldsymbol{u}_{i-1}, \lambda \boldsymbol{u}_i, \boldsymbol{u}_{i+1}, \dots, \boldsymbol{u}_k) = \lambda \Phi(\boldsymbol{u}_1, \dots, \boldsymbol{u}_{i-1}, \boldsymbol{u}_i, \boldsymbol{u}_{i+1}, \dots, \boldsymbol{u}_k).
\end{gathered}$$

In the special case $\boldsymbol{U}_1 = \boldsymbol{U}_2 = \cdots = \boldsymbol{U}_k = \boldsymbol{U}$ and $\boldsymbol{V} = \mathbb{F}$, the resulting map

$$\Phi : \underbrace{\boldsymbol{U} \times \cdots \times \boldsymbol{U}}_{k} \to \mathbb{F}$$

is called a k*-linear form* on $\boldsymbol{U}$. When $k = 2$, we will refer to 2-linear forms as *bilinear forms*. We will denote by $\mathcal{T}^k(\boldsymbol{U}^*)$ the space of k-linear forms on $\boldsymbol{U}$. □

Example 1.2. Suppose that $\boldsymbol{U}$ is an $\mathbb{F}$-vector space and $\boldsymbol{U}^*$ is its dual, $\boldsymbol{U}^* := \operatorname{Hom}(\boldsymbol{U}, \mathbb{F})$. We have a natural bilinear map

$$\langle -, - \rangle : \boldsymbol{U}^* \times \boldsymbol{U} \to \mathbb{F}, \;\; \boldsymbol{U}^* \times \boldsymbol{U} \ni (\alpha, \boldsymbol{u}) \mapsto \langle \alpha, \boldsymbol{u} \rangle := \alpha(\boldsymbol{u}).$$

The bilinear map is called the *canonical pairing* between the vector space $\boldsymbol{U}$ and its dual. □

Example 1.3. Suppose that $A = (a_{ij})_{1 \leq i,j \leq n}$ is an $n \times n$ matrix with real entries. Define

$$\Phi_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \;\; \Phi(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i,j} a_{ij} x_i y_j,$$

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \;\; \boldsymbol{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

To show that Φ is indeed a bilinear form we need to prove that for any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$ and any $\lambda \in \mathbb{R}$ we have

$$\Phi_A(\boldsymbol{x} + \boldsymbol{z}, \boldsymbol{y}) = \Phi_A(\boldsymbol{x}, \boldsymbol{y}) + \Phi_A(\boldsymbol{z}, \boldsymbol{y}), \tag{1.1a}$$

$$\Phi_A(\boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z}) = \Phi_A(\boldsymbol{x}, \boldsymbol{y}) + \Phi_A(\boldsymbol{x}, \boldsymbol{z}), \tag{1.1b}$$

$$\phi_A(\lambda \boldsymbol{x}, \boldsymbol{y}) = \Phi_A(\boldsymbol{x}, \lambda \boldsymbol{y}) = \lambda \Phi_A(\boldsymbol{x}, by). \tag{1.1c}$$

To verify (1.1a) we observe that

$$\begin{aligned}
\Phi_A(\boldsymbol{x} + \boldsymbol{z}, \boldsymbol{y}) = \sum_{i,j} a_{ij}(x_i + z_i) y_j = \sum_{i,j} (a_{ij} x_i y_j + a_{ij} z_i y_j) = \sum_{i,j} a_{ij} x_i y_j + \sum_{ij} a_{ij} z_i y_j \\
= \Phi_A(\boldsymbol{x}, \boldsymbol{y}) + \Phi_A(\boldsymbol{z}, \boldsymbol{z}).
\end{aligned}$$

The equalities (1.1b) and (1.1c) are proved in a similar fashion. Observe that if $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ is the natural basis of $\mathbb{R}^n$, then

$$\Phi_A(\boldsymbol{e}_i, \boldsymbol{e}_j) = a_{ij}.$$

This shows that Φ_A is completely determined by its action on the basic vectors $\boldsymbol{e}_1, \dots \boldsymbol{e}_n$. $\square$

Proposition 1.4. *For any bilinear form $\Phi \in \mathcal{T}^2(\mathbb{R}^n)$ there exists an $n \times n$ real matrix A such that $\Phi = \Phi_A$, where Φ_A is defined as in Example 1.3.* $\square$

The proof is left as an exercise.

1.2. **The symmetric group.** For any finite sets A, B we denote $\operatorname{Bij}(A, B)$ the collection of bijective maps $\varphi : A \to B$. We set $\mathcal{S}(A) := \operatorname{Bij}(A, A)$. We will refer to $\mathcal{S}(A)$ as the *symmetric group on A* and to its elements as *permutations of A*. Note that if $\varphi, \sigma \in \mathcal{S}(A)$ then

$$\varphi \circ \sigma, \;\; \varphi^{-1} \in \mathcal{S}(A).$$

The composition of two permutations is often referred to as the *product* of the permutations. We denote by $\mathbb{1}$, or $\mathbb{1}_A$ the *identity permutation* that does not permute anything, i.e., $\mathbb{1}_A(a) = a$, $\forall a \in A$.

For any finite set S we denote by $|S|$ its cardinality, i.e., the number of elements of S. Observe that

$$\operatorname{Bij}(A, B) \neq \emptyset \Longleftrightarrow |A| = |B|.$$

In the special case when A is the discrete interval $A = \boldsymbol{I}_n = \{1, \dots, n\}$ we set

$$\mathcal{S}_n := \mathcal{S}(\boldsymbol{I}_n).$$

The collection $\mathcal{S}_n$ is called the symmetric group on n objects. We will indicate the elements $\varphi \in \mathcal{S}_n$ by diagrams of the form

$$\left(\begin{array}{cccc} 1 & 2 & \dots & n \\ \varphi_1 & \varphi_2 & \dots & \varphi_n \end{array}\right).$$

For any finite set S we denote by $|S|$ its cardinality, i.e., the number of elements of S.

Proposition 1.5. *(a) If A, B are finite sets and $|A| = |B|$, then*

$$|\operatorname{Bij}(A, B)| = |\operatorname{Bij}(B, A)| = |\mathcal{S}(A)| = |\mathcal{S}(B)|.$$

(b) For any positive integer n we have $|\mathcal{S}_n| = n! := 1 \cdot 2 \cdot 3 \cdots n$.

Proof. (a) Observe that we have a bijective correspondence

$$\operatorname{Bij}(A, B) \ni \varphi \mapsto \varphi^{-1} \in \operatorname{Bij}(B, A)$$

so that

$$|\operatorname{Bij}(A, B)| = |\operatorname{Bij}(B, A)|.$$

Next, fix a bijection $\psi : A \to B$. We get a correspondence

$$F_\psi : \operatorname{Bij}(A, A) \to \operatorname{Bij}(A, B), \;\; \varphi \mapsto F_\psi(\varphi) = \psi \circ \varphi.$$

This correspondence is injective because

$$F_\psi(\varphi_1) = F_\psi(\varphi_2) \Rightarrow \psi \circ \varphi_1 = \psi \circ \varphi_2 \Rightarrow \psi^{-1} \circ (\psi \circ \varphi_1) = \psi^{-1} \circ (\psi \circ \varphi_2) \Rightarrow \varphi_1 = \varphi_2.$$

This correspondence is also surjective. Indeed, if $\phi \in \operatorname{Bij}(A, B)$ then $\psi^{-1} \circ \phi \in \operatorname{Bij}(A, A)$ and

$$F_\psi(\psi^{-1} \circ \phi) = \psi \circ (\psi^{-1} \circ \phi) = \phi.$$

This, F_ψ is a bijection so that

$$|\mathcal{S}(A)| = |\operatorname{Bij}(A, B)|.$$

Finally we observe that

$$|\mathcal{S}(B)| = |\operatorname{Bij}(B, A)| = |\operatorname{Bij}(A, B)| = |\mathcal{S}(A)|.$$

This takes care of (a).

To prove (b) we argue by induction. Observe that $|\mathcal{S}_1| = 1$ because there exists a single bijection $\{1\} \to \{1\}$. We assume that $|\mathcal{S}_{n-1}| = (n-1)!$ and we prove that $\mathcal{S}_n| = n!$. For each $k \in \boldsymbol{I}_n$ we set

$$\mathcal{S}_n^k := \big\{ \varphi \in \mathcal{S}_n;\ \varphi(n) = k \big\}.$$

A permutation $\varphi \in \mathcal{S}_n^k$ is uniquely detetrimined by its restriction to $\boldsymbol{I}_n \setminus \{n\} = \boldsymbol{I}_{n-1}$ and this restriction is a bijection $\boldsymbol{I}_{n-1} \to \boldsymbol{I}_n \setminus \{k\}$. Hence

$$|\mathcal{S}_n^k| = |\operatorname{Bij}(\boldsymbol{I}_{n-1}, \boldsymbol{I}_n \setminus \{k\})| = |\mathcal{S}_{n-1}|,$$

where at the last equality we used part(a). We deduce

$$|\mathcal{S}_n| = |\mathcal{S}_n^1| + \cdots + |\mathcal{S}_n^n| = \underbrace{|\mathcal{S}_{n-1}| + \cdots + |\mathcal{S}_{n-1}|}_{n}$$

$$= n|\mathcal{S}_{n-1}| = n(n-1)!,$$

where at the last step we invoked the inductive assumption. □

Definition 1.6. An *inversion* of a permutation $\sigma \in \mathcal{S}_n$ is a pair $(i, j) \in \boldsymbol{I}_n \times \boldsymbol{I}_n$ with the following properties.

- $i < j$.
- $\sigma(i) > \sigma(j)$.

We denote by $|\sigma|$ the number of inversions of the permutation σ. The *signature* of σ is then the quantity

$$\operatorname{sign}(\sigma) := (-1)^{|\sigma|} \in \{-1, 1\}.$$

A permutation σ is called *even/odd* if $\operatorname{sign}(\sigma) = \pm 1$. We denote by $\mathcal{S}_n^{\pm}$ the collection of even/odd permulations. □

Example 1.7. (a) Consider the permutation $\sigma \in \mathcal{S}_5$ given by

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix}.$$

The inversions of σ are

$$(1,2),\ (1,3),\ (1,4),\ (1,5),$$
$$(2,3),\ (2,4),\ (2,5),$$
$$(3,4),\ (3,5),\ \ (4,5),$$

so that $|\sigma| = 4 + 3 + 2 + 1 = 10$, $\operatorname{sign}(\sigma) = 1$.

(b) For any $i \neq j$ in $\boldsymbol{I}_n$ we denote by τ_{ij} the permutation defined by the equalities

$$\tau_{ij}(k) = \begin{cases} k, & k \neq i, j \\ j, & k = i \\ i, & k = j. \end{cases}$$

A transposition is defined to be a permutation of the form τ_{ij} for some $i < j$. Observe that

$$|\tau_{ij}| = 2|j - i| - 1,$$

so that

$$\operatorname{sign}(\tau_{ij}) = -1, \ \ \forall i \neq j. \tag{1.2}$$

□

Proposition 1.8. *(a) For any $\sigma \in \mathcal{S}_n$ we have*

$$\operatorname{sign}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}. \tag{1.3}$$

(b) For any $\varphi, \sigma \in \mathcal{S}_n$ we have

$$\operatorname{sign}(\varphi \circ \sigma) = \operatorname{sign}(\varphi) \cdot \operatorname{sign}(\sigma). \tag{1.4}$$

(c) $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}(\sigma)$

Proof. (a) Observe that the ratio $\frac{\sigma(j)-\sigma(i)}{j-i}$ negative if and only if (i,j) is inversion. Thus the number of negative ratios $\frac{\sigma(j)-\sigma(i)}{j-i}$, $i < j$, is equal to the number of inversions of σ so that the product

$$\prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}$$

has the same sign as the signature of σ. Hence, to prove (1.3) it suffices to show that

$$\left| \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i} \right| = |\operatorname{sign}(\sigma)| = 1,$$

i.e.,

$$\prod_{i<j} |\sigma(j) - \sigma(i)| = \prod_{i<j} |j - i|. \tag{1.5}$$

This is now obvious because the factors in the left-hand side are exactly the factors in the right-hand side multiplied in a different order. Indeed, for any $i < j$ we can find a unique pair $i' < j'$ such that

$$\sigma(j') - \sigma(i') = \pm(j - i).$$

(b) Observe that

$$\operatorname{sign}(\varphi) = \prod_{i<j} \frac{\varphi(j) - \varphi(i)}{j - i} = \prod_{i<j} \frac{\varphi(\sigma(j)) - \varphi(\sigma(i)}{\sigma(j) - \sigma(i)}$$

and we deduce

$$\operatorname{sign}(\varphi) \cdot \operatorname{sign}(\sigma) = \left(\prod_{i<j} \frac{\varphi(\sigma(j)) - \varphi(\sigma(i))}{\sigma(j) - \sigma(i)} \right) \left(\prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i} \right)$$

$$\prod_{i<j} \frac{\varphi(\sigma(j)) - \varphi(\sigma(i))}{j - i} = \operatorname{sign}(\varphi \circ \sigma).$$

To prove (c) we observe that

$$1 = \operatorname{sign}(\mathbb{1}) = \operatorname{sign}(\sigma^{-1} \circ \sigma) = \operatorname{sign}(\sigma^{-1}) \operatorname{sign}(\sigma).$$

□

1.3. **Symmetric and skew-symmetric forms.**

Definition 1.9. Let $\boldsymbol{U}$ be an $\mathbb{F}$-vector space, $\boldsymbol{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$.
(a) A k-linear form $\Phi \in \mathcal{T}^k(\boldsymbol{U}^*)$ is called *symmetric* if for any $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k \in \boldsymbol{U}$, and any permutation $\sigma \in \mathcal{S}_k$ we have

$$\Phi(\boldsymbol{u}_{\sigma(1)}, \dots, \boldsymbol{u}_{\sigma(k)}) = \Psi(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k).$$

We denote by $S^k\boldsymbol{U}^*$ the collection of symmetric k-linear forms on $\boldsymbol{U}$.
(b) A k-linear form $\Phi \in \mathcal{T}^k(\boldsymbol{U}^*)$ is called *skew-symmetric* if for any $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k \in \boldsymbol{U}$, and any permutation $\sigma \in \mathcal{S}_k$ we have

$$\Phi(\boldsymbol{u}_{\sigma(1)}, \dots, \boldsymbol{u}_{\sigma(k)}) = \operatorname{sign}(\sigma)\Psi(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k).$$

We denote by $\Lambda^k\boldsymbol{U}^*$ the space of skew-symmetric k-linear forms on $\boldsymbol{U}$. □

Example 1.10. Suppose that $\Phi \in \Lambda^n\boldsymbol{U}^*$, and $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n \in \boldsymbol{U}$. The skew-linearity implies that for any $i < j$ we have

$$\Phi(\boldsymbol{u}_1, \dots, \boldsymbol{u}_{i-1}, \boldsymbol{u}_i, \boldsymbol{u}_{i+1}, \dots, \boldsymbol{u}_{j-1}, \boldsymbol{u}_j, \boldsymbol{u}_{j+1}, \dots, \boldsymbol{u}_n)$$
$$= -\Phi(\boldsymbol{u}_1, \dots, \boldsymbol{u}_{i-1}, \boldsymbol{u}_j, \boldsymbol{u}_{i+1}, \dots, \boldsymbol{u}_{j-1}, \boldsymbol{u}_i, \boldsymbol{u}_{j+1}, \dots, \boldsymbol{u}_n).$$

Indeed, we have

$$\Phi(\boldsymbol{u}_1, \dots, \boldsymbol{u}_{i-1}, \boldsymbol{u}_j, \boldsymbol{u}_{i+1}, \dots, \boldsymbol{u}_{j-1}, \boldsymbol{u}_i, \boldsymbol{u}_{j+1}, \dots, \boldsymbol{u}_n)$$
$$= \Phi(\boldsymbol{u}_{\tau_{ij}(1)}, \dots, \boldsymbol{u}_{\tau_{ij}(k)}, \dots, \boldsymbol{u}_{\tau_{ij}(n)})$$

and $\operatorname{sign}(\tau_{ij}) = -1$. In particular, this implies that if $i \neq j$, but $\boldsymbol{u}_i = \boldsymbol{u}_j$ then

$$\Phi(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n) = 0.$$ □

Proposition 1.11. *Suppose that $\boldsymbol{U}$ is an n-dimensional $\mathbb{F}$-vector space and $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ is a basis of $\boldsymbol{U}$. Then for any scalar $c \in \mathbb{F}$ there exists a unique skew-symmetric n-linear form $\Phi \in \Lambda^n\boldsymbol{U}^*$ such that*

$$\Phi(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n) = c.$$

Proof. To understand what is happening we consider first the special case $n = 2$. Thus $\dim \boldsymbol{U} = 2$. If $\Phi \in \Lambda^2\boldsymbol{U}^*$ and $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \boldsymbol{U}$ we can write

$$\boldsymbol{u}_1 = a_{11}\boldsymbol{e}_1 + a_{21}\boldsymbol{e}_2, \;\; \boldsymbol{u}_2 = a_{12}\boldsymbol{e}_1 + a_{22}\boldsymbol{e}_2,$$

for some scalars $a_{ij} \in \boldsymbol{F}$, $i, j \in \{1, 2\}$. We have

$$\Phi(\boldsymbol{u}_1, \boldsymbol{u}_2) = \Phi(a_{11}\boldsymbol{e}_1 + a_{21}\boldsymbol{e}_2, a_{12}\boldsymbol{e}_1 + a_{22}\boldsymbol{e}_2)$$
$$= a_{11}\Phi(\boldsymbol{e}_1, a_{12}\boldsymbol{e}_1 + a_{22}\boldsymbol{e}_2) + a_{21}\Phi(\boldsymbol{e}_2, a_{12}\boldsymbol{e}_1 + a_{22}\boldsymbol{e}_2)$$
$$= a_{11}a_{12}\Phi(\boldsymbol{e}_1, \boldsymbol{e}_1) + a_{11}a_{22}\Phi(\boldsymbol{e}_1, \boldsymbol{e}_2) + a_{21}a_{12}\Phi(\boldsymbol{e}_2, \boldsymbol{e}_1) + a_{21}a_{22}\Phi(\boldsymbol{e}_2, \boldsymbol{e}_2).$$

The skew-symmetry of Φ implies that

$$\Phi(\boldsymbol{e}_1, \boldsymbol{e}_1) = \Phi(\boldsymbol{e}_2, \boldsymbol{e}_2) = 0, \;\; \Phi(\boldsymbol{e}_2, \boldsymbol{e}_1) = -\Phi(\boldsymbol{e}_1, \boldsymbol{e}_2).$$

Hence

$$\Phi(\boldsymbol{u}_1, \boldsymbol{u}_2) = (a_{11}a_{22} - a_{21}a_{12})\Phi(\boldsymbol{e}_1, \boldsymbol{e}_2).$$

If $\dim \boldsymbol{U} = n$ and $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n \in \boldsymbol{U}$, then we can write

$$\boldsymbol{u}_1 = \sum_{i_1=1}^{n} a_{i_1 1}\boldsymbol{e}_{i_1}, \dots, \boldsymbol{u}_k = \sum_{i_k=1}^{n} a_{i_k k}\boldsymbol{e}_{i_k}$$

$$\Phi(\boldsymbol{u}_1,\dots,\boldsymbol{u}_n)=\Phi\left(\sum_{i_1=1}^n a_{i_11}\boldsymbol{e}_{i_1},\dots,\sum_{i_n=1}^n a_{i_nn}\boldsymbol{e}_{i_n}\right)$$
$$=\sum_{i_1,\dots,i_n=1}^n a_{i_11}\cdots a_{i_nn}\Phi(\boldsymbol{e}_{i_1},\dots,\boldsymbol{e}_{i_n}).$$

Observe that if the indices $i_1,\dots,i_n$ are not pairwise distinct then

$$\Phi(\boldsymbol{e}_{i_1},\dots,\boldsymbol{e}_{i_n})=0.$$

Thus, in the above sum we get contributions only from pairwise distinct choices of indices $i_1,\dots,i_n$, Such a choice corresponds to a permutation $\sigma\in\mathcal{S}_n$, $\sigma(k)=i_k$. We deduce that

$$\Phi(\boldsymbol{u}_1,\dots,\boldsymbol{u}_n)=\sum_{\sigma\in\mathcal{S}_n}a_{\sigma(1)1}\cdots a_{\sigma(n)n}\Phi(\boldsymbol{e}_{\sigma(1)},\dots,\boldsymbol{e}_{\sigma(n)}).$$
$$=\left(\sum_{\sigma\in\mathcal{S}_n}\operatorname{sign}(\sigma)a_{\sigma(1)1}\cdots a_{\sigma(n)n}\right)\Phi(\boldsymbol{e}_1,\dots,\boldsymbol{e}_n).$$

Thus, $\Phi\in\Lambda^n\boldsymbol{U}^*$ is uniquely determined by its value on $(\boldsymbol{e}_1,\dots,\boldsymbol{e}_n)$.

Conversely, the map

$$(\boldsymbol{u}_1,\dots,\boldsymbol{u}_n)\to c\sum_{\sigma\in\mathcal{S}_n}\operatorname{sign}(\sigma)a_{\sigma(1)1}\cdots a_{\sigma(n)n},\quad \boldsymbol{u}_k=\sum_{i=1}^n a_{ik}\boldsymbol{e}_i,$$

is indeed n-linear, and skew-symmetric. The proof is notationally bushy, but it does not involve any subtle idea so I will skip it. Instead, I'll leave the proof in the case $n=2$ as an exercise.

□

1.4. **The determinant of a square matrix.** Consider the vector space $\mathbb{F}^n$ we canonical basis

$$\boldsymbol{e}_1=\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix},\quad \boldsymbol{e}_2=\begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix},\dots,\boldsymbol{e}_n=\begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix}.$$

According to Proposition 1.11 there exists a unique, n-linear skew-symmetric form Φ on $\mathbb{F}^n$ such that

$$\Phi(\boldsymbol{e}_1,\dots,\boldsymbol{e}_n)=1.$$

We will denote this form by $\det$ and we will refer to it as the *determinant form* on $\mathbb{F}^n$. The proof of Proposition 1.11 shows that if $\boldsymbol{u}_1,\dots,\boldsymbol{u}_n\in\mathbb{F}^n$,

$$\boldsymbol{u}_k=\begin{bmatrix}u_{1k}\\u_{2k}\\\vdots\\u_{nk}\end{bmatrix},\quad k=1,\dots,n,$$

then

$$\det(\boldsymbol{u}_1,\dots\boldsymbol{u}_n)=\sum_{\sigma\in\mathcal{S}_n}\operatorname{sign}(\sigma)u_{\sigma(1)1}u_{\sigma(2)2}\cdots u_{\sigma(n)n}. \tag{1.6}$$

Note that

$$\det(\boldsymbol{u}_1,\dots\boldsymbol{u}_n)=\sum_{\varphi\in\mathcal{S}_n}\operatorname{sign}(\varphi)u_{1\varphi(1)}u_{2\varphi(2)}\cdots u_{n\varphi(n)}. \tag{1.7}$$

Definition 1.12. Suppose that $A = (a_{ij})_{1\leq i,j\leq n}$ is an $n \times n$-matrix with entries in $\mathbb{F}$ which we regard as a linear operator $A : \mathbb{F}^n \to \mathbb{F}^n$. The determinant of A is the scalar

$$\det A := \det(A\boldsymbol{e}_1, \dots, A\boldsymbol{e}_n)$$

where $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ is the canonical basis of $\mathbb{F}^n$, and $A\boldsymbol{e}_k$ is the k-th column of A,

$$A\boldsymbol{e}_k = \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{bmatrix}, \quad k = 1, \dots, n.$$

□

Thus, according to (1.6) we have

$$\det A = \sum_{\sigma\in\mathcal{S}_n} \operatorname{sign}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} \overset{(1.7)}{=} \sum_{\sigma\in\mathcal{S}_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}. \tag{1.8}$$

Remark 1.13. Consider a typical summand in the first sum in (1.8), $a_{\sigma(1)1} \cdots a_{\sigma(n)n}$. Observe that the n entries

$$a_{\sigma(1)1}, a_{\sigma(2)2}, \dots, a_{\sigma(n)n}$$

lie on different columns of A and thus occupy all the n columns of A. Similarly, these entries lie on different rows of A.

A collection of n entries so that no two lie on the same row or the same column is called a *rook placement*.[1] Observe that in order to describe a rook placement, you need to indicate the position of the entry on the first column, by indicating the row $\sigma(1)$ on which it lies, then you need to indicate the position of the entry on the second column etc. Thus, the sum in (1.8) has one term for each rook placement. □

If $A^\dagger$ denotes the transpose of the $n \times n$-matrix A with entries

$$a^\dagger_{ij} = a_{ji}$$

we deduce that

$$\det A^\dagger = \sum_{\sigma\in\mathcal{S}_n} \operatorname{sign}(\sigma) a^\dagger_{\sigma(1)1} \cdots a^\dagger_{\sigma(n)n} = \sum_{\sigma\in\mathcal{S}_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \det A. \tag{1.9}$$

Example 1.14. Suppose that A is a 2×2 matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Then

$$\det A = a_{11}a_{22} - a_{12}a_{21}.$$

□

Proposition 1.15. *If A is an upper triangular $n \times n$-matrix, then* $\det A$ *is the product of the diagonal entries. A similar result holds if A is lower triangular.*

[1] If you are familiar with chess, a rook controls the row and the column at whose intersection it is situated.

Proof. To keep the ideas as transparent as possible, we carry the proof in the special case $n = 3$. Suppose first that A is upper tringular, Then

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}$$

so that

$$A\boldsymbol{e}_1 = a_{11}\boldsymbol{e}_1, \quad A\boldsymbol{e}_2 = a_{12}\boldsymbol{e}_1 + a_{22}\boldsymbol{e}_2, \quad A\boldsymbol{e}_3 = a_{13}\boldsymbol{e}_1 + a_{23}\boldsymbol{e}_2 + a_{33}\boldsymbol{e}_3$$

Then

$$\begin{aligned} \det A &= \det(A\boldsymbol{e}_1, A\boldsymbol{e}_2, A\boldsymbol{e}_3) \\ &= \det(a_{11}\boldsymbol{e}_1, a_{12}\boldsymbol{e}_1 + a_{22}\boldsymbol{e}_2, a_{13}\boldsymbol{e}_1 + a_{23}\boldsymbol{e}_2 + a_{33}\boldsymbol{e}_3) \\ &= \det(a_{11}\boldsymbol{e}_1, a_{12}\boldsymbol{e}_1, a_{13}\boldsymbol{e}_1 + a_{23}\boldsymbol{e}_2 + a_{33}\boldsymbol{e}_3) + \det(a_{11}\boldsymbol{e}_1, a_{22}\boldsymbol{e}_2, a_{13}\boldsymbol{e}_1 + a_{23}\boldsymbol{e}_2 + a_{33}\boldsymbol{e}_3) \\ &= a_{11}a_{12}\underbrace{\det(\boldsymbol{e}_1, \boldsymbol{e}_1, a_{13}\boldsymbol{e}_1 + a_{23}\boldsymbol{e}_2 + a_{33}\boldsymbol{e}_3)}_{=0} + a_{11}a_{22}\det(\boldsymbol{e}_1, \boldsymbol{e}_2, a_{13}\boldsymbol{e}_1 + a_{23}\boldsymbol{e}_2 + a_{33}\boldsymbol{e}_3) \\ &= a_{11}a_{22}\Big(\underbrace{\det(\boldsymbol{e}_1, \boldsymbol{e}_2, a_{13}\boldsymbol{e}_1)}_{=0} + \underbrace{\det(\boldsymbol{e}_1, \boldsymbol{e}_2, a_{23}\boldsymbol{e}_2)}_{=0} + \det(\boldsymbol{e}_1, \boldsymbol{e}_2, a_{33}\boldsymbol{e}_3)\Big) \\ &= a_{11}a_{22}a_{33}\det(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3) = a_{11}a_{22}a_{33}. \end{aligned}$$

This proves the proposition when A is upper triangular. If A is lower triangular, then its transpose $A^\dagger$ is upper triangular and we deduce

$$\det A = \det A^\dagger = a^\dagger_{11}a^\dagger_{22}a^\dagger_{33} = a_{11}a_{22}a_{33}.$$

□

Recall that we have a collection of elementary column (row) operations on a matrix. The next result explains the effect of these operations on the determinant of a matrix.

Proposition 1.16. *Suppose that A is an $n \times n$-matrix. The following hold.*

(a) If the matrix B is obtained from A by multiplying the elements of the i-th column of A by the same nonzero scalar λ, then

$$\det B = \lambda \det A.$$

(b) If the matrix B is obtained from A by switching the order of the columns i and j, $i \neq j$ then

$$\det B = -\det A.$$

(c) If the matrix B is obtained from A by adding to the i-th column, the j-th column, $j \neq i$ then

$$\det B = \det A.$$

(d) Similar results hold if we perform row operations of the same type.

Proof. (a) We have

$$\begin{aligned} \det B &= \det(B\boldsymbol{e}_1, \dots, B\boldsymbol{e}_n) = \det(A\boldsymbol{e}_1, \dots, \lambda A\boldsymbol{e}_i, A\boldsymbol{e}_n) \\ &= \lambda \det(A\boldsymbol{e}_1, \dots, A\boldsymbol{e}_i, A\boldsymbol{e}_n) = \lambda \det A. \end{aligned}$$

(b) Observe that for any $\sigma \in \mathcal{S}_n$ we have

$$\det(A\boldsymbol{e}_{\sigma(1)}, \dots, A\boldsymbol{e}_{\sigma(n)}) = \operatorname{sign}(\sigma)\det(A\boldsymbol{e}_1, \dots, A\boldsymbol{e}_{\sigma(n)}) = \operatorname{sign}(\sigma)\det A.$$

Now observe that the columns of B are

$$B\boldsymbol{e}_1 = A\boldsymbol{e}_{\tau_{ij}(1)}, \dots, B\boldsymbol{e}_n = A\boldsymbol{e}_{\tau_{ij}(n)}$$

and $\operatorname{sign}(\tau_{ij}) = 1$.

For (c) we observe that

$$\begin{aligned}\det B &= \det(A\boldsymbol{e}_1, \dots, A\boldsymbol{e}_{i-1}, A\boldsymbol{e}_i + A\boldsymbol{e}_j, A\boldsymbol{e}_{i+1}, \dots, A\boldsymbol{e}_j, \dots, A\boldsymbol{e}_n)\\ &= \det(A\boldsymbol{e}_1, \dots, A\boldsymbol{e}_{i-1}, A\boldsymbol{e}_i, A\boldsymbol{e}_{i+1}, \dots, A\boldsymbol{e}_j, \dots, A\boldsymbol{e}_n)\\ &+ \underbrace{\det(A\boldsymbol{e}_1, \dots, A\boldsymbol{e}_{i-1}, A\boldsymbol{e}_j, A\boldsymbol{e}_{i+1}, \dots, A\boldsymbol{e}_j, \dots, A\boldsymbol{e}_n)}_{=0}\\ &= \det A.\end{aligned}$$

Part (d) follows by applying (a), (b), (c) to the transpose of A, observing that the rows of A are the columns of $A^\dagger$ and then using the equality $\det C = \det C^\dagger$.

□

The above results represents one efficient method for computing determinants because we know that by performing elementary row operations on a square matrix we can reduce it to upper triangular form.

Here is a first application of determinants.

Proposition 1.17. *Suppose that A is an $n \times n$-matrix with entries in $\mathbb{F}$. Then the following statements are equivalent.*

(a) *The matrix A is invertible.*
(b) $\det A \neq 0$.

Proof. A matrix A is invertible if and only if by performing elementary row operations we can reduce to an upper triangular matrix B whose diagonal entries are nonzero, i.e., $\det B \neq 0$. By performing elementary row operation the determinant changes by a nonzero factor so that

$$\det A \neq 0 \Longleftrightarrow \det B \neq 0.$$

□

Corollary 1.18. *Suppose that $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n \in \mathbb{F}^n$. The following statements are equivalent.*

(a) *The vectors $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ are linearly independent.*
(b) $\det(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n) \neq 0$.

Proof. Consider the linear operator $A : \mathbb{F}^n \to \mathbb{F}^n$ given by $A\boldsymbol{e}_i = \boldsymbol{u}_i$, $i = 1, \dots, n$. We can tautologically identify it with a matrix and we have

$$\det(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n) = \det A.$$

Now observe that $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ are linearly independent of and only if A is invertible and according to the previous propostion, this happens if and only if $\det A \neq 0$.

□

1.5. **Additional properties of determinants.**

Proposition 1.19. *If A, B are two $n \times n$-matrices , then*

$$\det AB = \det A \det B. \tag{1.10}$$

Proof. We have

$$\det AB = \det(AB\boldsymbol{e}_1, \dots, AB\boldsymbol{e}_n) = \det\Big(\sum_{i_1=1}^n b_{i_1 1} A\boldsymbol{e}_{i_1}, \dots \sum_{i_n=1}^n b_{i_n n} A\boldsymbol{e}_{i_n}\Big)$$

$$= \sum_{i_1,\dots,i_n=1}^{b} b_{i_1 1}\cdots b_{i_n n}\det(A\boldsymbol{e}_{i_1},\dots A\boldsymbol{e}_{i_n})$$

In the above sum, the only nontrivial terms correspond to choices of pairwise distinct indices $i_1,\dots,i_n$. For such a choice, the sequence $i_1,\dots,i_n$ describes a permutation of $\boldsymbol{I}_n$. We deduce

$$\det AB = \sum_{\sigma\in\mathcal{S}_n} b_{\sigma(1)1}\cdots b_{\sigma(n)n}\underbrace{\det(A\boldsymbol{e}_{\sigma(1)},\dots,A\boldsymbol{e}_{\sigma(n)})}_{=\mathrm{sign}(\sigma)\det A}$$

$$= \det A\sum_{\sigma\in\mathcal{S}_n}\mathrm{sign}(\sigma)b_{\sigma(1)1}\cdots b_{\sigma(n)n} = \det A\det B.$$

□

Corollary 1.20. *If A is an invertible matrix, then*

$$\det A^{-1} = \frac{1}{\det A}.$$

Proof. Indeed, we have

$$A\cdot A^{-1} = \mathbb{1}$$

so that

$$\det A\det A^{-1} = \det\mathbb{1} = 1.$$

□

Proposition 1.21. *Suppose that m,n are positive integers and S is an $(m+n)\times(m+n)$-matrix that has the block form*

$$S = \left[\begin{array}{cc} A & C\\ 0 & B\end{array}\right],$$

where A is an $m\times m$-matrix, B is an $n\times n$-matrix and C is an $m\times n$-matrix. Then

$$\det S = \det A\cdot\det B.$$

Proof. We denote by s_{ij} the (i,j)-entry of S, $i,j\in\boldsymbol{I}_{m+n}$. From the block description of S we deduce that

$$j\leq m \ \text{ and } \ i>n\Rightarrow s_{ij}=0. \tag{1.11}$$

We have

$$\det S = \sum_{\sigma\in\mathcal{S}_{m+n}}\mathrm{sign}(\sigma)\prod_{i=1}^{m+n}s_{\sigma(i)i},$$

From (1.11) we deduce that in the above sum the nonzero terms correspond to permutations $\sigma\in\mathcal{S}_{m+n}$ such that

$$\sigma(i)\leq m,\ \ \forall i\leq m. \tag{1.12}$$

If σ is such a permutation, ten its restriction to $\boldsymbol{I}_m$ is a permutation α of $\boldsymbol{I}_m$ and its restriction to $\boldsymbol{I}_{m+n}\setminus\boldsymbol{I}_m$ is a permutation of this set, which we regard as a permutation β of $\boldsymbol{I}_n$. Conversely, given $\alpha\in\mathcal{S}_m$ and $\beta\in\mathcal{S}_n$ we obtain a permutation $\sigma=\alpha*\beta\in\mathcal{S}_{m_n}$ satisfying (1.12) given by

$$\alpha*\beta(i) = \begin{cases}\alpha(i), & i\leq m,\\ m+\beta(i-m), & i>m.\end{cases}$$

Observe that

$$\operatorname{sign}(\alpha * \beta) = \operatorname{sign}(\alpha)\operatorname{sign}(\beta),$$

and we deduce

$$\det S = \sum_{\alpha\in\mathcal{S}_m,\ \beta\in\mathcal{S}_n} \operatorname{sign}(\alpha*\beta)\prod_{i=1}^{m+n} s_{\alpha*\beta(i)i}$$

$$= \left(\sum_{\alpha\in\mathcal{S}_m}\operatorname{sign}(\alpha)\prod_{i=1}^{m} s_{\alpha(i)i}\right)\left(\sum_{\beta\in\mathcal{S}_n}\operatorname{sign}(\beta)\prod_{j=1}^{n} s_{m+\beta(j),j+m}\right) = \det A\det B.$$

□

Definition 1.22. If A is an $n\times n$-matrix and $i,j\in\boldsymbol{I}_n$, we denote by $A(i,j)$ the matrix obtained from A by removing the i-th row and the j-th column. □

Corollary 1.23. *Suppose that the j-th column of an $n\times n$-matrix A is sparse, i.e., all the elements on the j-th column, with the possible exception of the element on the i-th row, are equal to zero. Then*

$$\det A = (-1)^{i+j}a_{ij}\det A(i,j).$$

Proof. Observe that if $i=j=1$ then A has the block form

$$A = \begin{bmatrix} a_{11} & * \\ 0 & A(1,1)\end{bmatrix}$$

and the result follows from Proposition 1.21.

We can reduce the general case to this special case by permuting rows and columns of A. If we switch the j-th column with $(j-1)$-th column we can arrange that the $(j-1)$-th column is the sparse column. Iterating this procedure we deduce after $(j-1)$ such switches that the first column is the sparse column.

By performing $(i-1)$ row-switches we can arrange that the nontrivial element on this sparse column is situated on the first row. Thus, after a total of $i+j-2$ row and column switches we obtain a new matrix A' with the block form

$$A' = \begin{bmatrix} a_{ij} & * \\ 0 & A(i,j)\end{bmatrix}$$

We have

$$(-1)^{i+j}\det A = \det A' = a_{ij}\det A(i,j).$$

□

Corollary 1.24 (Row and column expansion)**.** *Fix $j\in\boldsymbol{I}_n$. Then for any $n\times n$-matrix we have*

$$\det A = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}\det A(i,j) = \sum_{k=1}^{n}(-1)^{i+k}a_{jk}A(j,k).$$

*The first equality is referred to as the j-*th column expansion *of* $\det A$*, while the second equality is referred to as the j-*th row expansion *of* $\det A$*.*

Proof. We prove only the column expansion. The row expansion is obtained by applying to column expansion to the transpose matrix. For simplicity we assume that $j=1$. We have

$$\det A = \det(A\boldsymbol{e}_1, A\boldsymbol{e}_2,\dots,A\boldsymbol{e}_n) = \det\Big(\sum_{i-1}^{n} a_{i1}\boldsymbol{e}_i, A\boldsymbol{e}_2,\dots,A\boldsymbol{e}_n\Big)$$

$$= \sum_{i-1}^{n} a_{i1} \det\bigl(\boldsymbol{e}_i, A\boldsymbol{e}_2, \dots, A\boldsymbol{e}_n\bigr).$$

Denote by A_i the matrix whose first column is the column basic vector $\boldsymbol{e}_i$, and the other columns are the corresponding columns of A, $A\boldsymbol{e}_2, \dots, A\boldsymbol{e}_n$. We can rewrite the last equality as

$$\det A = \sum_{i=1}^{n} a_{i1} \det A_i.$$

The first column of A_i is sparse, and the submatrix $A_i(i,1)$ is equal to the submatrix $A(i,1)$. We deduce from the previous corollary that

$$\det A_i = (-1)^{i+1} \det A_i(i,1) = (-1)^{i+1} \det A(i,1).$$

This completes the proof of the column expansion formula. □

Corollary 1.25. *If* $k \neq j$ *then*

$$\sum_{i=1}^{n} (-1)^{i+j} a_{ik} \det A(i,j) = 0.$$

Proof. Denote by A' the matrix obtained from A by removing the j-th column and replacing with the k-th column of A. Thus, in the new matrix A' the j-th and the k-th columns are identical so that $\det A' = 0$. On the other hand $A'(i,j) = A(i,j)$ Expanding $\det A'$ along the j-th column we deduce

$$0 = \det A' = \sum_{i=1}^{n} (-1)^{i+j} a'_{ij} \det A(i,j) = \sum_{i=1}^{n} (-1)^{ij} a_{ik} \det A(i,j).$$

□

Definition 1.26. For any $n \times n$ matrix A we define the *adjoint matrix* $\check{A}$ to be the $n \times n$-matrix with entries

$$\check{a}_{ij} = (-1)^{i+j} \det A(j,i), \;\; \forall i,j \in \boldsymbol{I}_n.$$ □

Form Corollary 1.24 we deduce that for any j we have

$$\sum_{i=1}^{n} \check{a}_{ji} a_{ij} = \det A,$$

while Corollary 1.25 implies that for any $j \neq k$ we have

$$\sum_{i=1}^{n} \check{a}_{ji} a_{ik} = 0.$$

The last two identities can be rewritten in the compact form

$$\check{A}A = (\det A)\mathbb{1}. \tag{1.13}$$

If A is invertible, then from the above equality we conclude that

$$A^{-1} = \frac{1}{\det A}\check{A}. \tag{1.14}$$

Example 1.27. Suppose that A is a 2×2 matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Then

$$\det A = a_{11}a_{22} - a_{12}a_{21},$$

$$A(1,1) = [a_{22}], \quad A(1,2) = [a_{21}], \quad A(2,1) = [a_{12}], \quad A(2,2) = [a_{11}],$$

$$\check{a}_{11} = \det A(1,1) = a_{22}, \quad \check{a}_{12} = -\det A(2,1) = -a_{12},$$

$$\check{a}_{21} = -\det A(1,2) = -a_{21}, \quad \check{a}_{22} = \det A(2,2) = a_{11},$$

so that

$$\check{A} = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

and we observe that

$$\check{A}A = \begin{bmatrix} \det A & 0 \\ 0 & \det A \end{bmatrix}.$$

□

Proposition 1.28 (Cramer's Rule)**.** *Suppose that A is an invertible $n \times n$-matrix and $\boldsymbol{u}, \boldsymbol{x} \in \mathbb{F}^n$ are two column vectors such that*

$$A\boldsymbol{x} = \boldsymbol{u},$$

i.e., $\boldsymbol{x}$ is a solution of the linear system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & u_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & u_2 \\ \vdots & \vdots & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n & = & u_n. \end{cases}$$

Denote by $A_j(\boldsymbol{u})$ the matrix obtained from A by replacing the j-th column with the column vector $\boldsymbol{u}$. Then

$$x_j = \frac{\det A_j(\boldsymbol{u})}{\det A}, \quad \forall j = 1, \ldots, n. \tag{1.15}$$

Proof. By expanding along the j-th column of $A_j(\boldsymbol{u})$ we deduce

$$\det A_j(u) = \sum_{k=1}^{n} (-1)^{j+k} \det A(k,j). \tag{1.16}$$

On the other hand,

$$(\det A)\boldsymbol{x} = (\check{A}A)\boldsymbol{x} = \check{A}\boldsymbol{u}.$$

Hence

$$(\det A)x_j = \sum_{k=1}^{n} \check{a}_{jk}u_k = \sum_{k} (-1)^{k+j} u_k \det A(k,j) \stackrel{(1.16)}{=} \det A_j(\boldsymbol{u}).$$

□

1.6. **Examples.** To any list of complex numbers $(x_1, \dots, x_n)$ we associate the $n \times n$ matrix

$$V(x_1, \dots, x_n) = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & \vdots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{bmatrix}. \tag{1.17}$$

This matrix is called the *Vandermonde matrix* associated to the list of numbers $(x_1, \dots, x_n)$. We want to compute its determinant. Observe first that

$$\det V(x_1, \dots x_n) = 0.$$

if the numbers $z_1, \dots, z_n$ are not distinct. Observe next that

$$\det V(x_1, x_2) = \det \begin{bmatrix} 1 & 1 \\ x_1 & x_2 \end{bmatrix} = (x_2 - x_1).$$

Consider now the 3×3 situation. We have

$$\det V(x_1, x_2, x_3) = \det \begin{bmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ x_1^2 & x_2^2 & x_3^2 \end{bmatrix}.$$

Subtract from the 3rd row the second row multiplied by x_1 to deduce

$$\det V(x_1, x_2, x_3) = \det \begin{bmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ 0 & x_2^2 - x_1x_2 & x_3^2 - x_3x_1 \end{bmatrix}$$

$$= \det \begin{bmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ 0 & x_2(x_2 - x_1) & x_3^2 - x_3x_1 \end{bmatrix}.$$

Subtract from the 2nd row the first row multiplied by x_1 to deduce

$$\det V(x_1, x_2, x_3) = \det \begin{bmatrix} 1 & 1 & 1 \\ 0 & x_2 - x_1 & x_3 - x_1 \\ 0 & x_2(x_2 - x_1) & x_3(x_3 - x_1) \end{bmatrix} = \det \begin{bmatrix} x_2 - x_1 & x_3 - x_1 \\ x_2(x_2 - x_1) & x_3(x_3 - x_1) \end{bmatrix}$$

$$= (x_2 - x_1)(x_3 - x_1) \det \begin{bmatrix} 1 & 1 \\ x_2 & x_3 \end{bmatrix} = (x_2 - x_1)(x_3 - x_1) \det V(x_2, x_3).$$

$$= (x_2 - x_1)(x_3 - x_1)(x_3 - x_1).$$

We can write the above equalities in a more compact form

$$\det V(x_1, x_2) = \prod_{1 \le i < j \le 2} (x_j - x_i), \ \ \det V(x_1, x_2, x_3) = \prod_{1 \le i < j \le 3} (x_j - x_i). \tag{1.18}$$

A similar row manipulation argument (left to you as an exercise) shows that

$$\det V(x_1, \dots, x_n) = (x_2 - x_1) \cdots (x_n - x_1) \det V(x_2, \dots, x_n). \tag{1.19}$$

We have the following general result.

Proposition 1.29. *For any integer $n \ge 2$ and any complex numbers $x_1, \dots, x_n$ we have*

$$\det V_n(x_1, \dots, x_n) = \prod_{1 \le i < j \le n} (x_j - x_i). \tag{1.20}$$

Proof. We will argue by induction on n. The case $n = 2$ is contained in (1.18). Assume now that (1.20) is true for $n-1$. This means that

$$\det V(x_2, \dots, x_n) = \prod_{2 \leq i < j \leq n} (x_j - x_i).$$

Using this in (1.19) we deduce

$$\det V_n(x_1, \dots, x_n) = (x_2 - x_1) \cdots (x_n - x_1) \cdot \prod_{2 \leq i < j \leq n} (x_j - x_i) = \prod_{1 \leq i < j \leq n} (x_j - x_i). \qquad \square$$

Here is a simple application of the above computation.

Corollary 1.30. *If* $x_1, \dots, x_n$ *are* distinct *complex numbers then for any complex numbers* $r_1, \dots, r_n$ *there exists a polynomial of degree* $\leq n-1$ *uniquely determined by the conditions*

$$P(x_1) = r_1, \dots, p(x_n) = r_n. \tag{1.21}$$

Proof. The polynomial P must have the form

$$P(x) = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1},$$

where the coefficients $a_0, \dots, a_{n-1}$ are to be determined. We will do this using (1.21) which can be rewritten as a system of linear equations in which the unknown are the coefficients $a_0, \dots, a_{n-1}$,

$$\begin{cases} a_0 + a_0 x_1 + \cdots + a_{n-1} x_1^{n-1} & = & r_1 \\ a_0 + a_1 x_2 + \cdots + a_{n-1} x_2^{n-1} & = & r_2 \\ \vdots & \vdots & \vdots \\ a_0 + a_1 x_n + \cdots + a_{n-1} x_n^{n-1} & = & r_n. \end{cases}$$

We can rewrite this in matrix form

$$\underbrace{\begin{bmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{bmatrix}}_{=V(x_1, \dots, x_n)^\dagger} \cdot \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}.$$

Because the numbers $x_1, \dots, x_n$ are distinct, we deduce from (1.20) that

$$\det V(x_1, \dots, x_n)^\dagger = V(x_1, \dots, x_n) \neq 0.$$

Hence the above linear system has a unique solution $a_0, \dots, a_{n-1}$. $\square$

1.7. **Exercises.**

Exercise 1.1. Prove that the map in Example 1.2 is indeed a bilinear map. □

Exercise 1.2. Prove Proposition 1.4. □

Exercise* 1.3. (a) Show that for any $i \neq\in \boldsymbol{I}_n$ we have $\tau_{ij}\circ\tau_{ij} = \mathbb{1}_{\boldsymbol{I}_n}$.

(b) Prove that for any permutation $\sigma\in\mathcal{S}_n$ there exists a sequence of transpositions $\tau_{i_1j_1},\dots,\tau_{i_mj_m}$, $m<n$, such that

$$\tau_{i_mj_m}\circ\cdots\circ\tau_{i_1j_1}\circ\sigma = \mathbb{1}_{\boldsymbol{I}_n}.$$

Conclude that any permutation is a product of transpositions. □

Exercise 1.4. Decompose the permutation

$$\left(\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5\\ 5 & 4 & 3 & 2 & 1 \end{array}\right)$$

as a composition of transpositions. □

Exercise 1.5. Suppose that $\Phi\in\mathcal{T}^2(\boldsymbol{U})$ is a *symmetric* bilinear map. Define $Q:\boldsymbol{U}\to\mathbb{F}$ by setting

$$Q(\boldsymbol{u}) = \Phi(\boldsymbol{u},\boldsymbol{u}),\ \ \forall\boldsymbol{u}\in\boldsymbol{U}.$$

Show that for any $\boldsymbol{u},\boldsymbol{v}\in\boldsymbol{U}$ we have

$$\Phi(\boldsymbol{u},\boldsymbol{v}) = \frac{1}{4}\Big(Q(\boldsymbol{u}+\boldsymbol{v}) - Q(\boldsymbol{u}-\boldsymbol{v})\Big).$$ □

Exercise 1.6. Prove that the map

$$\Phi:\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R},\ \ \Phi(\boldsymbol{u},\boldsymbol{v}) = u_1v_2 - u_2v_1,$$

is bilinear, and skew-symmetric. □

Exercise 1.7. (a) Show that a bilinear form $\Phi:\boldsymbol{U}\times\boldsymbol{U}\to\mathbb{F}$ is skew-symmetric if and only if $\Phi(\boldsymbol{u},\boldsymbol{u})=0$, $\forall\boldsymbol{u}\in\boldsymbol{U}$.

Hint: Expand $\Phi(\boldsymbol{u}+\boldsymbol{v},\boldsymbol{u}+\boldsymbol{v})$ using the bilinearity of Φ.

(b) Prove that an n-linear form $\Phi\in\mathcal{T}^n(\boldsymbol{U})$ is skew-symmetric if and only if for any $i\neq j$ and any vectors $\boldsymbol{u}_1,\dots,\boldsymbol{u}_n\in\boldsymbol{U}$ such that $\boldsymbol{u}_i=\boldsymbol{u}_j$ we have

$$\Phi(\boldsymbol{u}_1,\dots,\boldsymbol{u}_n) = 0.$$

Hint. Use the trick in part (a) and Exercise 1.3. □

Exercise 1.8. Compute the determinant of the following 5×5-matrix

$$\left[\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5\\ 2 & 1 & 2 & 3 & 4\\ 3 & 2 & 1 & 2 & 3\\ 4 & 3 & 2 & 1 & 2\\ 5 & 4 & 3 & 2 & 1 \end{array}\right].$$

Exercise 1.9. Fix complex numbers x and h. Compute the determinant of the matrix

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ x & h & -1 & 0 \\ x^2 & hx & h & -1 \\ x^3 & hx^2 & hx & h \end{bmatrix}.$$

Can you generalize thus example? □

Exercise 1.10. Prove the equality (1.19). □

Exercise 1.11. (a) Consider a degree $(n-1)$ polynomial

$$P(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1 x + a_0,\ \ a_{n-1} \neq 0.$$

Compute the determinant of the following matrix.

$$V = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & \vdots & \vdots \\ x_1^{n-2} & x_2^{n-2} & \cdots & x_n^{n-2} \\ P(x_1) & P(x_2) & \cdots & P(x_n) \end{bmatrix}.$$

(b) Compute the determinants of the following $n \times n$ matrices

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & \vdots & \vdots \\ x_1^{n-2} & x_2^{n-2} & \cdots & x_n^{n-2} \\ x_2x_3\cdots x_n & x_1x_3x_4\cdots x_n & \cdots & x_1x_2\cdots x_{n-1} \end{bmatrix},$$

and

$$B = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & \vdots & \vdots \\ x_1^{n-2} & x_2^{n-2} & \cdots & x_n^{n-2} \\ (x_2+x_3+\cdots+x_n)^{n-1} & (x_1+x_3+x_4+\cdots+x_n)^{n-1} & \cdots & (x_1+x_2+\cdots+x_{n-1})^{n-1} \end{bmatrix}.$$

Hint. To compute $\det B$ it is wise to write $S = x_1 + \cdots + x_n$ so that $x_2 + x_3 + \cdots + x_n = (S - x_1)$, $x_1 + x_3 + \cdots + x_n = S - x_2$ etc. Next observe that $(S-x)^k$ is a polynomial of degree k in x. □

Exercise 1.12. Suppose that A is skew-symmetric $n \times n$ matrix, i.e.,

$$A^\dagger = -A.$$

Show that $\det A = 0$ if n is odd. □

Exercise 1.13. Suppose that $A = (a_{ij})_{1\leq i,j\leq n}$ is an $n \times n$ matrix with complex entries.
(a) Fix complex numbers $x_1, \dots, x_n, y_1, \dots, y_n$ and consider the $n \times n$ matrix B with entries

$$b_{ij} = x_i y_j a_{ij}.$$

Show that

$$\det B = (x_1 y_1 \cdots x_n y_n) \det A.$$

(b) Suppose that C is the $n \times n$ matrix with entries

$$c_{ij} = (-1)^{i+j} a_{ij}.$$

Show that $\det C = \det A$. □

Exercise 1.14. (a) Suppose we are given three sequences of numbers $\underline{a} = (a_k)_{k\geq 1}$, $\underline{b} = (b_k)_{k\geq 1}$ and $\underline{c} = (c_k)_{k\geq 1}$. To these seqences we associate a sequence of *Jacobi matrices*

$$J_n = \begin{bmatrix} a_1 & b_1 & 0 & 0 & \cdots & 0 & 0 \\ c_1 & a_2 & b_2 & 0 & \cdots & 0 & 0 \\ 0 & c_2 & a_3 & b_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & c_{n-1} & a_n \end{bmatrix}. \tag{$\boldsymbol{J}$}$$

Show that

$$\det J_n = a_n \det J_{n-1} - b_{n-1}c_{n-1} \det J_{n-2}. \tag{1.22}$$

Hint: Expand along the last row.
(b) Suppose that above we have

$$c_k = 1, \;\; b_k = 2, \;\; a_k = 3, \;\; \forall k \geq 1.$$

Compute J_1, J_2. Using (1.22) determine J_3, J_4, J_5, J_6, J_7. Can you detect a pattern? □

Exercise 1.15. Suppose we are given a sequence of polynomials with complex coefficients $(P_n(x))_{n\geq 0}$, $\deg P_n = n$, for all $n \geq 0$,

$$P_n(x) = a_n x^n + \cdots, \;\; a_n \neq 0.$$

Denote by $\boldsymbol{V}_n$ the space of polynomials with complex coefficients and degree $\leq n$.
(a) Show that the collection $\{P_0(x), \dots, P_n(x)\}$ is a basis of $\boldsymbol{V}_n$.
(b) Show that for any $x_1, \dots, x_n \in \mathbb{C}$ we have

$$\det \begin{bmatrix} P_0(x_1) & P_0(x_1) & \cdots & P_0(x_n) \\ P_1(x_1) & P_1(x_2) & \cdots & P_1(x_n) \\ \vdots & \vdots & \vdots & \vdots \\ P_{n-1}(x_1) & P_{n-1}(x_2) & \cdots & P_{n-1}(x_n) \end{bmatrix} = a_0 a_1 \cdots a_{n-1} \prod_{i<j} (x_j - x_i).$$ □

Exercise 1.16. To any polynomial $P(x) = c_0 + c_1 x + \ldots + c_{n-1}x^{n-1}$ of degree $\leq n-1$ with complex coefficients we associate the $n \times n$ *circulant matrix*

$$C_P = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_{n-2} & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \cdots & c_{n-3} & c_{n-2} \\ c_{n-2} & c_{n-1} & c_0 & \cdots & c_{n-4} & c_{n-3} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_1 & c_2 & c_3 & \cdots & c_{n-1} & c_0 \end{bmatrix},$$

Set

$$\rho = e^{\frac{2\pi \boldsymbol{i}}{n}}, \;\; \boldsymbol{i} = \sqrt{-1},$$

so that $\rho^n = 1$. Consider the $n \times n$ Vandermonde matrix $V_\rho = V(1, \rho, \dots, \rho^{n-1})$ defined as in (1.17)
(a) Show that for any $j = 1, \dots, n-1$ we have

$$1 + \rho^j + \rho^{2j} + \cdots + \rho^{(n-1)j} = 0.$$

(b) Show that

$$C_P \cdot V_\rho = V_\rho \cdot \mathrm{Diag}\big(P(1), P(\rho), \dots, P(\rho^{n-1}) \big),$$

where $\mathrm{Diag}(a_1, \dots, a_n)$ denotes the diagonal $n \times n$-matrix with diagonal entries $a_1, \dots, a_n$.
(c) Show that

$$\det C_P = P(1)P(\rho) \cdots P(\rho^{n-1}).$$

□

(d)* Suppose that $P(x) = 1 + 2x + 3x^2 + 4x^3$ so that C_P is a 4×4-matrix with integer entries and thus $\det C_P$ is an integer. Find this *integer*. Can you generalize this computation?

Exercise 1.17. Consider the $n \times n$-matrix

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

(a) Find the matrices

$$A^2, A^3, \dots, A^n.$$

(b) Compute $(I - A)(I + A + \cdots + A^{n-1})$.
(c) Find the inverse of $(I - A)$. □

Exercise 1.18. Let

$$P(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0$$

be a polynomial of degree d with complex coefficients. We denote by $\mathcal{S}$ the collection of sequences of complex numbers, i.e., functions

$$f : \{0, 1, 2, \dots\} \to \mathbb{C}, \;\; n \mapsto f(n).$$

This is a complex vector space in a standard fashion. We denote by $\mathcal{S}_P$ the subcollection of sequences $f \in \mathcal{S}$ satisfying the *recurrence relation*

$$f(n+d) + a_{d-1}f(n+d-1) + \cdots a_1 f(n+1) + a_0 f(n) = 0, \;\; \forall n \geq 0. \qquad (\boldsymbol{R}_P)$$

(a) Show that $\mathcal{S}_P$ is a vector subspace of $\mathcal{S}$.
(b) Show that the map $\mathcal{I} : \mathcal{S}_P \to \mathbb{C}^d$ which associates to $f \in \mathcal{S}_P$ its initial values $\mathcal{I}f$,

$$\mathcal{I}f = \begin{bmatrix} f(0) \\ f(1) \\ \vdots \\ f(d-1) \end{bmatrix} \in \mathbb{C}^d$$

is an isomorphism of vector spaces.
(c) For any $\lambda \in \mathbb{C}$ we consider the sequence f_λ defined by

$$f_\lambda(n) = \lambda^n, \;\; \forall n \geq 0.$$

(Above it is understood that $\lambda^0 = 1$.) Show that $f_\lambda \in \mathcal{S}_P$ if and only if $P(\lambda) = 0$, i.e., λ is a root of P.
(d) Suppose P has d *distinct* roots $\lambda_1, \dots, \lambda_d \in \mathbb{C}$. Show that the collection of sequences $f_{\lambda_1}, \dots, f_{\lambda_d}$ is a basis of $\mathcal{S}_P$.
(e) Consider the *Fibonacci sequence* $\big(f(n) \big)_{n \geq 0}$ defined by

$$f(0) = f(1) = 1, \;\; f(n+2) = f(n+1) + f(n), \;\; \forall n \geq 0.$$

Thus,

$$f(2)=2,\quad f(3)=3,\quad f(4)=5,\quad f(5)=8,\quad f(6)=13,\dots .$$

Use the results (a)–(d) above to find a short formula describing $f(n)$. □

Exercise 1.19. Let b,c be two distinct complex numbers. Consider the $n\times n$ Jacobi matrix

$$J_n=\begin{bmatrix} b+c & b & 0 & 0 & \cdots & 0 & 0\\ c & b+c & b & 0 & \cdots & 0 & 0\\ 0 & c & b+c & b & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & 0 & c & b+c & b\\ 0 & 0 & 0 & 0 & \cdots & c & b+c \end{bmatrix}.$$

Find a short formula for $\det J_n$.

Hint: Use the results in Exercises 1.14 and 1.18. □

2. SPECTRAL DECOMPOSITION OF LINEAR OPERATORS

2.1. **Invariants of linear operators.** Suppose that $\boldsymbol{U}$ is an n-dimensional $\mathbb{F}$-vector space. We denote by $L(\boldsymbol{U})$ the space of linear operators (maps) $T : \boldsymbol{U} \to \boldsymbol{U}$. We already know that once we choose a basis $\underline{be} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ of $\boldsymbol{U}$ we can represent T by a matrix

$$A = \mathcal{M}(\underline{\boldsymbol{e}}, T) = (a_{ij})_{1 \leq i,j \leq n},$$

where the elements of the k-th column of A describe the coordinates of $T\boldsymbol{e}_k$ in the basis $\underline{\boldsymbol{e}}$, i.e.,

$$T\boldsymbol{e}_k = a_{1k}\boldsymbol{e}_1 + \cdots + a_{nk}\boldsymbol{e}_n = \sum_{j=1}^{n} a_{jk}\boldsymbol{e}_j.$$

A priori, there is no good reason of choosing the basis $\underline{\boldsymbol{e}} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ over another $\underline{\boldsymbol{f}} = (\boldsymbol{f}_1, \dots, \boldsymbol{f}_n)$. With respect to this new basis the operator T is represented by another matrix

$$B = \mathcal{M}(\underline{\boldsymbol{f}}, T) = (b_{ij})_{1 \leq i,j \leq n}, \;\; T\boldsymbol{f}_k = \sum_{j=1}^{n} b_{jk}\boldsymbol{f}_j.$$

The basis $\underline{\boldsymbol{f}}$ is related to the basis $\underline{\boldsymbol{e}}$ by a *transition matrix*

$$C = (c_{ij})_{1 \leq i,j \leq n}, \;\; \boldsymbol{f}_k = \sum_{j=1}^{n} c_{jk}\boldsymbol{e}_j.$$

Thus the, k-th column of C describes the coordinates of the vector $\boldsymbol{f}_k$ in the basis $\underline{\boldsymbol{e}}$. Then C is invertible and

$$B = C^{-1}AC. \tag{2.1}$$

The space $\boldsymbol{U}$ has lots of bases, so *the same* operator T can be represented by many different matrices. The question we want to address in this section can be loosely stated as follows.

Find bases of $\boldsymbol{U}$ so that, in these bases, the operator T represented by "very simple" matrices.

We will not define what a a "very simple" matrix is but we will agree that the more zeros a matrix has, the simpler it is. We already know that we can find bases in which the operator T is represented by upper triangular matrices. These have lots of zero entries, but it turns out that we can do much better than this.

The above question is closely related to the concept of *invariant* of a linear operator. An invariant is roughly speaking a quantity naturally associated to the operator that does not change when we change bases.

Definition 2.1. (a) A subspace $\boldsymbol{V} \subset \boldsymbol{U}$ is called an *invariant subspace* of the linear operator $T \in L(\boldsymbol{U})$ if

$$T\boldsymbol{v} \in \boldsymbol{V}, \;\; \forall \boldsymbol{v} \in \boldsymbol{V}.$$

(b) A *nonzero* vector $\boldsymbol{u}_0 \in \boldsymbol{U}$ is called an *eigenvector* of the linear operator T if and only if the linear subspace spanned by $\boldsymbol{u}_0$ is an invariant subspace of T. □

Example 2.2. (a) Suppose that $T : \boldsymbol{U} \to \boldsymbol{U}$ is a linear operator. Its *null space* or *kernel*

$$\ker T := \{\, \boldsymbol{u} \in \boldsymbol{U}; \;\; T\boldsymbol{u} = 0 \,\},$$

is an invariant subspace of T. Its dimension, $\dim \ker T$, is an *invariant* of T because in its definition we have not mentioned any particular basis. We have already encountered this dimension under a different guise.

If we choose a basis $\underline{\boldsymbol{e}} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ of $\boldsymbol{U}$ and use it to represent T as an $n \times n$ matrix $A = (A_{ij})_{1 \leq i,j \leq n}$, then $\dim \ker T$ is equal to the nullity of A, i.e., the dimension of the vector space of solutions of the linear system

$$A\boldsymbol{x} = 0, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{F}^n.$$

The *range*

$$\boldsymbol{R}(T) = \bigl\{ T\boldsymbol{u}; \;\; \boldsymbol{u} \in \boldsymbol{U} \bigr\}$$

is also an invariant subspace of T. Its dimension $\dim \boldsymbol{R}(T)$ can be identified with the rank of the matrix A above. The rank nullity theorem implies that

$$\dim \ker T + \dim \boldsymbol{R}(T) = \dim \boldsymbol{U}. \tag{2.2}$$

(b) Suppose that $\boldsymbol{u}_0 \in \boldsymbol{U}$ is an eigenvector of T. Then $T\boldsymbol{u}_0 \in \operatorname{span}(u_0)$ so that there exists $\lambda \in \mathbb{F}$ such that

$$T\boldsymbol{u}_0 = \lambda \boldsymbol{u}_0.$$ □

2.2. **The determinant and the characteristic polynomial of an operator.** Assume again that $\boldsymbol{U}$ is an n-dimensional $\mathbb{F}$-vector space. A more subtle invariant of an operator $T \in L(\boldsymbol{U})$ is its determinant. This is a scalar $\det T \in \mathbb{F}$. Its definition requires a choice of a basis of $\boldsymbol{U}$, but the end result *is independent any choice of basis*. Here are the details.

Fix a basis

$$\underline{\boldsymbol{e}} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$$

of $\boldsymbol{U}$. We use it to represent T as an $n \times n$ real matrix $A = (a_{ij})_{1 \leq i,j \leq n}$. More precisely, this means that

$$T\boldsymbol{e}_j = \sum_{i=1}^n a_{ij}\boldsymbol{e}_i, \quad \forall j = 1, \dots, n.$$

If we choose another basis of $\boldsymbol{U}$,

$$\underline{\boldsymbol{f}} = (\boldsymbol{f}_1, \dots, \boldsymbol{f}_n),$$

then we can represent T by another $n \times n$ matrix $B = (b_{ij})_{1 \leq i,j \leq n}$, i.e.,

$$T\boldsymbol{f}_j = \sum_{i=1}^n b_{ij}\boldsymbol{f}_i, \quad j = 1, \dots, n.$$

As we have discussed above the basis $\underline{\boldsymbol{f}}$ is obtained from $\underline{\boldsymbol{e}}$ via a *change-of-basis* matrix $C = (c_{ij})_{1 \leq i,j \leq n}$, i.e.,

$$\boldsymbol{f}_j = \sum_{i=1}^n c_{ij}\boldsymbol{e}_j, \quad j = 1, \dots, n.$$

Moreover the matrices A, B, C are related by the *transition rule* (2.1),

$$B = C^{-1}AC.$$

Thus

$$\det B = \det(C^{-1}AC) = \det C^{-1} \det A \det C = \det A.$$

The upshot is that the matrices A and B *have the same determinant*. Thus, no mater what basis of $\boldsymbol{U}$ we choose to represent T as an $n \times n$ matrix, the determinate of that matrix *is independent of the basis used*. This number, denoted by $\det T$ is an invariant of T called the *determinant* of the *operator* T. Here is a simple application of this concept.

Corollary 2.3.

$$\ker T \neq 0 \Longleftrightarrow \det T = 0. \qquad \square$$

More generally, for any $x \in \mathbb{F}$ consider the operator

$$x\mathbb{1} - T : \boldsymbol{U} \to \boldsymbol{U},$$

defined by

$$(\lambda\mathbb{1} - T)\boldsymbol{u} = x\boldsymbol{u} - T\boldsymbol{u}, \;\; \forall \boldsymbol{u} \in \boldsymbol{U}.$$

We set

$$P_T(x) = \det(x\mathbb{1} - T).$$

Proposition 2.4. *The quantity $P_T(x)$ is a polynomial of degree $n = \dim \boldsymbol{U}$ in the variable λ.*

Proof. Choose a basis $\underline{\boldsymbol{e}} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$. In this basis T is represented by an $n \times m$ matrix $A = (a_{ij})_{1 \leq i,j \leq n}$ and the operator $x\mathbb{1} - T$ is represented by the matrix

$$xI - A = \begin{bmatrix} x - a_{11} & -a_{12} & -a_{13} & \cdots & -a_{1n} \\ -a_{21} & x - a_{22} & -a_{23} & \cdots & -a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -a_{n1} & -a_{n2} & -a_{n3} & \cdots & x - a_{nn} \end{bmatrix}.$$

As explained in Remark 1.13, the determinant of this matrix is a sum of products of certain choices of n entries of this matrix, namely the entries that form a rook placement. Since there are exactly n entries in this matrix that contain the variable x, we see that each product associated to a rook placement of entries is a polynomial in x of degree $\leq n$. There exists *exactly one* rook placement so that each of the entries of this placement contain the term x This pavement is easily described, it consists of the terms situated on the diagonal of this matrix, and the product associated to these entries is

$$(x - a_{11}) \cdots (x - a_{nn}).$$

Any other rook placement contains at most $(n-1)$ entries that involve the term x, so the corresponding product of these entries is a polynomial of degree at most $n - 1$. Hence

$$\det(xI - A) = (x - a_{11}) \cdots (x - a_{nn}) + \text{polynomial of degree} \leq n - 1.$$

Hence $P_T(x) = \det(xI - A)$ is a polynomial of degree n in x. $\square$

Definition 2.5. The polynomial $P_T(x)$ is called the *characteristic polynomial* of the operator T. $\square$

Recall that a number $\lambda \in \mathbb{F}$ is called an *eigenvalue* of the operator T if and only if there exists $\boldsymbol{u} \in \boldsymbol{U} \setminus 0$ such that $T\boldsymbol{u} = \lambda\boldsymbol{u}$, i.e.,

$$(\lambda\mathbb{1} - T)\boldsymbol{u} = 0.$$

Thus λ is an eigenvalue of T if and only if $\ker(\lambda I - T) \neq 0$. Invoking Corollary 2.3 we obtain the following important result.

Corollary 2.6. *A scalar $\lambda \in \mathbb{F}$ is an eigenvalue of T if and only if it is a root of the characteristic polynomial of T, i.e., $P_T(\lambda) = 0$.* $\square$

The collection of eigenvalues of an operator T is called the *spectrum* of T and it is denoted by $\operatorname{spec}(T)$. If $\lambda \in \operatorname{spec}(T)$, then the subspace $\ker(\lambda \mathbb{1} - T) \subset \boldsymbol{U}$ is called the *eigenspace* corresponding to the eigenvalue λ.

From the above corollary and the fundamental theorem of algebra we obtain the following important consequence.

Corollary 2.7. *If $T : \boldsymbol{U} \to \boldsymbol{U}$ is a linear operator on a* complex *vector space $\boldsymbol{U}$, then $\operatorname{spec}(T) \neq \emptyset$.* $\square$

We say that a linear operator $T : \boldsymbol{U} \to \boldsymbol{U}$ is *triangulable* if there exists a basis $\underline{\boldsymbol{e}} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ of $\boldsymbol{U}$ such that the matrix representing T in this basis is upper triangular. We will refer to A as a *triangular representation* of T. Triangular representations, if they exist, *are not unique*. We already know that any linear operator on a *complex* vector space is triangulable.

Corollary 2.8. *Suppose that $T : \boldsymbol{U} \to \boldsymbol{U}$ is a triangulable operator. Then for any basis $\underline{\boldsymbol{e}} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ of $\boldsymbol{U}$ such that the matrix $A = (a_{ij})_{1 \leq i,j \leq n}$ representing T in this basis is upper triangular, we have*

$$P_T(x) = (x - a_{11}) \cdots (x - a_{nn}).$$

Thus, the eigenvalues of T are the elements along the diagonal of any triangular representation of T. $\square$

2.3. **Generalized eigenspaces.** Suppose that $T : \boldsymbol{U} \to \boldsymbol{U}$ is a linear operator on the n-dimensional $\mathbb{F}$-vector space. Suppose that $\operatorname{spec}(T) \neq \emptyset$. Choose an eigenvalue $\lambda \in \operatorname{spec}(T)$.

Lemma 2.9. *Let k be a positive integer. Then*

$$\ker(\lambda \mathbb{1} - T)^k \subset \ker(\lambda \mathbb{1} - T)^{k+1}.$$

Moreover, if $\ker(\lambda \mathbb{1} - T)^k = \ker(\lambda \mathbb{1} - T)^{k+1}$, then

$$\ker(\lambda \mathbb{1} - T)^k = \ker(\lambda \mathbb{1} - T)^{k+1} = \ker(\lambda \mathbb{1} - T)^{k+2} = \ker(\lambda \mathbb{1} - T)^{k+3} = \cdots .$$

Proof. Observe that if $(\lambda \mathbb{1} - T)^k \boldsymbol{u} = 0$, then

$$(\lambda \mathbb{1} - T)^{k+1} \boldsymbol{u} = (\lambda \mathbb{1} - T)(\lambda \mathbb{1} - T)^k \boldsymbol{u} = 0,$$

so that $\ker(\lambda \mathbb{1} - T)^k \subset \ker(\lambda \mathbb{1} - T)^{k+1}$.

Suppose that

$$\ker(\lambda \mathbb{1} - T)^k = \ker(\lambda \mathbb{1} - T)^{k+1}$$

To prove that $\ker(\lambda \mathbb{1} - T)^{k+1} = \ker(\lambda \mathbb{1} - T)^{k+2}$ it suffices to show that

$$\ker(\lambda \mathbb{1} - T)^{k+1} \supset \ker(\lambda \mathbb{1} - T)^{k+2}.$$

Let $\boldsymbol{v} \in \ker(\lambda \mathbb{1} - T)^{k+2}$. Then

$$(\lambda \mathbb{1} - T)^{k+1}(\mathbb{1} - \lambda T)\boldsymbol{v} = 0,$$

so that $(\mathbb{1} - \lambda T)\boldsymbol{v} \in \ker \ker(\lambda \mathbb{1} - T)^{k+1} = \ker(\lambda \mathbb{1} - T)^k$ so that

$$(\lambda \mathbb{1} - T)^k (\mathbb{1} - \lambda T)\boldsymbol{v} = 0,$$

i.e., $\boldsymbol{v} \in \ker(\lambda \mathbb{1} - T)^{k+1}$. We have thus shown that

$$\ker(\lambda \mathbb{1} - T)^{k+1} = \ker(\lambda \mathbb{1} - T)^{k+2}.$$

The remaining equalities $\ker(\lambda \mathbb{1} - T)^{k+2} = \ker(\lambda \mathbb{1} - T)^{k+3} = \cdots$ are proven in a similar fashion. $\square$

Corollary 2.10. *For any $m \geq n = \dim \boldsymbol{U}$ we have*

$$\ker(\lambda\mathbb{1} - T)^m = \ker(\lambda\mathbb{1} - T)^n, \tag{2.3a}$$

$$\boldsymbol{R}(\lambda\mathbb{1} - T)^m = \boldsymbol{R}(\lambda\mathbb{1} - T)^n. \tag{2.3b}$$

Proof. Consider the sequence of positive integers

$$d_1(\lambda) = \dim_{\mathbb{F}}(\lambda\mathbb{1} - T), \dots, d_k(\lambda) = \dim_{\mathbb{F}}(\lambda\mathbb{1} - T)^k, \dots .$$

Lemma 2.9 shows that

$$d_1(\lambda) \leq d_2(\lambda) \leq \cdots \leq n = \dim \boldsymbol{U}.$$

Thus there must exist k such that $d_k(\lambda) = d_{k+1}(\lambda)$. We set

$$k_0 = \min\{k;\ \ d_k(\lambda) = d_{k+1}(\lambda)\}.$$

Thus

$$d_(\lambda) < \cdots < d_{k_0}(\lambda) \leq n,$$

so that $k_0 \leq n$. On the other hand, since $d_{k_0}(\lambda) = d_{k_0+1}(\lambda)$ we deduce that

$$\ker(\lambda\mathbb{1} - T)^{k_0} = \ker(\lambda\mathbb{1} - T)^m,\ \ \forall m \geq k_0.$$

Since $n \geq k_0$ we deduce

$$\ker(\lambda\mathbb{1} - T)^n = \ker(\lambda\mathbb{1} - T)^{k_0} = \ker(\lambda\mathbb{1} - T)^m,\ \ \forall m \geq k_0.$$

This proves (2.3a). To prove (2.3b) observe that if $m > n$, then

$$\boldsymbol{R}(\lambda\mathbb{1} - T)^m = (\lambda\mathbb{1} - T)^n\Big(\,(\lambda\mathbb{1} - T)^{m-n}V\,\Big) \subset (\lambda\mathbb{1} - R)^n\big(\,V\,\big) = \boldsymbol{R}(\lambda\mathbb{1} - T)^n.$$

On the other hand, the rank-nullity formula (2.2) implies that

$$\dim \boldsymbol{R}(\lambda\mathbb{1} - T)^n = \dim \boldsymbol{U} - \dim\ker(\lambda\mathbb{1} - T)^n$$

$$= \dim \boldsymbol{U} - (\lambda\mathbb{1} - T)^m = \dim \boldsymbol{R}(\lambda\mathbb{1} - T)^m.$$

This proves (2.3b). □

Definition 2.11. Let $T : \boldsymbol{U} \to \boldsymbol{U}$ be a linear operator on the n-dimensional $\mathbb{F}$-vector space $\boldsymbol{U}$. Then for any $\lambda \in \operatorname{spec}(T)$ the subspace $\ker(\lambda\mathbb{1} - T)^n$ is called the *generalized eigenspace* of T corresponding to the eigenvalue λ and it is denoted by $E_\lambda(T)$. We will denote its dimension by $m_\lambda(T)$, or m_λ, and we will refer to it as the *multiplicity* of the eigenvalue λ. □

Proposition 2.12. *Let $T \in L(\boldsymbol{U})$, $\dim_{\mathbb{F}} \boldsymbol{U} = n$, and $\lambda \in \operatorname{spec}(T)$. Then the generalized eigenspace $E_\lambda(T)$ is an invariant subspace of T.*

Proof. We need to show that $TE_\lambda(T) \subset E_\lambda(T)$. Let $\boldsymbol{u} \in E_\lambda(T)$, i.e.,

$$(\lambda\mathbb{1} - T)^n\boldsymbol{u} = 0.$$

Clearly $\lambda\boldsymbol{u} - T\boldsymbol{u} \in \ker(\lambda\mathbb{1} - T)^{n+1} = E_\lambda(T)$. Since $\lambda\boldsymbol{u} \in E_\lambda(T)$ we deduce that

$$T\boldsymbol{u} = \lambda\boldsymbol{u} - (\lambda\boldsymbol{u} - T) \in E_\lambda(T).$$

□

Theorem 2.13. *Suppose that $T : \boldsymbol{U} \to \boldsymbol{U}$ is a triangulable operator on the the n-dimensional $\boldsymbol{F}$-vector space $\boldsymbol{U}$. Then the following hold.*
(a) For any $\lambda \in \operatorname{spec}(T)$ the multiplicity m_λ is equal to the number of times λ appears along the diagonal of a triangular representation of T.
(b)

$$\det T = \prod_{\lambda \in \operatorname{spec}(T)} \lambda^{m_\lambda(T)}, \tag{2.4a}$$

$$P_T(x) = \prod_{\lambda \in \operatorname{spec}(T)} (x-\lambda)^{m_\lambda(T)}, \tag{2.4b}$$

$$\sum_{\lambda \in \operatorname{spec}(T)} m_\lambda(T) = \deg P_T = \dim \boldsymbol{U} = n. \tag{2.4c}$$

Proof. To prove (a) we will argue by induction on n. For $n = 1$ the result is trivially true. For the inductive step we assume that the result is true for any triangulable operator on an $(n-1)$-dimensional $\mathbb{F}$-vector space V, and we will prove that the same is true for triangulable operators acting on an n-dimensional space $\boldsymbol{U}$.

Let $T \in L(\boldsymbol{U})$ be such an operator. We can then find a basis $\underline{\boldsymbol{e}} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ of $\boldsymbol{U}$ such that, in this basis, the operator T is represented by the upper triangular matrix

$$A = \begin{bmatrix} \lambda_1 & * & * & * & * \\ 0 & \lambda_2 & * & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \lambda_{n-1} & * \\ 0 & \cdots & 0 & 0 & \lambda_n \end{bmatrix}$$

Suppose that $\lambda \in \operatorname{spec}(T)$. For simplicity we assume $\lambda = 0$. Otherwise, we carry the discussion of the operator $T - \lambda\mathbb{1}$. Let ν be the number of times 0 appears on the diagonal of A we have to show that

$$\nu = \dim \ker T^n.$$

Denote by $\boldsymbol{V}$ the subspace spanned by the vectors $\boldsymbol{e}_1, \dots, \boldsymbol{e}_{n-1}$. Observe that $\boldsymbol{V}$ is an invariant subspace of T, i.e., $T\boldsymbol{V} \subset \boldsymbol{V}$. If we denote by S the restriction of T to $\boldsymbol{V}$ we can regard S as a linear operator $S : \boldsymbol{V} \to \boldsymbol{V}$.

The operator S is triangulable because in the basis $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_{n-1})$ of $\boldsymbol{V}$ it is represented by the upper triangular matrix

$$B = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \lambda_{n-1} \end{bmatrix}.$$

Denote by μ the number of times 0 appears on the diagonal of B. The induction hypothesis implies that

$$\mu = \dim \ker S^{n-1} = \dim \ker S^n.$$

Clearly $\mu \leq \nu$. Note that

$$\ker S^n \subset \ker T^n$$

so that

$$\mu = \dim \ker S^n \leq \dim \ker T^n.$$

We distinguish two cases.

1. $\lambda_n \neq 0$. In this case we have $\mu = \nu$ so it suffices to show that

$$\ker T^n \subset \boldsymbol{V}.$$

Indeed, if that were the case, we would conclude that $\ker T^n \subset \ker S^n$, and thus

$$\dim \ker T^n = \dim \ker S^n = \mu = \nu.$$

We argue by contradiction. Suppose that there exists $\boldsymbol{u} \in \ker T^n$ such that $\boldsymbol{u} \notin \boldsymbol{V}$. Thus, we can find $\boldsymbol{v} \in \boldsymbol{V}$ and $c \in \mathbb{F} \setminus 0$ such that

$$\boldsymbol{u} = \boldsymbol{v} + c\boldsymbol{e}_n.$$

Note that $T^n \boldsymbol{v} \in \boldsymbol{V}$ and

$$\boldsymbol{e}_n = \lambda_n \boldsymbol{e}_n + \text{vector in } \boldsymbol{V}.$$

Thus

$$T^n c\boldsymbol{e}_n = c\lambda_n^n \boldsymbol{e}_n + \text{vector in } \boldsymbol{V}$$

so that

$$T^n \boldsymbol{u} = c\lambda_n^n \boldsymbol{e}_n + \text{vector in } \boldsymbol{V} \neq 0.$$

This contradiction completes the discussion of Case **1.**

2. $\lambda_n = 0$. In this case we have $\nu = \mu + 1$ so we have to show that

$$\dim \ker T^n = \mu + 1.$$

We need an auxiliary result.

Lemma 2.14. *There exists $\boldsymbol{u} \in \boldsymbol{U} \setminus \boldsymbol{V}$ such that $T^n \boldsymbol{u} = 0$ so that*

$$\dim(\boldsymbol{V} + \ker T^n) \geq \dim \boldsymbol{V} + 1 = n. \tag{2.5}$$

Proof. Set

$$\boldsymbol{v}_n := T\boldsymbol{e}_n.$$

Observe that $\boldsymbol{v}_n \in \boldsymbol{V}$. From (2.3b) we deduce that $\boldsymbol{R}\, S^{n-1} = \boldsymbol{R}\, S^n$ so that there exists $\boldsymbol{v}_0 \in \boldsymbol{V}$ such that

$$S^{n-1}\boldsymbol{v}_n = S^n \boldsymbol{v}_0.$$

Set $\boldsymbol{u} := \boldsymbol{e}_n - \boldsymbol{v}_0$. Note that $\boldsymbol{u} \in \boldsymbol{U} \setminus \boldsymbol{V}$,

$$T\boldsymbol{u} = \boldsymbol{v}_n - T\boldsymbol{v}_0 = \boldsymbol{v}_n - S\boldsymbol{v},$$

$$T^n \boldsymbol{u} = T^{n-1}(\boldsymbol{v}_n - S\boldsymbol{v}_0) = S^{n-1}\boldsymbol{v}_n - S^n \boldsymbol{v}_0 = 0.$$

□

Now observe that

$$n = \dim \boldsymbol{U} \geq \dim(\boldsymbol{V} + \ker T^n) \overset{(2.5)}{\geq} n,$$

so that

$$\dim(\boldsymbol{V} + \ker T^n) = n.$$

We conclude that

$$n = \dim(\boldsymbol{V} + \ker T^n) = \dim(\ker T^n) + \underbrace{\dim \boldsymbol{V}}_{n-1} - \underbrace{\dim(\boldsymbol{U} \cap \ker T^n)}_{=\mu}$$

$$= \dim(\ker T^n) + n - 1 - \mu,$$

which shows that

$$\dim \ker T^n = \mu + 1 = \nu.$$

This proves (a). The equalities (2.4c), (2.4b), (2.4c) follow easily from (a). □

☞ *In the remainder of this section we will assume that $\mathbb{F}$ is the field of complex numbers, $\mathbb{C}$.*

Suppose that $\boldsymbol{U}$ is a complex vector space and $T \in L(\boldsymbol{U})$ is a linear operator. We already know that T is triangulable and we deduce from the above theorem the following important consequence.

Corollary 2.15. *Suppose that T is a linear operator on the complex vector space $\mathbb{U}$. Then*

$$\det T = \prod_{\lambda \in \operatorname{spec}(T)} \lambda^{m_\lambda(T)}, \quad P_T(x) = \prod_{\lambda \in \operatorname{spec}(T)} (x-\lambda)^{m_\lambda(T)},$$

$$\sum_{\lambda \in \operatorname{spec}(T)} m_\lambda(T).$$ □

For any polynomial with complex coefficients

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{C}[x]$$

and any linear operator T on a complex vector space $\boldsymbol{U}$ we set

$$p(T) = a_0 \mathbb{1} + a_1 T + \cdots + a_n T^n.$$

Note that if $p(x), q(x) \in \mathbb{C}[x]$, and if we set $r(x) = p(x)q(x)$, then

$$r(T) = p(T)q(T).$$

Theorem 2.16 (Cayley-Hamilton)**.** *Suppose T is a linear operator on the complex vector space $\boldsymbol{U}$. If $P_T(x)$ is the characteristic polynomial of T, then*

$$P_T(T) = 0.$$

Proof. Fix a basis $\underline{\boldsymbol{e}} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ in which T is represented by the upper triangular matrix

$$A = \begin{bmatrix} \lambda_1 & * & * & * & * \\ 0 & \lambda_2 & * & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \lambda_{n-1} & * \\ 0 & \cdots & 0 & 0 & \lambda_n \end{bmatrix}$$

Note that

$$P_T(x) = \det(x\mathbb{1} - T) = \prod_{j=1}^n (x - \lambda_j)$$

so that

$$P_T(T) = \prod_{j=1}^n (T - \lambda_j \mathbb{1}).$$

For $j = 1, \dots, N$ we define

$$\boldsymbol{U}_j := \operatorname{span}\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_j\}.$$

and we set $\boldsymbol{U}_0 = \{0\}$. Note that for any $j = 1, \dots, n$ we have

$$(T - \lambda_j \mathbb{1})\boldsymbol{U}_j \subset \boldsymbol{U}_{j-1}.$$

Thus

$$P_T(T)\boldsymbol{U} = \prod_{j=1}^n (T - \lambda_j)\boldsymbol{U}_n = \prod_{j=1}^{n-1} (T - \lambda_j)\Big((T - \lambda_n \mathbb{1})\boldsymbol{U}_n \Big)$$

$$\subset \prod_{j=1}^{n-1}(T-\lambda_j)\boldsymbol{U}_{n-1} \subset \prod_{j=1}^{n-2}(T-\lambda_j)\boldsymbol{U}_{n-2} \subset \cdots \subset (T-\lambda_1)\boldsymbol{U}_1 \subset \{0\}.$$

In other words,

$$P_T(T)\boldsymbol{u} = 0, \ \ \forall \boldsymbol{u} \in \boldsymbol{U}.$$

□

Example 2.17. Consider the 2×2-matrix

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

Its characteristic polynomial is

$$P_A(x) = \det(xI - A) = \det \begin{bmatrix} x-3 & -2 \\ 2 & x+1 \end{bmatrix} = (x-3)(x+1)+4 = x^2-2x-3+4 = x^2-2x+1.$$

The Cayley-Hamilton theorem shows that

$$A^2 - 2A + 1 = 0.$$

Let us verify this directly. We have

$$A^2 = \begin{bmatrix} 5 & 8 \\ -4 & -3 \end{bmatrix}$$

and

$$A^2 - 2A + I = \begin{bmatrix} 5 & 4 \\ -4 & -3 \end{bmatrix} - 2\begin{bmatrix} 3 & 2 \\ -2 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0.$$

We can rewrite the last equality as

$$A^2 = 2A - I$$

so that

$$A^{n+2} = 2A^{n+1} - A^n,$$

We can rewrite this as

$$A^{n+2} - A^{n+1} = A^{n+1} - A^n = A^n - A^{n-1} = \cdots = A - I.$$

Hence

$$A^n = \underbrace{(A^n - A^{n-1}) + (A^{n-1} - A^{n-2}) + \cdots + (A - I)}_{=n(A-I)} + I = nA - (n-1)I.$$

□

2.4. **The Jordan normal form of a complex operator.** Let $\boldsymbol{U}$ be a complex n-dimensional vector space and $T : \boldsymbol{U} \to \boldsymbol{U}$. For each eigenvalue $\lambda \in \operatorname{spec}(T)$ we denote by $E_\lambda(T)$ the corresponding generalized eigenspace, i.e.,

$$\boldsymbol{u} \in E_\lambda(T) \Longleftrightarrow \exists k > 0 : \ \ (T - \lambda\mathbb{1})^k \boldsymbol{u} = 0.$$

From Proposition 2.12 we know that $E_\lambda(T)$ is an invariant subspace of T. Suppose that the spectrum of T consists of ℓ distinct eigenvalues,

$$\operatorname{spec}(T) = \{\, \lambda_1, \dots, \lambda_\ell \,\}.$$

Proposition 2.18.

$$\boldsymbol{U} = E_{\lambda_1}(T) \oplus \cdots \oplus E_{\lambda_\ell}(T).$$

Proof. It suffices to show that

$$\boldsymbol{U} = E_{\lambda_1}(T) + \cdots + E_{\lambda_\ell}(T), \tag{2.6a}$$

$$\dim \boldsymbol{U} = \dim E_{\lambda_1}(T) + \cdots + \dim E_{\lambda_\ell}(T). \tag{2.6b}$$

The equality (2.6a) follows from (2.4c) since

$$\dim \boldsymbol{U} = m_{\lambda_1}(T) + \cdots + m_{\lambda_\ell}(T) = \dim E_{\lambda_1}(T) + \cdots + \dim E_{\lambda_\ell}(T),$$

so we only need to prove (2.6b). Set

$$\boldsymbol{V} := E_{\lambda_1}(T) + \cdots + E_{\lambda_\ell}(T) \subset \boldsymbol{U}.$$

We have to show that $\boldsymbol{V} = \boldsymbol{U}$.

Note that since each of the generalized eigenspaces $E_\lambda(T)$ are invariant subspaces of T, so is there sum $\boldsymbol{V}$. Denote by S the restriction of T to $\boldsymbol{V}$, which we regard as an operator $S : \boldsymbol{V} \to \boldsymbol{V}$.

If $\lambda \in \operatorname{spec}(T)$ and $\boldsymbol{v} \in E_\lambda(T) \subset \boldsymbol{V}$, then

$$(S - \lambda\mathbb{1})^k \boldsymbol{v} = (T - \lambda\mathbb{1})^k \boldsymbol{v} = 0$$

for some $k \geq 0$. Thus λ is also an eigenvalue of S and $\boldsymbol{v}$ is also a generalized eigenvector of S. This proves that

$$\operatorname{spec}(T) \subset \operatorname{spec}(S),$$

and

$$E_\lambda(T) \subset E_\lambda(S), \;\; \forall \lambda \in \operatorname{spec}(T).$$

In particular, this implies that

$$\dim \boldsymbol{U} = \sum_{\lambda \in \operatorname{spec}(T)} \dim E_\lambda(T) \leq \sum_{\mu \in \operatorname{spec}(S)} \dim E_\mu(S) = \dim \boldsymbol{V} \leq \dim \boldsymbol{U}.$$

This shows that $\dim \boldsymbol{V} = \dim \boldsymbol{U}$ and thus $\boldsymbol{V} = \boldsymbol{U}$. □

For any $\lambda \in \operatorname{spec}(T)$ we denote by S_λ the restriction of T on the generalized eigenspace $E_\lambda(T)$. Since this is an invariant subspace of T we can regard S_λ as a linear operator

$$S_\lambda : E_\lambda(T) \to E_\lambda(T).$$

Arguing as in the proof of the above proposition we deduce that $E_\lambda(T)$ is also a generalized eigenspace of S_λ. Thus, the spectrum of S_λ consists of a single eigenvalue and

$$E_\lambda(T) = E_\lambda(S) = \ker(\lambda\mathbb{1} - S_\lambda)^{\dim E_\lambda(T)} = \ker(\lambda\mathbb{1} - S_\lambda)^{m_\lambda(T)}.$$

Thus, for any $\boldsymbol{u} \in E_\lambda(T)$ we have

$$(\lambda\mathbb{1} - S_\lambda)^{m_\lambda(T)} \boldsymbol{u} = 0,$$

i.e.,

$$(S_\lambda - \lambda\mathbb{1})^{m_\lambda(T)} = (-1)^{m_\lambda(T)} (\lambda\mathbb{1} - S_\lambda)^{m_\lambda(T)} = 0.$$

Definition 2.19. A linear operator $\boldsymbol{N} : \boldsymbol{U} \to \boldsymbol{U}$ is called *nilpotent* if $\boldsymbol{N}^k = 0$ for some $k > 0$. □

If we set $N_\lambda = S_\lambda - \lambda\mathbb{1}$ we deduce that the operator N_λ is nilpotent.

Definition 2.20. Let $\boldsymbol{N} : \boldsymbol{U} \to \boldsymbol{U}$ be a nilpotent operator on a finite dimensional complex vector spec $\boldsymbol{V}$. A *tower* of $\boldsymbol{N}$ is an *ordered* collection $\mathcal{T}$ of vectors

$$\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_k \in \boldsymbol{U}$$

satisfying the equalities

$$\boldsymbol{N}\boldsymbol{u}_1 = 0, \quad \boldsymbol{N}\boldsymbol{u}_2 = \boldsymbol{u}_1, \dots, \boldsymbol{N}\boldsymbol{u}_k = \boldsymbol{u}_{k-1}.$$

The vector $\boldsymbol{u}_1$ is called the *bottom* of the tower, the vector $\boldsymbol{u}_k$ is called the *top* of the tower, while the integer k is called the *height* of the tower. □

In Figure 1 we depicted a tower of height 4. Observe that the vectors in a tower are generalized eigenvectors of the corresponding nilpotent operator.

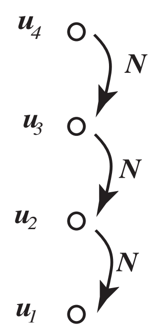

FIGURE 1. *Pancaking a tower of height* 4.

Towers interact in a rather pleasant way.

Proposition 2.21. *Suppose that $\boldsymbol{N} : \boldsymbol{U} \to \boldsymbol{U}$ is a nilpotent operator on a complex vector space $\boldsymbol{U}$ and $\mathcal{T}_1, \dots, \mathcal{T}_r$ are towers of $\boldsymbol{N}$ with bottoms $\boldsymbol{b}_1, \dots, \boldsymbol{b}_r$.*

If the bottom vectors $\boldsymbol{b}_1, \dots, \boldsymbol{b}_r$ are linearly independent, then the following hold.

(i) *The towers $\mathcal{T}_1, \dots, \mathcal{T}_r$ are mutually disjoint, i.e., $\mathcal{T}_i \cap \mathcal{T}_j = \emptyset$ if $i \neq j$.*

(ii) *The union*

$$\mathcal{T} = \mathcal{T}_1 \cup \cdots \cup \mathcal{T}_r$$

is a linearly independent family of vectors.

Proof. Denote by k_i the height of the tower $\mathcal{T}_i$ and set

$$k = k_1 + \cdots + k_r.$$

We will argue by induction on k, the sum of the heights of the towers.

For $k = 1$ the result is trivially true. Assume the result is true for all collections of towers with total heights $< k$ and linear independent bases, and we will prove that it is true for collection of towers with total heights $= k$.

Denote by $\boldsymbol{V}$ the subspace spanned by the union $\mathcal{T}$. It is an invariant subspace of $\boldsymbol{N}$, and we denote by S the restriction of $\boldsymbol{N}$ to $\boldsymbol{V}$. We regard S as a linear operator $S : \boldsymbol{V} \to \boldsymbol{V}$.

Denote by $\mathcal{T}_i'$ the tower obtained by removing the top of the tower $\mathcal{T}_i$ and set (see Figure 2)

$$\mathcal{T}' = \mathcal{T}_1' \cup \cdots \cup \mathcal{T}_r'.$$

Note that

$$\boldsymbol{R}(S) = \operatorname{span}(\mathcal{T}'). \tag{2.7}$$

The collection of towers $\mathcal{T}_1', \dots \mathcal{T}_r'$ has total height

$$k' = (k_1 - 1) + \cdots + (k_r - 1) = k - r < k.$$

Moreover the collection of bottoms is a subcollection of $\{\boldsymbol{b}_1, \dots, \boldsymbol{b}_r\}$ so it is linearly independent.

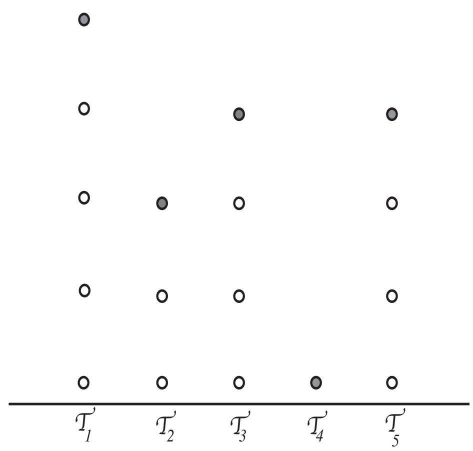

FIGURE 2. *A group of* 5 *towers. The tops are shaded and about to be removed.*

The inductive assumption implies that the towers $\mathcal{T}_1', \dots, \mathcal{T}_r'$ are mutually disjoint and their union $\mathcal{T}'$ is a linear independent family of vectors. Hence $\mathcal{T}'$ is a basis of $\boldsymbol{R}(S)$, and thus

$$\dim \boldsymbol{R}(S) = k' = k - r.$$

On the other hand

$$S\boldsymbol{b}_j = \boldsymbol{N}\boldsymbol{b}_j = 0, \quad \forall j = 1, \dots, r.$$

Since the vectors $\boldsymbol{b}_1, \dots, \boldsymbol{b}_r$ are linearly independent we deduce that

$$\dim \ker S \geq r,$$

From the rank-nullity theorem we deduce that

$$\dim \boldsymbol{V} = \dim \ker S + \dim \boldsymbol{R}(S) \geq r + k - r = r.$$

on the other hand

$$\dim \boldsymbol{V} = \dim \operatorname{span}(\mathcal{T}) \leq |\mathcal{T}| \leq k.$$

Hence

$$|\mathcal{T}| = \dim \boldsymbol{V} = k.$$

This proves that the towers $\mathcal{T}_1, \dots \mathcal{T}_r$ are mutually disjoint and their union is linearly independent. □

Theorem 2.22 (Jordan normal form of a nilpotent operator)**.** *Let $\boldsymbol{N} : \boldsymbol{U} \to \boldsymbol{U}$ be a nilpotent operator on an n-dimensional complex vector space $\boldsymbol{U}$. Then $\boldsymbol{U}$ has a basis consisting of a disjoint union of towers of $\boldsymbol{N}$.*

Proof. We will argue by induction on the dimension n of $\boldsymbol{U}$. For $n = 1$ the result is trivially true. We assume that the result is true for any nilpotent operator on a space of dimension $< n$ and we prove it is true for any nilpotent operator $\boldsymbol{N}$ on a space $\boldsymbol{U}$ of dimension n.

Observe that $\boldsymbol{V} = \boldsymbol{R}(\boldsymbol{N})$ is an invariant subspace of N. Moreover, since $\ker N \neq 0$, we deduce from the rank nullity formula that

$$\dim \boldsymbol{V} = \dim \boldsymbol{U} - \dim \ker \boldsymbol{N} < \dim \boldsymbol{U}.$$

Denote by M the restriction of $\boldsymbol{N}$ to $\boldsymbol{V}$. We view M as a linear operator $M : \boldsymbol{V} \to \boldsymbol{V}$. Clearly M is nilpotent. The induction assumption implies that there exist a basis of $\boldsymbol{V}$ consisting a mutually disjoint towers of M,

$$\mathcal{T}_1, \dots, \mathcal{T}_r.$$

For any $j = 1, \dots, r$ we denote by k_j the height of $\mathcal{T}_j$, by $\boldsymbol{b}_j$ the bottom of $\mathcal{T}_j$ and by $\boldsymbol{t}_j$ the top of $\mathcal{T}_j$. By construction

$$\dim \boldsymbol{R}(\boldsymbol{N}) = k_1 + \cdots + k_r.$$

Since $\boldsymbol{t}_j \in \boldsymbol{V} = \boldsymbol{R}(\boldsymbol{N})$ there exists $\boldsymbol{u}_j \in \boldsymbol{U}$ such that (see Figure 3)

$$\boldsymbol{t}_j = \boldsymbol{N}\boldsymbol{u}_j.$$

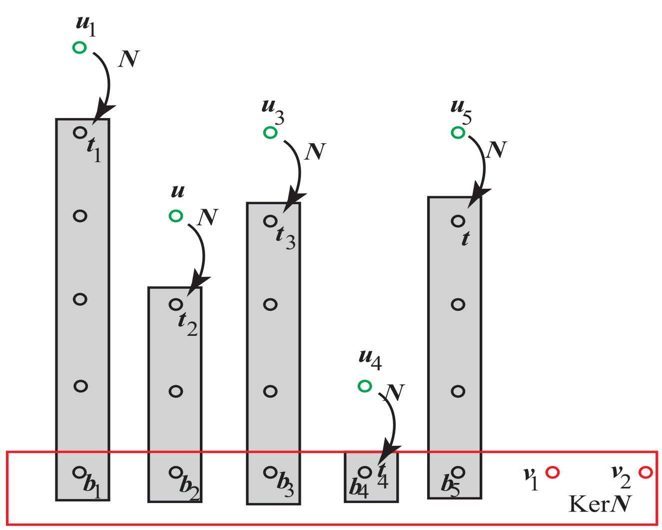

FIGURE 3. *Towers in* $\boldsymbol{R}(\boldsymbol{N})$.

Next observe that the bottoms $\boldsymbol{b}_1, \dots, \boldsymbol{b}_r$ belong to $\ker \boldsymbol{N}$ and are linearly independent, because they are a subfamily of the linearly independent family $\mathcal{T}_1 \cup \cdots \cup \mathcal{T}_r$. We can therefore extend the family $\{\boldsymbol{b}_1, \dots, \boldsymbol{b}_r\}$ to a basis of $\ker \boldsymbol{N}$,

$$\boldsymbol{b}_1, \dots, \boldsymbol{b}_r, \boldsymbol{v}_1, \dots, \boldsymbol{v}_s, \quad r + s = \dim \ker \boldsymbol{N}.$$

We obtain new towers $\hat{\mathcal{T}}_1, \dots, \hat{\mathcal{T}}_r, \hat{\mathcal{T}}_{r+1}, \dots, \hat{\mathcal{T}}_{r+s}$ defined by (see Figure 3)

$$\hat{\mathcal{T}}_1 := \mathcal{T}_1 \cup \{\boldsymbol{u}_1\}, \dots, \hat{\mathcal{T}}_r := \mathcal{T}_r \cup \{\boldsymbol{u}_r\}, \;\; \hat{\mathcal{T}}_{r+1} := \{\boldsymbol{v}_1\}, \dots, \hat{\mathcal{T}}_{r+s} := \{\boldsymbol{v}_s\}.$$

The sum of the heights of these towers is

$$(k_1 + 1) + (k_2 + 1) + \cdots + (k_r + 1) + \underbrace{1 + \cdots + 1}_{s}$$

$$= \underbrace{(k_1 + \cdots + k_r)}_{= \dim \boldsymbol{R}(\boldsymbol{N})} + \underbrace{(r + s)}_{= \dim \ker \boldsymbol{N}} = \dim \boldsymbol{U}.$$

By construction, their bottoms are linearly independent and Proposition 2.21 implies that they are mutually disjoint and their union is a linearly independent collection of vectors. The above computation shows that the number of elements in the union of these towers is equal to the dimension of $\boldsymbol{U}$. Thus, this union is a basis of $\boldsymbol{U}$. □

Definition 2.23. A *Jordan basis* of a nilpotent operator $\boldsymbol{N}: \boldsymbol{U} \to \boldsymbol{U}$ is a basis of $\boldsymbol{U}$ consisting of a disjoint union of towers of $\boldsymbol{N}$. □

Example 2.24. (a) Suppose that the nilpotent operator $\boldsymbol{N}: \boldsymbol{U} \to \boldsymbol{U}$ admits a Jordan basis consisting of a single tower

$$\boldsymbol{e}_1, \dots, \boldsymbol{e}_n.$$

Denote by C_n the matrix representing $\boldsymbol{N}$ in this basis. We use this basis to identify $\boldsymbol{U}$ with $\mathbb{C}^n$ and thus

$$\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \dots, \boldsymbol{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

From the equalities

$$\boldsymbol{N}\boldsymbol{e}_1 = 0, \quad \boldsymbol{N}\boldsymbol{e}_2 = \boldsymbol{e}_1, \quad \boldsymbol{N}\boldsymbol{e}_3 = \boldsymbol{e}_2, \dots$$

we deduce that the first column of C_n is trivial, the second column is $\boldsymbol{e}_1$, the 3-rd column is $\boldsymbol{e}_2$ etc. Thus C_n is the $n \times n$ matrix.

$$C_n = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

The matrix C_n is called a (nilpotent) *Jordan cell of size* n.

(b) Suppose that the nilpotent operator $\boldsymbol{N}: \boldsymbol{U} \to \boldsymbol{U}$ admits a Jordan basis consisting of mutually disjoint towers $\mathcal{T}_1, \dots, \mathcal{T}_r$ of heights $k_1, \dots, k_r$. For $j = 1, \dots, r$ we set

$$\boldsymbol{U}_j = \operatorname{span}(\mathcal{T}_j).$$

Observe that $\boldsymbol{U}_j$ is an invariant subspace of $\boldsymbol{N}$, $\mathcal{T}_j$ is a basis of $\boldsymbol{U}_j$ and we have a direct sum decomposition

$$\boldsymbol{U} = \boldsymbol{U}_1 \oplus \cdots \oplus \boldsymbol{U}_r.$$

The restriction of $\boldsymbol{N}$ to $\boldsymbol{U}_j$ is represented in the basis $\mathcal{T}_j$ by the Jordan cell C_{k_j} so that in the basis $\mathcal{T} = \mathcal{T}_1 \cup \cdots \cup \mathcal{T}_r$ the operator $\boldsymbol{N}$ has the block-matrix description

$$\begin{bmatrix} C_{k_1} & 0 & 0 & \cdots & 0 \\ 0 & C_{k_2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & C_{k_r} \end{bmatrix}. \qquad \square$$

We want to point out that the sizes of the Jordan cells correspond to the heights of the towers in a Jordan basis. While there may be several Jordan bases, the heights of the towers are the same in all of them; see Remark 2.26. In other words, these heights are invariants of the nilpotent operator. □

Definition 2.25. The *Jordan invariant* of a nilpotent operator N is the nonincreasing list of the sizes of the Jordan cells that describe the operator in a Jordan basis. □

Remark 2.26 (Algorithmic construction of a Jordan basis)**.** Here is how one constructs a Jordan basis of a nilpotent operator $\boldsymbol{N}:\boldsymbol{U}\to\boldsymbol{U}$ on a complex vector space $\boldsymbol{U}$ of dimension n.

(i) Compute $\boldsymbol{N}^2,\boldsymbol{N}^3,\dots$ and stop at the moment m when $\boldsymbol{N}^m=0$. Set

$$R_0=\boldsymbol{U},\ \ R_1=\boldsymbol{R}(\boldsymbol{N}),\ \ R_2=\boldsymbol{R}(\boldsymbol{N}^2),\dots,R_m=\boldsymbol{R}(\boldsymbol{N}^m)=\{0\}.$$

Observe that $R_1,R_2,\dots,R_m$ are invariant subspaces of $\boldsymbol{N}$, satisfying

$$R_0\supset R_1\supset R_2\supset\cdots,$$

(ii) Denote by N_k the restriction of $\boldsymbol{N}$ to R_k, viewed as an operator $N_k:R_k\to R_k$. Note that $N_{m-1}=0$, $N_0=\boldsymbol{N}$ and

$$R_k=\boldsymbol{R}(N_{k-1}),\ \ \forall k=1,\dots,m.$$

Set $r_k=\dim R_k$ so that $\dim\ker N_k=r_k-r_{k+1}$.

(iii) Construct a basis $\mathcal{B}_{m-1}$ of $R_{m-1}=\ker N_{m-1}$. $\mathcal{B}_{m-1}$ consists of r_{m-1} vectors.

(iv) For each $\boldsymbol{b}\in\mathcal{B}_{m-1}$ find a vector $\boldsymbol{t}(\boldsymbol{b})\in\boldsymbol{U}$ such that

$$\boldsymbol{b}=\boldsymbol{N}^{m-1}\boldsymbol{t}(\boldsymbol{b}).$$

For each $\boldsymbol{b}\in\mathcal{B}_{m-1}$ we thus obtain a tower of height m

$$\mathcal{T}_{m-1}(\boldsymbol{b})=\big\{\,\boldsymbol{b}=\boldsymbol{N}^{m-1}\boldsymbol{t}(\boldsymbol{b}),\boldsymbol{N}^{m-2}\boldsymbol{t}(\boldsymbol{b}),\dots,\boldsymbol{N}\boldsymbol{t}(\boldsymbol{b}),\boldsymbol{t}(\boldsymbol{b})\,\big\},\ \ \boldsymbol{b}\in\mathcal{B}_{m-1}.$$

(v) Extend $\mathcal{B}_{m-1}$ to a basis

$$\mathcal{B}_{m-2}=\mathcal{B}_{m-1}\cup\mathscr{C}_{m-2}$$

of $\ker N_{m-2}\subset R_{m-2}$.

(vi) For each $\boldsymbol{b}\in\mathscr{C}_{m-2}\subset R_{m-2}$ find $\boldsymbol{t}=\boldsymbol{t}(\boldsymbol{b})\in\boldsymbol{N}$ such that $\boldsymbol{N}^{m-2}\boldsymbol{t}=\boldsymbol{b}$. For each $\boldsymbol{b}\in\mathscr{C}_{m-2}$ we thus obtain a tower of height $m-1$

$$\mathcal{T}_{m-2}(\boldsymbol{b})=\big\{\,\boldsymbol{b}=\boldsymbol{N}^{m-2}\boldsymbol{t}(\boldsymbol{b}),\dots,\boldsymbol{N}\boldsymbol{t}(\boldsymbol{b}),\boldsymbol{t}(\boldsymbol{b})\,\big\},\ \ \boldsymbol{b}\in\mathcal{B}_{m-2}$$

(vii) Extent $\mathcal{B}_{m-2}$ to a basis

$$\mathcal{B}_{m-3}=\mathcal{B}_{m-2}\cup\mathscr{C}_{m-3}$$

of $\ker N_{m-3}\subset R_{m-3}$.

(viii) For each $\boldsymbol{b}\in\mathscr{C}_{m-3}\subset R_{m-3}$, find $\boldsymbol{t}(\boldsymbol{b})\in\boldsymbol{N}$ such that $\boldsymbol{N}^{m-3}\boldsymbol{t}(\boldsymbol{b})=\boldsymbol{b}$. For each $\boldsymbol{b}\in\mathscr{C}_{m-3}$ we thus obtain a tower of height $m-2$

$$\mathcal{T}_{m-3}(\boldsymbol{b})=\big\{\,\boldsymbol{b}=\boldsymbol{N}^{m-3}\boldsymbol{t}(\boldsymbol{b}),\dots,\boldsymbol{N}\boldsymbol{t}(\boldsymbol{b}),\boldsymbol{t}(\boldsymbol{b})\,\big\},\ \ \boldsymbol{b}\in\mathscr{C}_{m-3}$$

(ix) Iterate the previous two steps

(x) In the end we obtain a basis

$$\mathcal{B}_0=\mathcal{B}_{m-1}\cup\mathscr{C}_{m-2}\cup\cdots\cup\mathscr{C}_0$$

of $\ker N_0=\ker\boldsymbol{N}$, vectors $\boldsymbol{t}(\boldsymbol{b})$, $\boldsymbol{b}\in\mathscr{C}_j$, and towers

$$\mathcal{T}_j(\boldsymbol{b})=\big\{\,\boldsymbol{b}=\boldsymbol{N}^j\boldsymbol{t}(\boldsymbol{b}),\dots,\boldsymbol{N}\boldsymbol{t}(\boldsymbol{b}),\boldsymbol{t}(\boldsymbol{b})\,\big\},\ \ j=0,\dots,m-1,\ \ \boldsymbol{b}\in\mathscr{C}_j.$$

These towers form a Jordan basis of $\boldsymbol{N}$.

(xi) For uniformity set $\mathscr{C}_{m-1} = \mathcal{B}_{m-1}$, and for any $j = 1, \dots, m$ denote by c_j the cardinatlity of $\mathscr{C}_{j-1}$. In the above Jordan basis the operator N will be a dirrect sum of c_1 cells of dimension 1, c_2 cells of dimension 2, etc. We obtain the identities

$$\begin{gathered} n = c_1 + 2c_2 + \cdots + mc_m, \\ r_1 = c_2 + 2c_3 + \cdots + (m-1)c_{m-1}, \;\; r_2 = c_3 + \cdots + (m-2)c_{m-1}, \\ r_j = c_{j+1} + 2c_{j+2} + \cdots + (m-j)c_m, \;\; \forall j = 0, \dots, m-1. \end{gathered} \tag{2.8}$$

where $r_0 = n$.

If we treat the equalities (2.8) as a linear system with unknown $k_1, \dots, k_m$, we see that the matrix of this system is upper triangular with only 1-s along the diagonal. It is thus invertible so that the numbers $c_1, \dots, c_m$ are uniquely determined by the numbers r_j which are *invariants* of the operator N. This shows that the sizes of the Jordan cells are independent of the chosen Jordan basis.

If you are interested only in the sizes of the Jordan cells, all you have to do is find the integers m, $r_1, \dots, r_{m-1}$ and then solve the system (2.8). Exercise 2.13 explains how to explicitly express the c_j-s in terms of the r_j-s. □

Remark 2.27. To find the Jordan invariant of a nilpotent operator N on a complex n-dimensional space proceed as follows.

(i) Find the smallest integer m such that $N^m = 0$.
(ii) Find the ranks r_j of the matrices N^j, $j = 0, \dots, m-1$, where $N^0 : -I$.
(iii) Find the nonnegative integers $c_1, \dots, c_m$ by solving the linear system (2.8).
(iv) The Jordan invariant of N is the list

$$\underbrace{m, \dots, m}_{c_m}, \; \underbrace{(m-1), \dots, (m-1)}_{c_{m-1}}, \dots, \; \underbrace{1, \dots, 1}_{c_1}.$$

□

If $T : \boldsymbol{U} \to \boldsymbol{U}$ is an arbitrary linear operator on a complex n-dimensional space $\boldsymbol{U}$ with spectrum

$$\operatorname{spec}(T) = \{\lambda_1, \dots, \lambda_m\},$$

then we have a direct sum decomposition

$$\boldsymbol{U} = E_{\lambda_1}(T) \oplus \cdots \oplus E_{\lambda_m}(T),$$

where $E_{\lambda_j}(T)$ denotes the generalized eigenspace of T corresponding to the eigenvalue λ_j. The generalized eigenspace $E_{\lambda_j}(T)$ is an invariant subspace of T and we denote by S_{λ_j} the restriction of T to $E_{\lambda_j}(T)$. The operator $N_{\lambda_j} = S_{\lambda_j} - \lambda_j \mathbb{1}$ is nilpotent.

A *Jordan basis* of $\boldsymbol{U}$ is basis obtain as a union of Jordan bases of the nilpotent operators $N_{\lambda_1}, \dots, N_{\lambda_r}$. The matrix representing T in a Jordan basis is a direct sum of elementary *Jordan λ-cells*

$$C_n(\lambda) = C_n + \lambda I = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}.$$

Definition 2.28. The *Jordan invariant* of a complex operator T is a collection of lists, one list for every eigenvalue of of T. The list L_λ corresponding to the eigenvalue λ is the Jordan invariant of the nilpotent operator N_λ, the restriction of $T - \lambda\mathbb{1}$ to $E_\lambda(T)$ arranged in noincreasing order. □

Example 2.29. Consider the 4×4 matrix

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 2 & -4 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

viewed as a linear operator $\mathbb{C}^4 \to \mathbb{C}^4$.

Expanding along the first row and then along the last column we deduce that

$$P_A(x) = \det(xI - A) = (x-1)^2 \det \begin{bmatrix} x+1 & -1 \\ 4 & x-3 \end{bmatrix} = (x-1)^4.$$

Thus A has a single eigenvalue $\lambda = 1$ which has multiplicity 4. Set

$$N := A - I = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 2 & -4 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The matrix N is nilpotent. In fact we have

$$N^2 = 0$$

Upon inspecting N we see that ech of its columns is a multiple of the fiurst column. This means that the range of N is spanned by the vector

$$\boldsymbol{u}_1 := N\boldsymbol{e}_1 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 0 \end{bmatrix},$$

where $\boldsymbol{e}_1, \dots, \boldsymbol{e}_4$ denotes the canonical basis of $\mathbb{C}^4$.

The vector $\boldsymbol{u}_1$ is a tower in $\boldsymbol{R}(N)$ which we can extend to a taller tower of N

$$\mathcal{T}_1 = (\boldsymbol{u}_1, \boldsymbol{u}_2), \quad \boldsymbol{u}_2 = \boldsymbol{e}_1.$$

Next, we need to extend the basis $\{\boldsymbol{u}_1\}$ of $R(N)$ to a basis of $\ker N$. The rank nulity theorem tells us that

$$\dim \boldsymbol{R}(N) + \dim \ker N = 4,$$

so that $\dim \ker N = 3$. Thus, we need to find two more vectors $\boldsymbol{v}_1, \boldsymbol{v}_2$ so that the collection $\{\boldsymbol{u}_1, \boldsymbol{v}_1, \boldsymbol{v}_2\}$ is a basis of $\ker N$.

To find $\ker N$ we need to solve the linear system

$$N\boldsymbol{x} = 0, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}.$$

which we do using Gauss elimination, i.e., row operations on N. Observe that

$$N = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 2 & -4 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 & 0 \\ 2 & -4 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Hence

$$\ker N = \left\{ \boldsymbol{x} \in \mathbb{C}^4; \;\; x_1 - 2x_2 + x_3 = 0 \right\},$$

and thus a basis of $\ker N$ is

$$\boldsymbol{f}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \;\; \boldsymbol{f}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \;\; \boldsymbol{f}_3 := \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Observe that

$$\boldsymbol{u}_1 = 2\boldsymbol{f}_1 + \boldsymbol{f}_2,$$

and thus the collection $\{\boldsymbol{u}_1, \boldsymbol{f}_2, \boldsymbol{f}_3\}$ is also a basis of $\ker N$.

We now have a Jordan basis of N consisting of the towers

$$\mathcal{T}_1 = \{\boldsymbol{u}_1, \boldsymbol{u}_2\}, \;\; \mathcal{T}_2 = \{\boldsymbol{f}_2\}, \;\; \mathcal{T}_3 = \{\boldsymbol{f}_3\}.$$

In this basis the operator is described as a direct sum of three Jordan cells: a cell of dimension 2, and two cells of dimension 1. Thus the Jordan invariant of A consists of single list L_1 corresponding to the single eigenvalue 1. More precisely

$$L_1 = 2, 1, 1. \qquad \square$$

2.5. **Exercises.**

Exercise 2.1. Denote by $\boldsymbol{U}_3$ the space of polynomials of degree ≤ 3 with real coefficients in one variable x. We denote by $\mathcal{B}$ the canonical basis of $\boldsymbol{U}_3$,

$$\mathcal{B} = \big\{ 1, x, x^2, x^3 \big\}.$$

(a) Consider the linear operator $D : \boldsymbol{U}_3 \to \boldsymbol{U}_3$ given by

$$\boldsymbol{U}_3 \ni p \mapsto Dp = \frac{dp}{dx} \in \boldsymbol{U}_3.$$

Find the matrix that describes D in the canonical basis $\mathcal{B}$.
(b) Consider the operator $\Delta : \boldsymbol{U}_3 \to \boldsymbol{U}_3$ given by

$$(\Delta p)(x) = p(x+1) - p(x).$$

Find the matrix that describes Δ in the canonical basis $\mathcal{B}$.
(c) Show that for any $p \in \boldsymbol{U}_3$ the function

$$x \mapsto (\mathcal{L}p)(x) = \int_0^\infty e^{-t} \frac{dp(x+t)}{dx} dt$$

is also a polynomial of degree ≤ 3 and then find the matrix that describes $\mathcal{L}$ in the canonical basis $\mathcal{B}$. $\square$

Exercise 2.2. (a) For any $n \times n$ matrix $A = (a_{ij})_{1 \leq i,j \leq n}$ we define the *trace of* A as the sum of the diagonal entries of A,

$$\operatorname{tr} A := a_{11} + \cdots + a_{nn} = \sum_{i=1}^n a_{ij}.$$

Show that if A, B are two $n \times n$ matrices then

$$\operatorname{tr} AB = \operatorname{tr} BA.$$

(b) Let $\boldsymbol{U}$ be an n-dimensional $\mathbb{F}$-vector space, and $\underline{\boldsymbol{e}}$, $\underline{\boldsymbol{f}}$ be two bases of $\boldsymbol{U}$. Suppose that $T : \boldsymbol{U} \to \boldsymbol{U}$ is a linear operator represented in the basis $\underline{\boldsymbol{e}}$ by the matrix A and the basis $\underline{\boldsymbol{f}}$ by the matrix B. Prove that

$$\operatorname{tr} A = \operatorname{tr} B.$$

(The common value of these traces is called the *trace* of the operator T and it is denoted by $\operatorname{tr} T$.)
Hint: Use part (a) and (2.1).
(c) Consider the operator $A : \mathbb{F}^2 \to \mathbb{F}^2$ described by the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

Show that

$$P_T(x) = x^2 - \operatorname{tr} Ax + \det A.$$

(d) Let T be a linear operator on the n-dimensional $\mathbb{F}$-vector space. Prove that the characteristic polynomial of T has the form

$$P_T(x) = \det(x\mathbb{1} - T) = x^n - (\operatorname{tr} T)x^{n-1} + \cdots + (-1)^n \det T. \qquad \square$$

Exercise 2.3. Suppose $T : \boldsymbol{U} \to \boldsymbol{U}$ is a linear operator on the $\mathbb{F}$-vector space $\boldsymbol{U}$, and $\boldsymbol{V}_1, \boldsymbol{V}_2 \subset \boldsymbol{U}$ are invariant subspaces of T. Show that $\boldsymbol{V}_1 \cap \boldsymbol{V}_2$ and $\boldsymbol{V}_1 + \boldsymbol{V}_2$ are also invariant subspaces of T. $\square$

Exercise 2.4. Consider the Jacobi matrix

$$J_n = \begin{bmatrix} 2 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 2 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 2 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 2 \end{bmatrix}$$

(a) Let $P_n(x)$ denote the characteristic polynomial of J_n,

$$P_n(x) = \det(xI - J_n).$$

Show that

$$P_1(x) = x - 2, \quad P_2(x) = x^2 - 4x + 3 = (x-1)(x-3),$$
$$P_n(x) = (x-2)P_{n-1}(x) - P_{n-2}(x), \quad \forall n \geq 3.$$

(b) Show that all the eigenvalues of J_4 are real and distinct, and then conclude that the matrices $\mathbb{1}, J_4, J_4^2, J_4^3$ are linearly independent. □

Exercise 2.5. Consider the $n \times n$ matrix $A = (a_{ij})_{1 \leq i,j \leq n}$, where $a_{ij} = 1$, for all i, j, which we regard as a linear operator $\mathbb{C}^n \to \mathbb{C}^n$. Consider the vector

$$\boldsymbol{c} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{C}^n.$$

(a) Compute $A\boldsymbol{c}$.
(b) Compute $\dim \boldsymbol{R}(A)$, $\dim \ker A$ and then determine $\operatorname{spec}(A)$.
(c) Find the characteristic polynomial of A. □

Exercise 2.6. Find the eigenvalues and the eigenvectors of the circulant matrix described in Exercise 1.16. □

Exercise 2.7. Prove the equality (2.7) that appears in the proof of Proposition 2.21. □

Exercise 2.8. Let $T : \boldsymbol{U} \to \boldsymbol{U}$ be a linear operator on the finite dimensional complex vector space $\boldsymbol{U}$. Suppose that m is a positive integer and $\boldsymbol{u} \in \boldsymbol{U}$ is a vector such that

$$T^{m-1}\boldsymbol{u} \neq 0 \text{ but } T^m\boldsymbol{u} = 0.$$

Show that the vectors

$$\boldsymbol{u},\ T\boldsymbol{u}, \dots, T^{m-1}\boldsymbol{u}$$

are linearly independent. □

Exercise 2.9. Let $T : \boldsymbol{U} \to \boldsymbol{U}$ be a linear operator on the finite dimensional complex vector space $\boldsymbol{U}$. Show that if

$$\dim \ker T^{\dim \boldsymbol{U} - 1} \neq \dim \ker T^{\dim \boldsymbol{U}},$$

then T is a nilpotent operator and

$$\dim \ker T^j = j, \quad \forall j = 1, \dots, \dim \boldsymbol{U}.$$

Can you construct an example of operator T satisfying the above properties? □

Exercise 2.10. Let S, T be two linear operators on the finite dimensional complex vector space $\boldsymbol{U}$. Show that ST is nilpotent if and only if TS is nilpotent. $\square$

Exercise 2.11. Find a Jordan basis of the linear operator $\mathbb{C}^4 \to \mathbb{C}^4$ described by the $\times 4$ matrix

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

Exercise 2.12. Find a Jordan basis of the linear operator $\mathbb{C}^7 \to \mathbb{C}^7$ given by the matrix

$$A = \begin{bmatrix} 3 & 2 & 1 & 1 & 0 & -3 & 1 \\ -1 & 2 & 0 & 0 & 1 & -1 & 1 \\ 1 & 1 & 3 & 0 & -6 & 5 & -1 \\ 0 & 2 & 1 & 4 & 1 & -3 & 1 \\ 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 & 5 & 0 \\ 0 & 0 & 0 & 0 & -2 & 0 & 5 \end{bmatrix}. \qquad \square$$

Exercise 2.13. (a) Find the inverse of the $m \times m$ matrix

$$A = \begin{bmatrix} 1 & 2 & 3 & \cdots & m-1 & m \\ 0 & 1 & 2 & \cdots & m-2 & m-1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 2 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

(b) Find the Jordan normal form of the above matrix. $\square$

3. Euclidean spaces

In the sequel $\mathbb{F}$ will denote either the field $\mathbb{R}$ of real numbers, or the field $\mathbb{C}$ of complex numbers. Any complex number has the form $z = a + b\boldsymbol{i}$, $\boldsymbol{i} = \sqrt{-1}$, so that $\boldsymbol{i}^2 = -1$. The real number a is called the *real part* of z and it is denoted by $\mathbf{Re}\, z$. The real number b is called the *imaginary part* of z and its is denoted by $\mathbf{Im}\, z$.

The *conjugate* of a complex number $z = a + b\boldsymbol{i}$, is the complex number

$$\bar{z} = a - b\boldsymbol{i}.$$

In particular, any real number is equal to its conjugate. Note that

$$z + \bar{z} = 2\,\mathbf{Re}\, z, \quad z - \bar{z} = 2\boldsymbol{i}\,\mathbf{Im}\, z.$$

The *norm* or *absolute value* of a complex number $z = a + b\boldsymbol{i}$ is the real number

$$|z| = \sqrt{a^2 + b^2}.$$

Observe that

$$|z|^2 = a^2 + b^2 = (a + b\boldsymbol{i})(a - b\boldsymbol{i}) = z\bar{z}.$$

In particular, if $z \neq 0$ we have

$$\frac{1}{z} = \frac{1}{|z|^2}\bar{z}.$$

3.1. Inner products.

Let $\boldsymbol{U}$ be an $\mathbb{F}$-vector space.

Definition 3.1. An inner product on $\boldsymbol{U}$ is a map

$$\langle -, - \rangle : \boldsymbol{U} \times \boldsymbol{U} \to \mathbb{F}, \quad \boldsymbol{U} \times \boldsymbol{U} \ni (\boldsymbol{u}_1, \boldsymbol{u}_2) \mapsto \langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle \in \mathbb{F},$$

satisfying with the following condition.

(i) *Linearity in the first variable*, i.e., $\forall \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{u}_2 \in \boldsymbol{U}$, $x, y \in \mathbb{F}$ we have

$$\langle x\boldsymbol{u} + y\boldsymbol{v}, \boldsymbol{v}_2 \rangle = x\langle \boldsymbol{u}, \boldsymbol{u}_2 \rangle + y\langle \boldsymbol{v}, \boldsymbol{v}_2 \rangle.$$

(ii) *Conjugate linearity in the second variable*, i.e., $\forall \boldsymbol{u}_1, , \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{U}$, $x, y \in \mathbb{F}$ we have

$$\langle \boldsymbol{u}_1, x\boldsymbol{u} + y\boldsymbol{v} \rangle = \bar{x}\langle \boldsymbol{u}_1, \boldsymbol{u} \rangle + \bar{y}\langle \boldsymbol{u}_1, \boldsymbol{v} \rangle.$$

(iii) *Hermitian property*, i.e., $\forall \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{U}$ we have

$$\langle \boldsymbol{v}, \boldsymbol{u} \rangle = \overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}.$$

(iv) *Positive definiteness*, i.e.,

$$\langle \boldsymbol{u}, \boldsymbol{u} \rangle \geq 0, \quad \forall \boldsymbol{u} \in \boldsymbol{U},$$

and

$$\langle \boldsymbol{u}, \boldsymbol{u} \rangle = 0 \Longleftrightarrow \boldsymbol{u} = 0.$$

A vector space $\boldsymbol{u}$ equipped with an inner product is called an *Euclidean* space. □

Example 3.2. (a) ***The standard real n-dimensional Euclidean space.*** The vector space $\mathbb{R}^n$ is equipped with a canonical inner product

$$\langle -, - \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}.$$

More precisely if

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \quad \boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n,$$

then

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = u_1 v_1 + \cdots + u_n v_n = \sum_{k=1}^n u_k v_k = \boldsymbol{u}^\dagger \cdot \boldsymbol{v}.$$

You can verify that this is indeed an inner product, i.e., it satisfies the conditions (i)-(iv) in Definition 3.1.

(b) ***The standard complex n-dimensional Euclidean space.*** The vector space $\mathbb{C}^n$ is equipped with a canonical inner product

$$\langle -, - \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}.$$

More precisely if

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \quad \boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{C}^n$$

then

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = u_1 \bar{v}_1 + \cdots + u_n \bar{v}_n = \sum_{k=1}^n u_k \bar{v}_k.$$

You can verify that this is indeed an inner product.

(c) Denote by $\mathcal{P}_n$ the vector space of polynomials with real coefficients and degree $\leq n$. We can define an inner product on $\mathcal{P}_n$

$$\langle -, - \rangle : \mathcal{P}_n \times \mathcal{P}_n \to \mathbb{R},$$

by setting

$$\langle p, q \rangle = \int_0^1 p(x) q(x) dx, \quad \forall p, q \in \mathcal{P}_n.$$

You can verify that this is indeed an inner product

(d) Any finite dimensional $\mathbb{F}$-vector space $\boldsymbol{U}$ admits an inner product. Indeed, if we fix a basis $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ of $\boldsymbol{U}$, then we define

$$\langle -, - \rangle : \boldsymbol{U} \times \boldsymbol{U} \to \mathbb{F},$$

by setting

$$\Big\langle u_1 \boldsymbol{e}_1 + \cdots + u_n \boldsymbol{e}_n \, , \, v_1 \boldsymbol{e}_1 + \cdots + v_n \boldsymbol{e}_n \Big\rangle = u_1 \bar{v}_1 + \cdots + u_n \bar{v}_n. \qquad \square$$

3.2. **Basic properties of Euclidean spaces.** Suppose that $(\boldsymbol{U}, \langle -, - \rangle)$ is an Euclidean vector space. We define the *norm* or *length* of a vector $\boldsymbol{u} \in \boldsymbol{U}$ to be the nonnegative number

$$\|\boldsymbol{u}\| := \sqrt{\langle \boldsymbol{u}, \boldsymbol{u} \rangle}.$$

Example 3.3. In the standard Euclidean space $\mathbb{R}^n$ of Example 3.2(a) we have

$$\|\boldsymbol{u}\| = \sqrt{\sum_{k=1}^n u_k^2}. \qquad \square$$

Theorem 3.4 (Cauchy-Schwarz inequality). *For any $\boldsymbol{u},\boldsymbol{v}\in\boldsymbol{U}$ we have*

$$\bigl|\,\langle\boldsymbol{u},\boldsymbol{v}\rangle\,\bigr|\leq\|\boldsymbol{u}\|\cdot\|\boldsymbol{v}\|.$$

Moreover, the equality is achieved if and only if the vectors $\boldsymbol{u},\boldsymbol{v}$ are linearly dependent, i.e., collinear.

Proof. If $\langle\boldsymbol{u},\boldsymbol{v}\rangle=0$, then the inequality is trivially satisfied. Hence we can assume that $\langle\boldsymbol{u},\boldsymbol{v}\rangle=0$. In particular, $\boldsymbol{u},\boldsymbol{v}\neq 0$.

From the positive definiteness of the inner product we deduce that for any $x\in\mathbb{F}$ we have

$$\begin{aligned}0\leq\|\boldsymbol{u}-x\boldsymbol{v}\|^2=\langle\boldsymbol{u}-x\boldsymbol{v},\boldsymbol{u}-x\boldsymbol{v}\rangle&=\langle\boldsymbol{u},\boldsymbol{u}-x\boldsymbol{v}\rangle-x\langle\boldsymbol{v},\boldsymbol{u}-x\boldsymbol{v}\rangle\\&=\langle\boldsymbol{u},\boldsymbol{u}\rangle-\bar{x}\langle\boldsymbol{u},\boldsymbol{v}\rangle-x\langle\boldsymbol{v},\boldsymbol{u}\rangle+x\bar{x}\langle\boldsymbol{v},\boldsymbol{v}\rangle\\&=\|\boldsymbol{u}\|^2-\bar{x}\langle\boldsymbol{u},\boldsymbol{v}\rangle-x\langle\boldsymbol{v},\boldsymbol{u}\rangle+|x|^2\|\boldsymbol{v}\|^2.\end{aligned}$$

If we let

$$x_0=\frac{1}{\|\boldsymbol{v}\|^2}\langle\boldsymbol{u},\boldsymbol{v}\rangle,\tag{3.1}$$

then

$$\bar{x_0}\langle\boldsymbol{u},\boldsymbol{v}\rangle+x_0\langle\boldsymbol{v},\boldsymbol{u}\rangle=\frac{1}{\|\boldsymbol{v}\|^2}\overline{\langle\boldsymbol{u},\boldsymbol{v}\rangle}\langle\boldsymbol{u},\boldsymbol{v}\rangle+\frac{1}{\|\boldsymbol{v}\|^2}\langle\boldsymbol{u},\boldsymbol{v}\rangle\overline{\langle\boldsymbol{u},\boldsymbol{v}\rangle}=2\frac{|\langle\boldsymbol{u},\boldsymbol{v}\rangle|^2}{\|\boldsymbol{v}\|^2},$$

$$|x_0|^2\|\boldsymbol{v}\|^2=\frac{|\langle\boldsymbol{u},\boldsymbol{v}\rangle|^2}{\|\boldsymbol{v}\|^2},$$

and thus

$$0\leq\|\boldsymbol{u}-x_0\boldsymbol{v}\|^2=\|\boldsymbol{u}\|^2-2\frac{|\langle\boldsymbol{u},\boldsymbol{v}\rangle|^2}{\|\boldsymbol{v}\|^2}+\frac{|\langle\boldsymbol{u},\boldsymbol{v}\rangle|^2}{\|\boldsymbol{v}\|^2}=\|\boldsymbol{u}\|^2-\frac{|\langle\boldsymbol{u},\boldsymbol{v}\rangle|^2}{\|\boldsymbol{v}\|^2}.\tag{3.2}$$

Thus

$$\frac{|\langle\boldsymbol{u},\boldsymbol{v}\rangle|^2}{\|\boldsymbol{v}\|^2}\leq\|\boldsymbol{u}\|^2$$

so that

$$\bigl|\,\langle\boldsymbol{u},\boldsymbol{v}\rangle\,\bigr|^2\leq\|\boldsymbol{u}\|^2\cdot\|\boldsymbol{v}\|^2.$$

Note that if $0=\langle\boldsymbol{u},\boldsymbol{v}\rangle=\|\boldsymbol{u}\|\cdot\|\boldsymbol{v}\|$ then at least one of the vectors $\boldsymbol{u},\boldsymbol{v}$ must be zero.

If $0\neq\langle\boldsymbol{u},\boldsymbol{v}\rangle=\|\boldsymbol{u}\|\cdot\|\boldsymbol{v}\|$, then by choosing x_0 as in (3.1) we deduce as in (3.2) that

$$\|\boldsymbol{u}-x_0\boldsymbol{v}\|=0.$$

Hence $\boldsymbol{u}=x_0\boldsymbol{v}$. □

Remark 3.5. The Cauchy-Schwarz theorem is a rather nontrivial result, which in skilled hands can produce remarkable consequences. Observe that if $\boldsymbol{U}$ is the standard real Euclidean space of Example 3.2(a), then the Cauchy-Schwarz inequality implies that for any real numbers $u_1,v_1,\dots,u_n,v_n$ we have

$$\left|\sum_{k=1}^n u_kv_k\right|\leq\sqrt{\sum_{k=1}^n u_k^2}\cdot\sqrt{\sum_{k=1}^n v_k^2}$$

If we square both sides of the above inequality we deduce

$$\left(\sum_{k=1}^n u_kv_k\right)^2\leq\left(\sum_{k=1}^n u_k^2\right)\cdot\left(\sum_{k=1}^n v_k^2\right).\tag{3.3}$$

□

Observe that if $\boldsymbol{u},\boldsymbol{v}$ are two nonzero vectors in an Euclidean space $(\boldsymbol{U},\langle -,-\rangle)$, then the Cauchy-Schwarz inequality implies that

$$\frac{\mathbf{Re}\langle \boldsymbol{u},\boldsymbol{v}\rangle}{\|\boldsymbol{u}\|\cdot\|\boldsymbol{v}\|}\in[-1,1].$$

Thus there exists a unique $\theta\in[0,\pi]$ such that

$$\cos\theta=\frac{\mathbf{Re}\langle \boldsymbol{u},\boldsymbol{v}\rangle}{\|\boldsymbol{u}\|\cdot\|\boldsymbol{v}\|}.$$

This angle θ is called the *angle* between the two nonzero vectors $\boldsymbol{u},\boldsymbol{v}$. We denote it by $\measuredangle(\boldsymbol{u},\boldsymbol{v})$. In particular, we have, by definition,

$$\cos\measuredangle(\boldsymbol{u},\boldsymbol{v})=\frac{\mathbf{Re}\langle \boldsymbol{u},\boldsymbol{v}\rangle}{\|\boldsymbol{u}\|\cdot\|\boldsymbol{v}\|}. \tag{3.4}$$

Note that if the two vectors $\boldsymbol{u},\boldsymbol{v}$ were perpendicular in the classical sense, i.e., $\measuredangle(\boldsymbol{u},\boldsymbol{v})=\frac{\pi}{2}$, then $\mathbf{Re}\langle \boldsymbol{u},\boldsymbol{v}\rangle=0$. This justifies the following notion.

Definition 3.6. Two vectors $\boldsymbol{u},\boldsymbol{v}$ in an Euclidean vector space $(\boldsymbol{U},\langle -,-\rangle)$ are said to be *orthogonal* if $\langle \boldsymbol{u},\boldsymbol{v}\rangle=0$. We will indicate the orthogonality of two vectors $\boldsymbol{u},\boldsymbol{v}$ using the notation $\boldsymbol{u}\perp\boldsymbol{v}$. □

☞ *In the remainder of this subsection we fix an Euclidean space* $(\boldsymbol{U},\langle -,-\rangle)$.

Theorem 3.7 (Pythagora)**.** *If* $\boldsymbol{u}\perp\boldsymbol{v}$*, then*

$$\|\boldsymbol{u}+\boldsymbol{v}\|^2=\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2.$$

Proof. We have

$$\|\boldsymbol{u}+\boldsymbol{v}\|^2=\langle \boldsymbol{u}+\boldsymbol{v},\boldsymbol{u}+\boldsymbol{v}\rangle=\langle \boldsymbol{u},\boldsymbol{u}\rangle+\underbrace{\langle \boldsymbol{u},\boldsymbol{v}\rangle}_{=0}+\underbrace{\langle \boldsymbol{v},\boldsymbol{u}\rangle}_{=0}+\langle \boldsymbol{v},\boldsymbol{v}\rangle$$

$$=\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2.$$

□

Theorem 3.8 (Triangle inequality)**.**

$$\|\boldsymbol{u}+\boldsymbol{v}\|\leq\|\boldsymbol{u}\|+\|\boldsymbol{v}\|,\ \ \forall\boldsymbol{u},\boldsymbol{v}\in\boldsymbol{U}.$$

Proof. Observe that the inequality is can be rewritten equivalently as

$$\|\boldsymbol{u}+\boldsymbol{v}\|^2\leq\big(\,\|\boldsymbol{u}\|+\|\boldsymbol{v}\|\,\big)^2.$$

Observe that

$$\|\boldsymbol{u}+\boldsymbol{v}\|^2=\langle \boldsymbol{u}+\boldsymbol{v},\boldsymbol{u}+\boldsymbol{v}\rangle=\langle \boldsymbol{u},\boldsymbol{u}\rangle+\langle \boldsymbol{u},\boldsymbol{v}\rangle+\langle \boldsymbol{v},\boldsymbol{u}\rangle+\langle \boldsymbol{v},\boldsymbol{v}\rangle$$

$$\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2+2\Re\langle \boldsymbol{u},\boldsymbol{v}\rangle\leq\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2+2\big|\,\langle \boldsymbol{u},\boldsymbol{v}\rangle\,\big|$$

(use the Cauchy-Schwarz inequality

$$\leq\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2+2\|\boldsymbol{u}\|\cdot\|\boldsymbol{v}\|=\big(\,\|\boldsymbol{u}\|+\|\boldsymbol{v}\|\,\big)^2.$$

□

3.3. **Orthonormal systems and the Gramm-Schmidt procedure.** In the sequel $\boldsymbol{U}$ will denote an n-dimensional Euclidean $\mathbb{F}$-vector space. We will denote the inner product on $\boldsymbol{U}$ by $\langle -,-\rangle$.

Definition 3.9. A family of *nonzero* vectors

$$\{\nu_1,\dots,\boldsymbol{u}_k\}\subset\boldsymbol{U}$$

is called *orthogonal* if

$$\boldsymbol{u}_i\perp\boldsymbol{u}_j,\ \ \forall i\neq j.$$

An orthogonal family

$$\{\boldsymbol{u}_1,\dots,\boldsymbol{u}_k\}\subset\boldsymbol{U}$$

is called *orthonormal* if

$$\|\boldsymbol{u}_i\|=1,\ \ \forall i=1,\dots,k.$$

A basis of $\boldsymbol{U}$ is called *orthogonal* (respectively *orthonormal*) if it is an orthogonal (respectively orthonormal) family. □

Proposition 3.10. *Any orthogonal family in $\boldsymbol{U}$ is linearly independent.*

Proof. Suppose that

$$\{\boldsymbol{u}_1,\dots,\boldsymbol{u}_k\}\subset\boldsymbol{U}$$

is an orthogonal family. If $x_1,\dots,x_k\in\mathbb{F}$ are such that

$$x_1\boldsymbol{u}_1+\cdots+x_k\boldsymbol{u}_k=0,$$

then taking the inner product with $\boldsymbol{u}_j$ of both sides in the above equality we deduce

$$0=x_1\langle\boldsymbol{u}_1,\boldsymbol{u}_j\rangle+\cdots+x_{j-1}\langle\boldsymbol{u}_{j-1},\boldsymbol{u}_j\rangle+x_j\langle\boldsymbol{u}_j,\boldsymbol{u}_j\rangle+x_{j+1}\langle\boldsymbol{u}_{j+1},\boldsymbol{u}_j\rangle+\cdots+x_n\langle\boldsymbol{u}_n,\boldsymbol{u}_j\rangle$$

$(\langle\boldsymbol{u}_i,\boldsymbol{u}_j\rangle=0,\forall i\neq j)$

$$=x_j\|\boldsymbol{u}_j\|^2.$$

Since $\boldsymbol{u}_j\neq 0$ we deduce $x_j=0$. This happens for any $j=1,\dots,k$, proving that the family is linearly independent. □

Theorem 3.11 (Gramm-Schmidt)**.** *Suppose that*

$$\{\boldsymbol{u}_1,\dots,\boldsymbol{u}_k\}\subset\boldsymbol{U}$$

is a linearly independent *family. Then there exists an* orthonormal *family*

$$\{\boldsymbol{e}_1,\dots,\boldsymbol{e}_k\}\subset\boldsymbol{U}$$

such that

$$\operatorname{span}\{\boldsymbol{u}_1,\dots,\boldsymbol{u}_j\}=\operatorname{span}\{\boldsymbol{e}_1,\dots,\boldsymbol{e}_j\},\ \ \forall j=1,\dots,k.$$

Proof. We will argue by induction on k. For $k=1$, if $\{\boldsymbol{u}_1\}\subset\boldsymbol{U}$ is a linearly independent family, then $\boldsymbol{u}_1\neq 0$ and we set

$$\boldsymbol{e}_1:=\frac{1}{\|\boldsymbol{u}_1\|}\boldsymbol{u}_1.$$

Clearly $\{\boldsymbol{e}_1\}$ is an orthonormal family spanning the same subspace as $\{\boldsymbol{u}_1\}$.

Suppose that the result is true for any linearly independent family of vectors con sting of $(k-1)$ vectors. We need to prove that the result is true for linearly independent families consisting of k vectors. Let

$$\{\boldsymbol{u}_1,\dots,\boldsymbol{u}_k\}\subset\boldsymbol{U}$$

be such a family. The induction assumption implies that we can find an orthonormal system

$$\{\, \boldsymbol{e}_1, \dots, \boldsymbol{e}_{k-1} \,\}$$

such that

$$\operatorname{span}\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_j\,\} = \operatorname{span}\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_j\,\},\ \ \forall j = 1, \dots, k-1.$$

Define

$$\boldsymbol{v}_k := \langle \boldsymbol{u}_k, \boldsymbol{e}_1\rangle \boldsymbol{e}_1 + \ldots + \langle \boldsymbol{u}_k, \boldsymbol{e}_{k-1}\rangle \boldsymbol{e}_{k-1},$$
$$\boldsymbol{f}_k := \boldsymbol{u}_k - \boldsymbol{v}_k.$$

Observe that

$$\boldsymbol{v}_k \in \operatorname{span}\{\, \boldsymbol{e}_1, \dots, \boldsymbol{e}_{k-1}\,\} = \operatorname{span}\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_{k-1}\,\}.$$

Since $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_k\}$ is a linearly independent family we deduce that

$$\boldsymbol{u}_k \notin \{\, \boldsymbol{e}_1, \dots, \boldsymbol{e}_{k-1}\,\}$$

so that $\boldsymbol{f}_k = \boldsymbol{u}_k - \boldsymbol{v}_k \neq 0$. We can now set

$$\boldsymbol{e}_k := \frac{1}{\|\boldsymbol{f}_k\|}\boldsymbol{f}_k.$$

By construction $\|\boldsymbol{e}_k\| = 1$. Also note that if $1 \leq j < k$, then

$$\langle \boldsymbol{f}_k, \boldsymbol{e}_j\rangle = \langle \boldsymbol{u}_k - \boldsymbol{v}_k, \boldsymbol{e}_j\rangle = \langle \boldsymbol{u}_k, \boldsymbol{e}_j\rangle - \langle \boldsymbol{v}_k, \boldsymbol{e}_j\rangle$$
$$= \langle \boldsymbol{u}_k, \boldsymbol{e}_j\rangle - \Big\langle \underbrace{\langle \boldsymbol{u}_k, \boldsymbol{e}_1\rangle \boldsymbol{e}_1 + \ldots + \langle \boldsymbol{u}_k, \boldsymbol{e}_{k-1}\rangle \boldsymbol{e}_{k-1}}_{=\boldsymbol{v}_k}, \boldsymbol{e}_j \Big\rangle$$
$$\langle \boldsymbol{u}_k, \boldsymbol{e}_j\rangle - \langle \boldsymbol{u}_j, \boldsymbol{e}_j\rangle \cdot \langle \boldsymbol{e}_j \boldsymbol{e}_j\rangle = 0.$$

This proves that $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_{k-1}, \boldsymbol{e}_k\}$ is an orthonormal family.

Finally observe that

$$\boldsymbol{u}_k = \boldsymbol{v}_k + \boldsymbol{f}_k = \boldsymbol{v}_k + \|\boldsymbol{f}_k\|\boldsymbol{e}_k.$$

Since

$$\boldsymbol{v}_k \in \operatorname{span}\{\, \boldsymbol{e}_1, \dots, \boldsymbol{e}_{k-1}\,\}$$

we deduce

$$\boldsymbol{u}_k \in \operatorname{span}\{\, \boldsymbol{e}_1, \dots, \boldsymbol{e}_{k-1}, \boldsymbol{e}_k\,\}$$

and thus

$$\operatorname{span}\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_k\,\} = \operatorname{span}\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k\,\}$$

□

Remark 3.12. The strategy used in the proof of the above theorem is as important as the theorem itself. The procedure we used to produce the orthonormal family $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k\}$. This procedure goes by the name of the *Gramm-Schmidt* procedure. To understand how it works we consider a simple case, when $\boldsymbol{U}$ is the space $\mathbb{R}^2$ equipped with the canonical inner product.

Suppose that

$$\boldsymbol{u}_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix},\ \ \boldsymbol{u}_2 := \begin{bmatrix} 5 \\ 0 \end{bmatrix}.$$

Then

$$\|\boldsymbol{u}_1\|^2 = 3^2 + 4^2 = 9 + 16 = 25,$$

so that and we set

$$\boldsymbol{e}_1 := \frac{1}{\|\boldsymbol{u}_1\|}\boldsymbol{u}_1 = \begin{bmatrix} \frac{3}{5} \\ \\ \frac{4}{5} \end{bmatrix}.$$

Next

$$v_2 = \langle u_2, e_1\rangle e_1 = 3e_1, \quad f_2 = u_2 - v_2 = \left[\begin{array}{c} 5 \\ \\ 0 \end{array}\right] - \left[\begin{array}{c} \frac{9}{5} \\ \\ \frac{12}{5} \end{array}\right]$$

$$= \left[\begin{array}{c} \frac{16}{5} \\ \\ -\frac{12}{5} \end{array}\right] = 4\left[\begin{array}{c} \frac{4}{5} \\ \\ -\frac{3}{5} \end{array}\right].$$

We see that $f_2 = 4$ and thus

$$e_2 = \left[\begin{array}{c} \frac{4}{5} \\ \\ -\frac{3}{5} \end{array}\right]. \qquad \square$$

The Gramm-Schmidt theorem has many useful consequences. We will discuss a few of them

Corollary 3.13. *Any finite dimensional Euclidean vector space (over $\mathbb{F} = \mathbb{R}, \mathbb{C}$) admits an orthonormal basis.*

Proof. Apply Theorem 3.11 to a basis of the vector space. $\square$

The orthonormal bases of an Euclidean space have certain computational advantages. Suppose that

$$e_1, \dots, e_n$$

is an orthonormal basis of the Euclidean space U. Then the coordinates of a vector $u \in U$ in this basis are easily computed. More precisely, if

$$u = x_1e_1 + \cdots + x_ne_n, \quad x_1, \dots, x_n \in \mathbb{F}, \tag{3.5}$$

then

$$x_j = \langle u, e_j\rangle \quad \forall j = 1, \dots, n. \tag{3.6}$$

Indeed, the equality (3.5) implies that

$$\langle u, e_j\rangle = \big\langle\, x_1e_1 + \cdots + x_ne_n, e_j \,\big\rangle = x_j\langle e_j, e_j\rangle = x_j,$$

where at the last step we used the orthonormality condition which translates to

$$\langle e_i, e_j\rangle = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases}$$

Applying Pythagora's theorem we deduce

$$\|x_1e_1 + \cdots + x_ne_n\|^2 = x_1^2 + \cdots + x_n^2 = |\langle u, e_1\rangle|^2 + \cdots + |\langle u, e_n\rangle\,|^2. \tag{3.7}$$

Example 3.14. Consider the orthonormal basis e_1, e_2 of $\mathbb{R}^2$ constructed in Remark 3.12. If

$$u = \left[\begin{array}{c} 2 \\ 1 \end{array}\right],$$

then the coordinates x_1, x_2 of u in this basis are given by

$$x_1 = \langle u, e_1\rangle = \frac{6}{5} + \frac{4}{5} = 2,$$

$$x_2 = \langle u, e_2\rangle = \frac{8}{5} - \frac{3}{5} = 1,$$

so that

$$u = 2e_1 + e_2. \qquad \square$$

Corollary 3.15. *Any orthonormal family*

$$\{\, \boldsymbol{e}_1, \dots, \boldsymbol{e}_k \,\}$$

in a finite dimensional Euclidean vector can be extended to an orthonormal basis *of that space.*

Proof. According to Proposition 3.37, the family

$$\{\, \boldsymbol{e}_1, \dots, \boldsymbol{e}_k \,\}$$

is linearly independent. Therefore, we can extend it to a basis of $\boldsymbol{U}$,

$$\{\, \boldsymbol{e}_1, \dots, \boldsymbol{e}_k, \boldsymbol{u}_{k+1}, \dots, \boldsymbol{u}_n \,\}.$$

If we apply the Gramm-Schmidt procedure to the above linearly independent family we obtaine an orthonormal basis that extends our original orthonormal family.[2] □

3.4. **Orthogonal projections.** Suppose that $(\boldsymbol{U}, \langle -, - \rangle)$ is a finite dimensional Euclidean $\mathbb{F}$-vector space. If X is a subset of $\boldsymbol{U}$ then we set

$$X^{\perp} = \big\{ \boldsymbol{u} \in \boldsymbol{U};\ \ \langle \boldsymbol{u}, \boldsymbol{x} \rangle = 0,\ \ \forall \boldsymbol{x} \in X \,\big\}.$$

In other words, $X^{\perp}$ consists of the vectors orthogonal to *all* the vectors in X. For this reason we will obten write $\boldsymbol{u} \perp X$ to indicate that $\boldsymbol{u} \in X^{\perp}$. The following result is left to the reader.

Proposition 3.16. *(a) The subset $X^{\perp}$ is a vector subspace of $\boldsymbol{U}$.*
(b)

$$X \subset Y \Rightarrow X^{\perp} \supset Y^{\perp}.$$

(c)

$$X^{\perp} = \big(\operatorname{span}(X)\,\big)^{\perp}.$$

□

Theorem 3.17. *If $\boldsymbol{V}$ is a subspace of $\boldsymbol{U}$, then*

$$\boldsymbol{U} = \boldsymbol{V} \oplus \boldsymbol{V}^{\perp}.$$

Proof. We need to check two things.

$$\boldsymbol{V} \cap \boldsymbol{V}^{\perp} = 0, \tag{3.8a}$$

$$\boldsymbol{U} = \boldsymbol{V} + \boldsymbol{V}^{\perp}. \tag{3.8b}$$

Proof of (3.8a). If $\boldsymbol{x} \in \boldsymbol{V} \cap \boldsymbol{V}^{\perp}$ then

$$0 = \langle \boldsymbol{x}, \boldsymbol{x} \rangle$$

which implies that $\boldsymbol{x} = 0$.
Proof of (3.8b). Fix an orthonormal basis of $\boldsymbol{V}$,

$$\mathcal{B} := \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k\}.$$

Extend it to an orthonormal basis of $\boldsymbol{U}$,

$$\boldsymbol{e}_1, \dots, \boldsymbol{e}_k, \boldsymbol{e}_{k+1}, \dots, \boldsymbol{e}_n.$$

By construction,

$$\boldsymbol{e}_{k+1}, \dots, \boldsymbol{e}_n \in \mathcal{B}^{\perp} = \big(\operatorname{span}(\mathcal{B})\,\big)^{\perp} = \boldsymbol{V}^{\perp}$$

so that

$$\operatorname{span}\{\boldsymbol{e}_{k+1}, \dots, \boldsymbol{e}_n\} \subset \boldsymbol{V}^{\perp}.$$

[2] Exercise 3.5 asks you to verify this claim.

Clearly any vector $\boldsymbol{u} \in \boldsymbol{U}$ can be written as a sum of two vectors

$$\boldsymbol{u} = \boldsymbol{v} + \boldsymbol{w}, \;\; \boldsymbol{v} \in \operatorname{span}(\mathcal{B}) = \boldsymbol{V}, \;\; \boldsymbol{w} \in \operatorname{span}\{\boldsymbol{e}_{k+1}, \dots, \boldsymbol{e}_n\} \subset \boldsymbol{V}^{\perp}.$$

□

Corollary 3.18. *If $\boldsymbol{V}$ is a subsace of $\boldsymbol{U}$, then*

$$\boldsymbol{V} = (\boldsymbol{V}^{\perp})^{\perp}.$$

Proof. Theorem 3.17 implies that for any subspace $\boldsymbol{W}$ of $\boldsymbol{U}$. We have

$$\dim \boldsymbol{W}^{\perp} = \dim \boldsymbol{U} - \dim \boldsymbol{W}.$$

If we let $\boldsymbol{W} = \boldsymbol{V}^{\perp}$ we deduce that

$$\dim(\boldsymbol{V}^{\perp})^{\perp} = \dim \boldsymbol{U} - \dim \boldsymbol{V}^{\perp}.$$

If we let $\boldsymbol{W} = \boldsymbol{V}$ we deduce

$$\dim \boldsymbol{V}^{\perp} = \dim \boldsymbol{U} - \dim \boldsymbol{V}.$$

Hence

$$\dim \boldsymbol{V} = \dim(\boldsymbol{V}^{\perp})^{\perp}$$

so it suffices to show that

$$\boldsymbol{V} \subset (\boldsymbol{V}^{\perp})^{\perp},$$

i.e., we have to show that any vector $\boldsymbol{v}$ in $\boldsymbol{V}$ is orthogonal to any vector $\boldsymbol{w}$ in $\boldsymbol{V}^{\perp}$. Since $\boldsymbol{w} \in \boldsymbol{V}^{\perp}$ we have $\boldsymbol{w} \perp \boldsymbol{v}$ so that $\boldsymbol{v} \perp \boldsymbol{w}$. □

Suppose that $\boldsymbol{V}$ is a subspace of $\boldsymbol{U}$. Then any $\boldsymbol{u} \in \boldsymbol{U}$ admits a *unique* decomposition

$$\boldsymbol{u} = \boldsymbol{v} + \boldsymbol{w}, \;\; \boldsymbol{v} \in \boldsymbol{V}, \;\; \boldsymbol{w} \in \boldsymbol{V}^{\perp}.$$

We set

$$P_{\boldsymbol{V}} \boldsymbol{u} := \boldsymbol{v}.$$

Observe that if

$$\boldsymbol{u}_0 = \boldsymbol{v}_0 + \boldsymbol{w}_0, \;\; \boldsymbol{u}_1 = \boldsymbol{v}_1 + \boldsymbol{w}_2, \;\; \boldsymbol{v}_0, \boldsymbol{v}_1 \in \boldsymbol{V}, \;\; \boldsymbol{w}_0, \boldsymbol{w}_1 \in \boldsymbol{V}^{\perp},$$

then

$$(\boldsymbol{u}_0 + \boldsymbol{u}_1) = \underbrace{(\boldsymbol{v}_0 + \boldsymbol{v}_1)}_{\in \boldsymbol{V}} + \underbrace{(\boldsymbol{w}_0 + \boldsymbol{w}_1)}_{\in \boldsymbol{V}^{\perp}}$$

and we deduce

$$P_{\boldsymbol{V}}(\boldsymbol{u}_0 + \boldsymbol{u}_1) = \boldsymbol{v}_0 + \boldsymbol{v}_1 = P_{\boldsymbol{V}} \boldsymbol{u}_0 + P_{\boldsymbol{V}} \boldsymbol{u}_1.$$

Similarly, if $\lambda \in \mathbb{F}$ and $\boldsymbol{u} \in \boldsymbol{U}$, then

$$\lambda u = \lambda \boldsymbol{v} + \lambda \boldsymbol{w}$$

and we deduce

$$P_{\boldsymbol{V}}(\lambda \boldsymbol{u}) = \lambda \boldsymbol{v} = \lambda P_{\boldsymbol{V}} \boldsymbol{u}.$$

We have thus shown that the map

$$P_{\boldsymbol{V}} : \boldsymbol{U} \to \boldsymbol{U}, \;\; \boldsymbol{u} \mapsto P_{\boldsymbol{V}} \boldsymbol{u}$$

is a linear operator. It is called the *the orthogonal projection* onto the subpace $\boldsymbol{V}$. Observe that

$$\boldsymbol{R}(P_{\boldsymbol{V}}) = \boldsymbol{V}, \;\; \ker P_{\boldsymbol{V}} = \boldsymbol{V}^{\perp}. \tag{3.9}$$

Note that $P_{\boldsymbol{V}} \boldsymbol{u}$ is the *unique* vector $\boldsymbol{v}$ in $\boldsymbol{V}$ with the property that $(\boldsymbol{u} - \boldsymbol{v}) \perp \boldsymbol{V}$.

Proposition 3.19. *Suppose $\boldsymbol{V}$ is a subspace of $\boldsymbol{U}$ and $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k$ is an* orthonormal *basis of $\boldsymbol{V}$. Then*

$$P_{\boldsymbol{V}}\boldsymbol{u} = \langle \boldsymbol{u}, \boldsymbol{e}_1\rangle \boldsymbol{e}_1 + \cdots + \langle \boldsymbol{u}, \boldsymbol{e}_k\rangle \boldsymbol{e}_k, \quad \forall \boldsymbol{u} \in \boldsymbol{U}.$$

Proof. It suffices to show that the vector

$$\boldsymbol{w} = \boldsymbol{u} - \big(\, \langle \boldsymbol{u}, \boldsymbol{e}_1\rangle \boldsymbol{e}_1 + \cdots + \langle \boldsymbol{u}, \boldsymbol{e}_k\rangle \boldsymbol{e}_k \,\big)$$

is orthogonal to all the vectors $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k$ because then it will be orthogonal to any linear combination of these vectors. We have

$$\langle \boldsymbol{w}, \boldsymbol{e}_j\rangle = \langle \boldsymbol{u}, \boldsymbol{e}_j\rangle - \big(\langle \boldsymbol{u}, \langle \boldsymbol{e}_1, \rangle\langle \boldsymbol{e}_1, \boldsymbol{e}_j\rangle + \cdots + \langle \boldsymbol{u}, \boldsymbol{e}_k\rangle\langle \boldsymbol{e}_k, \boldsymbol{e}_j\rangle \,\big).$$

Since $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k$ is an *orthonormal* basis of $\boldsymbol{V}$ we deduce that

$$\langle \boldsymbol{e}_i, \boldsymbol{e}_j\rangle = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases}$$

Hence

$$\big(\langle \boldsymbol{u}, \langle \boldsymbol{e}_1, \rangle\langle \boldsymbol{e}_1, \boldsymbol{e}_j\rangle + \cdots + \langle \boldsymbol{u}, \boldsymbol{e}_k\rangle\langle \boldsymbol{e}_k, \boldsymbol{e}_j\rangle \,\big) = \langle \boldsymbol{u}, \boldsymbol{e}_j\rangle\langle \boldsymbol{e}_j \boldsymbol{e}_j\rangle = \langle \boldsymbol{u}, \boldsymbol{e}_j\rangle.$$

□

Theorem 3.20. *Let $\boldsymbol{V}$ be a subspace of $\boldsymbol{U}$. Fix $\boldsymbol{u}_0 \in \boldsymbol{U}$. Then*

$$\|\boldsymbol{u}_0 - P_{\boldsymbol{V}}\boldsymbol{u}_0\| \leq \|\boldsymbol{u}_0 - \boldsymbol{v}\|, \quad \forall \boldsymbol{v} \in \boldsymbol{V},$$

and we have equality if and only if $\boldsymbol{v} = \boldsymbol{v}_0$. In other words, $P_{\boldsymbol{V}}\boldsymbol{u}_0$ is <u>the</u> *vector in $\boldsymbol{V}$ closest to $\boldsymbol{u}_0$.*

Proof. Set $\boldsymbol{v}_0 := P_{\boldsymbol{V}}\boldsymbol{u}_0$, $\boldsymbol{w}_0 := \boldsymbol{u}_0 - P_{\boldsymbol{V}}\boldsymbol{u}_0 \in \boldsymbol{V}^{\perp}$. Then for any $\boldsymbol{v} \in \boldsymbol{V}$ we have

$$\boldsymbol{u} - \boldsymbol{v} = (\boldsymbol{v}_0 - \boldsymbol{v}) + \boldsymbol{w}_0.$$

Since $\boldsymbol{v}_0 \perp (\boldsymbol{u}_0 - \boldsymbol{v})$ we deduce from Pythagora's Theorem that

$$\|\boldsymbol{u} - \boldsymbol{v}\|^2 = \|\boldsymbol{v}_0 - \boldsymbol{v}\|^2 + \|\boldsymbol{w}_0\|^2 \geq \|\boldsymbol{w}_0\|^2 = \|\boldsymbol{u}_0 - P_{\boldsymbol{V}}\boldsymbol{u}_0\|^2.$$

Hence

$$\|\boldsymbol{u}_0 - P_{\boldsymbol{V}}\boldsymbol{u}_0\| \leq \|\boldsymbol{u}_0 - \boldsymbol{v}\|,$$

and we have equality if and only if $\boldsymbol{v} = \boldsymbol{v}_0 = P_{\boldsymbol{V}}\boldsymbol{u}_0$. □

Proposition 3.21. *Let $\boldsymbol{V}$ be a subspace of $\boldsymbol{U}$. Then*

$$P_{\boldsymbol{V}}^2 = P_{\boldsymbol{V}},$$

and

$$\|P_{\boldsymbol{V}}\boldsymbol{u}\| \leq \|\boldsymbol{u}\|, \quad \forall \boldsymbol{u} \in \boldsymbol{U}.$$

Proof. By construction

$$P_{\boldsymbol{V}}\boldsymbol{v} = \boldsymbol{v}, \quad \forall \boldsymbol{v} \in \boldsymbol{V}.$$

Hence

$$P_{\boldsymbol{V}}\big(\, P_{\boldsymbol{V}}\boldsymbol{u} \,\big) = P_{\boldsymbol{V}}\boldsymbol{u}, \quad \forall \boldsymbol{u} \in \boldsymbol{U}$$

because $P_{\boldsymbol{V}}\boldsymbol{u} \in \boldsymbol{V}$.

Next, observe that for any $\boldsymbol{u} \in \boldsymbol{U}$ we have $P_{\boldsymbol{V}}\boldsymbol{u} \perp (\boldsymbol{u} - P_{\boldsymbol{V}}\boldsymbol{u})$ so that

$$\|\boldsymbol{u}\|^2 = \|P_{\boldsymbol{V}}\boldsymbol{u}\|^2 + \|\boldsymbol{u} - P_{\boldsymbol{V}}\boldsymbol{u}\|^2 \geq \|P_{\boldsymbol{V}}\boldsymbol{u}\|^2.$$

□

3.5. **Linear functionals and adjoints on Euclidean spaces.** Suppose that $\boldsymbol{U}$ is a finite dimensional $\mathbb{F}$-vector space, $\boldsymbol{F} = \mathbb{R}, \mathbb{C}$. The *dual* of $\boldsymbol{U}$, is the $\mathbb{F}$-vector space of *linear functionals* on $\boldsymbol{U}$, i.e., linear maps

$$\alpha : \boldsymbol{U} \to \mathbb{F}.$$

The dual of $\boldsymbol{U}$ is denoted by $\boldsymbol{U}^*$. The vector space $\boldsymbol{U}^*$ has the same dimension as $\boldsymbol{U}$ and thus they are isomorphic. However, *there is no distinguished isomorphism between these two vector spaces!*

We want to show that if we fix an inner product on $\boldsymbol{U}$, then we can construct in a concrete way an isomorphism between these spaces. Before we proceed with this construction let us first observe that there is a natural bilinear map

$$B : \boldsymbol{U}^* \times \boldsymbol{U} \to \mathbb{F}, \;\; B(\alpha, u) = \alpha(\boldsymbol{u}).$$

Theorem 3.22 (Riesz representation theorem: the real case)**.** *Suppose that $\boldsymbol{U}$ is a finite dimensional real vector space equipped with an inner product*

$$\langle -, - \rangle : \boldsymbol{U} \times \boldsymbol{U} \to \mathbb{R}.$$

To any $\boldsymbol{u} \in \boldsymbol{U}$ we associate the linear functional $\boldsymbol{u}^* \in \boldsymbol{U}^*$ defined by the equality

$$B(\boldsymbol{u}^*, \boldsymbol{x}) = \boldsymbol{u}^*(\boldsymbol{x}) := \langle \boldsymbol{x}, \boldsymbol{u} \rangle, \;\; \forall \boldsymbol{x} \in \boldsymbol{U}.$$

Then the map

$$\boldsymbol{U} \ni \boldsymbol{U} \mapsto \boldsymbol{u}^* \in \boldsymbol{U}^*$$

is a linear isomorphism.

Proof. Let us show that the map $\boldsymbol{u} \mapsto \boldsymbol{u}^*$ is linear.

For any $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{x} \in \boldsymbol{U}$ we have

$$\begin{aligned}(\boldsymbol{u} + \boldsymbol{v})^*(\boldsymbol{x}) &= \langle \boldsymbol{x}, \boldsymbol{u} + \boldsymbol{v} \rangle = \langle \boldsymbol{x}, \boldsymbol{u}^* \rangle + \langle \boldsymbol{x}, \boldsymbol{v}^* \rangle \\ &= \boldsymbol{u}^*(\boldsymbol{x}) + \boldsymbol{v}^*(\boldsymbol{x}) = (\boldsymbol{u}^* + \boldsymbol{v}^*)(\boldsymbol{x}),\end{aligned}$$

which show that

$$(\boldsymbol{u} + \boldsymbol{v})^* = \boldsymbol{u}^* + \boldsymbol{v}^*.$$

For any $\boldsymbol{u}, \boldsymbol{x} \in \boldsymbol{U}$ and any $t \in \mathbb{R}$ we have

$$(t\boldsymbol{u})^*(\boldsymbol{x}) = \langle \boldsymbol{x}, t\boldsymbol{u} \rangle = t\langle \boldsymbol{x}, \boldsymbol{u} \rangle = t\boldsymbol{u}^*(\boldsymbol{x}),$$

i.e., $(t\boldsymbol{u})^* = t\boldsymbol{u}^*$. This proved the claimed linearity. To prove that it is an isomorphism, we need to show that it is both injective and surjective.

Injectivity. Suppose that $\boldsymbol{u} \in \boldsymbol{U}$ is such that $\boldsymbol{u}^* = 0$. This means that $\boldsymbol{u}^*(\boldsymbol{x}) = 0$, $\forall \boldsymbol{x} \in \boldsymbol{U}$. If we let $\boldsymbol{x} = \boldsymbol{u}$ we deduce

$$0 = \boldsymbol{u}^*(\boldsymbol{u}) = \langle \boldsymbol{u}, \boldsymbol{u} \rangle = \|\boldsymbol{u}\|^2,$$

so that $\boldsymbol{u} = 0$.

Surjectivity. We have to show that that for any $\alpha \in \boldsymbol{U}^*$ there exists $\boldsymbol{u} \in \boldsymbol{U}$ such that $\alpha = \boldsymbol{u}^*$.

Let $n = \dim \boldsymbol{U}$. Fix an *orthonormal basis* $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ of $\boldsymbol{U}$. The linear functional α is uniquely determined by its values on $\boldsymbol{e}_i$,

$$\alpha_i = \alpha(\boldsymbol{e}_i).$$

Define

$$\boldsymbol{u} := \sum_{k=1}^{n} \alpha_k \boldsymbol{e}_k.$$

Then

$$\boldsymbol{u}^*(\boldsymbol{e}_i) = \langle \boldsymbol{e}_i, \boldsymbol{u} \rangle = \langle \boldsymbol{e}_i, \alpha_1 \boldsymbol{e}_1 + \cdots + \alpha_i \boldsymbol{e}_i + \cdots + \alpha_n \boldsymbol{e}_n \rangle$$

$$= \langle \boldsymbol{e}_i, \alpha_1 \boldsymbol{e}_1 \rangle + \cdots + \langle \boldsymbol{e}_i, \alpha_i \boldsymbol{e}_i \rangle + \cdots + \langle \boldsymbol{e}_i, \alpha_n \boldsymbol{e}_n \rangle = \alpha_i.$$

Thus

$$u^*(\boldsymbol{e}_i) = \alpha(\boldsymbol{e}_i), \ \ \forall i = 1, \dots, n,$$

so that $\alpha = \boldsymbol{u}^*$. □

Theorem 3.23 (Riesz representation theorem: the complex case)**.** *Suppose that $\boldsymbol{U}$ is a finite dimensional complex vector space equipped with an inner product*

$$\langle -, - \rangle : \boldsymbol{U} \times \boldsymbol{U} \to \mathbb{C}.$$

To any $\boldsymbol{u} \in \boldsymbol{U}$ we associate the linear functional $\boldsymbol{u}^* \in \boldsymbol{U}^*$ defined by the equality

$$B(\boldsymbol{u}^*, \boldsymbol{x}) = \boldsymbol{u}^*(\boldsymbol{x}) := \langle \boldsymbol{x}, \boldsymbol{u} \rangle, \ \ \forall \boldsymbol{x} \in \boldsymbol{U}.$$

Then the map

$$\boldsymbol{U} \ni \boldsymbol{U} \mapsto \boldsymbol{u}^* \in \boldsymbol{U}^*$$

is bijective and conjugate linear, *i.e., for any $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{U}$ and any $z \in \mathbb{C}$ we have*

$$(\boldsymbol{u} + \boldsymbol{v})^* = \boldsymbol{u}^* + \boldsymbol{v}^*, \ \ (z\boldsymbol{u})^* = \bar{z}\boldsymbol{u}^*.$$

Proof. Let us first show that the map $\boldsymbol{u} \mapsto \boldsymbol{u}^*$ is conjugate linear.

For any $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{x} \in \boldsymbol{U}$ we have

$$(\boldsymbol{u} + \boldsymbol{v})^*(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{u} + \boldsymbol{v} \rangle = \langle \boldsymbol{x}, \boldsymbol{u}^* \rangle + \langle \boldsymbol{x}, \boldsymbol{v}^* \rangle$$
$$= \boldsymbol{u}^*(\boldsymbol{x}) + \boldsymbol{v}^*(\boldsymbol{x}) = (\boldsymbol{u}^* + \boldsymbol{v}^*)(\boldsymbol{x}),$$

which show that

$$(\boldsymbol{u} + \boldsymbol{v})^* = \boldsymbol{u}^* + \boldsymbol{v}^*.$$

For any $\boldsymbol{u}, \boldsymbol{x} \in \boldsymbol{U}$ and any $z \in \mathbb{C}$ we have

$$(z\boldsymbol{u})^*(\boldsymbol{x}) = \langle \boldsymbol{x}, z\boldsymbol{u} \rangle = \bar{z}\langle \boldsymbol{x}, \boldsymbol{u} \rangle = \bar{z}\boldsymbol{u}^*(\boldsymbol{x}),$$

i.e., $(z\boldsymbol{u})^* = \bar{z}\boldsymbol{u}^*$. This proved the claimed conjugate linearity. We now prove the bijectivity claim.

Injectivity. Suppose that $\boldsymbol{u} \in \boldsymbol{U}$ is such that $\boldsymbol{u}^* = 0$. This means that $\boldsymbol{u}^*(\boldsymbol{x}) = 0$, $\forall \boldsymbol{x} \in \boldsymbol{U}$. If we let $\boldsymbol{x} = \boldsymbol{u}$ we deduce

$$0 = \boldsymbol{u}^*(\boldsymbol{u}) = \langle \boldsymbol{u}, \boldsymbol{u} \rangle = \|\boldsymbol{u}\|^2,$$

so that $\boldsymbol{u} = 0$.

Surjectivity. We have to show that that for any $\alpha \in \boldsymbol{U}^*$ there exists $\boldsymbol{u} \in \boldsymbol{U}$ such that $\alpha = \boldsymbol{u}^*$.

Let $n = \dim \boldsymbol{U}$. Fix an *orthonormal basis* $\mathbf{e}_1, \dots, \mathbf{e}_n$ of $\boldsymbol{U}$. The linear functional α is uniquely determined by its values on $\boldsymbol{e}_i$,

$$\alpha_i = \alpha(\boldsymbol{e}_i).$$

Define

$$\boldsymbol{u} := \sum_{k=1}^{n} \bar{\alpha}_k \boldsymbol{e}_k.$$

Then

$$\boldsymbol{u}^*(\boldsymbol{e}_i) = \langle \boldsymbol{e}_i, \boldsymbol{u} \rangle = \langle \boldsymbol{e}_i, \bar{\alpha}_1 \boldsymbol{e}_1 + \cdots + \bar{\alpha}_i \boldsymbol{e}_i + \cdots + \bar{\alpha}_n \boldsymbol{e}_n \rangle$$
$$= \langle \boldsymbol{e}_i, \bar{\alpha}_1 \boldsymbol{e}_1 \rangle + \cdots + \langle \boldsymbol{e}_i, \bar{\alpha}_i \boldsymbol{e}_i \rangle + \cdots + \langle \boldsymbol{e}_i, \bar{\alpha}_n \boldsymbol{e}_n \rangle = \boldsymbol{e}_i, \bar{\alpha}_i \boldsymbol{e}_i = \alpha_i \langle \boldsymbol{e}_i, \boldsymbol{e}_i \rangle = \alpha_i.$$

Thus

$$u^*(\boldsymbol{e}_i) = \alpha(\boldsymbol{e}_i), \ \ \forall i = 1, \dots, n,$$

so that $\alpha = \boldsymbol{u}^*$. □

Suppose that $\boldsymbol{U}$ and $\boldsymbol{V}$ are two $\mathbb{F}$-vector spaces equipped with inner products $\langle -,-\rangle_{\boldsymbol{U}}$ and respectively $\langle -,-\rangle_{\boldsymbol{V}}$. Next, assume that $T:\boldsymbol{U}\to\boldsymbol{V}$ is a linear map.

Theorem 3.24. *There exists a unique linear map $S:\boldsymbol{V}\to\boldsymbol{U}$ satisfying the equality*

$$\langle T\boldsymbol{u},\boldsymbol{v}\rangle_{\boldsymbol{V}}=\langle \boldsymbol{u},S\boldsymbol{v}\rangle_{\boldsymbol{U}},\ \ \forall \boldsymbol{u}\in\boldsymbol{U},\boldsymbol{v}\in\boldsymbol{V}. \tag{3.10}$$

This map is called the adjoint *of T with respect to the inner products $\langle -,-\rangle_{\boldsymbol{U}}$, $\langle -,-\rangle_{\boldsymbol{V}}$ and it is denoted by T^*. The equality (3.10) can then be rewritten*

$$\langle T\boldsymbol{u},\boldsymbol{v}\rangle_{\boldsymbol{V}}=\langle \boldsymbol{u},T^*\boldsymbol{v}\rangle_{\boldsymbol{U}},\ \ \langle T^*\boldsymbol{v},\boldsymbol{u}\rangle_{\boldsymbol{U}}=\langle \boldsymbol{v},T\boldsymbol{u}\rangle_{\boldsymbol{U}},\ \ \forall \boldsymbol{u}\in\boldsymbol{U},\ \ \boldsymbol{v}\in\boldsymbol{V}. \tag{3.11}$$

Proof. *Uniqueness.* Suppose there are two linear maps $S_1,S_2:\boldsymbol{V}\to\boldsymbol{U}$ satisfying (3.10). Thus

$$0=\langle \boldsymbol{u},S_1\boldsymbol{v}\rangle_{\boldsymbol{U}}-\langle \boldsymbol{u},S_2\boldsymbol{v}\rangle_{\boldsymbol{U}}=\langle \boldsymbol{u},(S_1-S_2)\boldsymbol{v}\rangle_{\boldsymbol{U}},\ \ \forall \boldsymbol{u}\in\boldsymbol{U},\boldsymbol{v}\in\boldsymbol{V}.$$

For fixed $\boldsymbol{v}\in\boldsymbol{V}$ we let $\boldsymbol{u}=(S_1-S_2)\boldsymbol{v}$ and we deduce from the above equality

$$0=\langle\,(S_1-S_2)\boldsymbol{v},(S_1-S_2)\boldsymbol{v}\,\rangle_{\boldsymbol{U}}=\|(S_1-S_2)\boldsymbol{v}\|^2_{\boldsymbol{U}},$$

so that $(S_1-S_2)\boldsymbol{v}=0$, for any $\boldsymbol{v}$ in $\boldsymbol{V}$. This shows that $S_1=S_2$, thus proving the uniqueness part of the theorem.

Existence. Any $\boldsymbol{v}\in\boldsymbol{V}$ defines a linear functional

$$L_{\boldsymbol{v}}:\boldsymbol{U}\to\mathbb{F},\ \ L_{\boldsymbol{v}}(\boldsymbol{u})=\langle T\boldsymbol{u},\boldsymbol{v}\rangle_{\boldsymbol{V}}.$$

Thus there exists a unique vector $S\boldsymbol{v}\in\boldsymbol{U}$ such that

$$L_{\boldsymbol{v}}=(S\boldsymbol{v})^*\Longleftrightarrow\langle T\boldsymbol{u},\boldsymbol{v}\rangle_{\boldsymbol{V}}=\langle \boldsymbol{u},S\boldsymbol{v}\rangle_{\boldsymbol{U}},\ \ \forall \boldsymbol{u}\in\boldsymbol{U}.$$

One can verify easily that the correspondence $\boldsymbol{V}\ni\boldsymbol{v}\mapsto S\boldsymbol{v}\in\boldsymbol{U}$ described above is a linear map; see Exercise 3.13. □

Example 3.25. Let $T:\boldsymbol{U}\to\boldsymbol{V}$ be as in the statement of Theorem 3.24. Assume $m=\dim_{\mathbb{F}}\boldsymbol{U}$, $n=\dim_{\mathbb{F}}\boldsymbol{V}$. Fix an *orthonormal* basis

$$\underline{\boldsymbol{e}}:=\{\boldsymbol{e}_1,\dots,\boldsymbol{e}_m\}$$

of $\boldsymbol{U}$ and an *orthonormla* basis

$$\underline{\boldsymbol{f}}:=\{\boldsymbol{f}_1,\dots,\boldsymbol{f}_n\}$$

of $\boldsymbol{V}$. With respect to these bases the operator T is represented by an $n\times m$ matrix

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m}\\ a_{21} & a_{22} & \cdots & a_{2m}\\ \vdots & \vdots & \vdots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nm}\end{bmatrix},$$

while the adjoint operator $T^*:\boldsymbol{V}\to\boldsymbol{U}$ is represented by an $m\times n$ matrix

$$A^*=\begin{bmatrix} a^*_{11} & a^*_{12} & \cdots & a^*_{1n}\\ a^*_{21} & a^*_{22} & \cdots & a^*_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ a^*_{m1} & a^*_{m2} & \cdots & a^*_{mn}\end{bmatrix}.$$

The j-th column of AS describes the coordinates of the vector $T\boldsymbol{e}_j$ in the basis $\underline{\boldsymbol{f}}$,

$$T\boldsymbol{e}_j=a_{1j}\boldsymbol{f}_1+\cdots a_{nj}\boldsymbol{f}_n.$$

We deduce that

$$\langle T\boldsymbol{e}_j, \boldsymbol{f}_i\rangle_{\boldsymbol{V}} = a_{ij}.$$

Hence

$$\langle \boldsymbol{f}_i, T\boldsymbol{e}_j\rangle_{\boldsymbol{V}} = \overline{\langle T\boldsymbol{e}_j, \boldsymbol{f}_i\rangle_{\boldsymbol{V}}} = \bar{a}_{ij}.$$

On the other hand, the i-th column of A^* describes the coordinates of $T^*\boldsymbol{f}_i$ in the basis $\underline{\boldsymbol{e}}$ so that

$$T^*\boldsymbol{f}_i = a^*_{1i}\boldsymbol{e}_1 + \cdots + a^*_{mi}\boldsymbol{e}_m$$

and we deduce that

$$\langle T^*\boldsymbol{f}_i, \boldsymbol{e}_j\rangle = a^*_{ji}.$$

On the other hand, we have

$$\bar{a}_{ij} = \langle \boldsymbol{f}_i, T\boldsymbol{e}_j\rangle_{\boldsymbol{V}} \stackrel{(3.11)}{=} \langle T^*\boldsymbol{f}_i, \boldsymbol{e}_j\rangle = a^*_{ji}.$$

Thus, A^* is the *conjugate transpose* of A. In other words, *the entries of A^* are the conjugates of the corresponding entries of the transpose of A*,

$$A^* = \overline{A^\dagger}. \qquad \square$$

Definition 3.26. Let $\boldsymbol{U}$ be a finite dimensional Euclidean $\mathbb{F}$-space with inner product $\langle -, -\rangle$. A linear operator $T : \boldsymbol{U} \to \boldsymbol{U}$ is called *selfadjoint* or *symmetric* if $T = T^*$, i.e.,

$$\langle T\boldsymbol{u}_1, \boldsymbol{u}_2\rangle = \langle \boldsymbol{u}_1, T\boldsymbol{u}_2, \;\; \forall \boldsymbol{u}_1, \boldsymbol{u}_2 \in \boldsymbol{U}. \qquad \square$$

Example 3.27. (a) Consider the standard real Euclidean n-dimensional space $\mathbb{R}^n$. Any real $n \times n$ matrix A can be identified with a linear operator $T_A : \mathbb{R}^n \to \mathbb{R}^n$. The operator T_A is selfadjoint if and only if the matrix A is *symmetric*, i.e.,

$$a_{ij} = a_{ji}, \;\; \forall i, j$$

or, equivalently $A = A^\dagger =$ the transpose of A.

(b) Consider the standard complex Euclidean n-dimensional space $\mathbb{C}^n$. Any complex $n \times n$ matrix A can be identified with a line operator $T_A : \mathbb{C}^n \to \mathbb{C}^n$. The operator T_A is selfadjoint if and only if the matrix A is *Hermitian*, i.e.,

$$a_{ij} = \bar{a}_{ji}, \;\; \forall i, j$$

or, equivalently $A = A^* =$ the conjugate transpose of A.

(c) Suppose that $\boldsymbol{V}$ is a subspace of the finite dimensional Euclidean space $\boldsymbol{U}$. Then the orthogonal projection $P_{\boldsymbol{V}} : \boldsymbol{U} \to \boldsymbol{U}$ is a selfadjoint operator, i.e.,

$$\langle P_{\boldsymbol{V}}\boldsymbol{u}_1, \boldsymbol{u}_2\rangle = \langle \boldsymbol{u}_1, P_{\boldsymbol{V}}\boldsymbol{u}_2\rangle, \;\; \forall \boldsymbol{u}_1, \boldsymbol{u}_2 \in \boldsymbol{U}.$$

Indeed, let $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \boldsymbol{U}$. They decompose uniquely as

$$\boldsymbol{u}_1 = \boldsymbol{v}_1 + \boldsymbol{w}_1, \;\; \boldsymbol{u}_2 = \boldsymbol{v}_2 + \boldsymbol{w}_2, \;\; \boldsymbol{v}_1, \boldsymbol{v}_2 \in \boldsymbol{V}, \;\; \boldsymbol{w}_1, \boldsymbol{w}_2 \in \boldsymbol{V}^\perp.$$

Then $P_{\boldsymbol{V}}\boldsymbol{u}_1 = \boldsymbol{v}_1$ so that

$$\langle P_{\boldsymbol{V}}\boldsymbol{u}_1, \boldsymbol{u}_2\rangle = \langle \boldsymbol{v}_1, \boldsymbol{v}_2 + \boldsymbol{w}_2\rangle = \langle \boldsymbol{v}_1, \boldsymbol{v}_2\rangle + \underbrace{\langle \boldsymbol{v}_1, \boldsymbol{w}_2\rangle}_{\boldsymbol{w}_2 \perp \boldsymbol{v}_1} = \langle \boldsymbol{v}_1, \boldsymbol{v}_2\rangle.$$

Similarly, $P_{\boldsymbol{V}}\boldsymbol{u}_2 = \boldsymbol{v}_2$ and we deduce

$$\langle \boldsymbol{u}_1, P_{\boldsymbol{V}}\boldsymbol{u}_2\rangle = \langle \boldsymbol{v}_1 + \boldsymbol{w}_1, \boldsymbol{v}_2\rangle = \langle \boldsymbol{v}_1, \boldsymbol{v}_2\rangle + \underbrace{\langle \boldsymbol{w}_1, \boldsymbol{v}_2\rangle}_{\boldsymbol{w}_1 \perp \boldsymbol{v}_2} = \langle \boldsymbol{v}_1, \boldsymbol{v}_2\rangle. \qquad \square$$

Proposition 3.28. *Suppose that $\boldsymbol{U}$ is a finite dimensional Euclidean $\mathbb{F}$-space and $T:\boldsymbol{U}\to\boldsymbol{U}$ is a selfadjoint operator. Then*

$$\operatorname{spec} T\subset\mathbb{R}.$$

In other words, the eigenvalues of a selfadjoint operator are real.

Proof. Let λ be an eigenvalue T and $\boldsymbol{u}\neq 0$ an eigenvector of T corresponding to the eigenvalue λ. We have

$$T\boldsymbol{u}=\lambda\boldsymbol{u}$$

so that

$$\lambda\|\boldsymbol{u}\|^2=\langle\lambda\boldsymbol{u},\boldsymbol{u}\rangle=\langle T\boldsymbol{u},\boldsymbol{u}\rangle\Rightarrow\lambda=\frac{1}{\|\boldsymbol{u}\|^2}\langle T\boldsymbol{u},\boldsymbol{u}\rangle.$$

On the other hand

$$\langle T\boldsymbol{u},\boldsymbol{u}\rangle=\overline{\langle\boldsymbol{u},T\boldsymbol{u}\rangle}.$$

Since T is selfadjoint we deduce

$$\langle\boldsymbol{u},T\boldsymbol{u}\rangle=\langle T\boldsymbol{u},\boldsymbol{u}\rangle$$

so that

$$\langle T\boldsymbol{u},\boldsymbol{u}\rangle=\overline{\langle T\boldsymbol{u},\boldsymbol{u}\rangle}.$$

Hence the inner product $\langle T\boldsymbol{u},\boldsymbol{u}\rangle$ is a real number. From the equality

$$\lambda=\frac{1}{\|\boldsymbol{u}\|^2}\langle T\boldsymbol{u},\boldsymbol{u}\rangle$$

we deduce that λ is a real number as well. □

Corollary 3.29. *If A is an $n\times n$ complex matrix such that $A=A^*$, then all the roots of the characteristic polynomial $P_A(\lambda)=\det(\lambda\mathbb{1}-A)$ are real.*

Proof. The roots of $P_A(\lambda)$ are the eigenvalues of the linear operator $T_A:\mathbb{C}^n\to\mathbb{C}^n$ defined by A. Since $A=A^*$ we deduce that T_A is selfadjoint with respect to the natural inner product on $\mathbb{C}^n$ so that all its eigenvalues are real. □

Theorem 3.30. *Suppose that $\boldsymbol{U},\boldsymbol{V}$ are two finite dimensional Euclidean $\mathbb{F}$-vector spaces and $T:\boldsymbol{U}\to\boldsymbol{V}$ is a linear operator. Then the following hold.*

(a) $(T^*)^*=T$.
(b) $\ker T=\boldsymbol{R}(T^*)^\perp$.
(c) $\ker T^*=\boldsymbol{R}(T)^\perp$.
(d) $\boldsymbol{R}(T)=(\ker T^*)^\perp$.
(e) $\boldsymbol{R}(T^*)=(\ker T^\perp$.

Proof. (a) The operator $(T^*)^*$ is a linear operator $\boldsymbol{U}\to\boldsymbol{V}$. We need to prove that, for any $\boldsymbol{u}\in\boldsymbol{U}$ we have $\boldsymbol{x}:=(T^*)^*\boldsymbol{u}-T\boldsymbol{u}$=0. Let

Because $(T^*)^*$ is the adjoint of T^* we deduce from (3.11) that

$$\langle(T^*)^*\boldsymbol{u},\boldsymbol{v}\rangle_{\boldsymbol{V}}=\langle\boldsymbol{u},T^*\boldsymbol{v}\rangle_{\boldsymbol{U}}.$$

Because T^* is the adjoint of T we deduce from (3.11) that

$$\langle\boldsymbol{u},T^*\boldsymbol{v}\rangle_{\boldsymbol{U}}=\langle T\boldsymbol{u},\boldsymbol{v}\rangle_{\boldsymbol{V}}.$$

Hence, *for any* $\boldsymbol{v}\in\boldsymbol{V}$ we have

$$0=\langle(T^*)^*\boldsymbol{u},\boldsymbol{v}\rangle_{\boldsymbol{V}}-\langle T\boldsymbol{u},\boldsymbol{v}\rangle_{\boldsymbol{V}}=\langle\boldsymbol{x},\boldsymbol{v}\rangle_{\boldsymbol{V}}.$$

By choosing $\boldsymbol{v}=\boldsymbol{x}$ we deduce $\boldsymbol{x}=0$.

(b) We need to prove that

$$\boldsymbol{u} \in \ker T \Longleftrightarrow \boldsymbol{u} \perp \boldsymbol{R}(T^*).$$

Let $\boldsymbol{u} \in \ker T$, i.e., $T\boldsymbol{u} = 0$. To prove that $\boldsymbol{u} \perp \boldsymbol{R}(T^*)$ we need to show that $\boldsymbol{u} \perp T^*\boldsymbol{v}$, $\forall \boldsymbol{v} \in \boldsymbol{V}$. For $\boldsymbol{v} \in \boldsymbol{V}$ we have

$$\langle \boldsymbol{u}, T^*\boldsymbol{v}\rangle \stackrel{(3.11)}{=} \langle T\boldsymbol{u}, \boldsymbol{v}\rangle = 0,$$

so that $\boldsymbol{u} \perp T^*\boldsymbol{v}$ for any $\boldsymbol{v} \in \boldsymbol{V}$.

Conversely, let us assume that $u \perp T^*\boldsymbol{v}$ for any $\boldsymbol{v} \in \boldsymbol{V}$. We have to show that $\boldsymbol{x} = Tu = 0$. Observe that $\boldsymbol{x} \in \boldsymbol{V}$ so that $u \perp T^*\boldsymbol{x}$. We deduce

$$0 = \langle \boldsymbol{u}, T^*x\rangle = \langle T\boldsymbol{u}, \boldsymbol{x}\langle = \langle \boldsymbol{x}, \boldsymbol{x}\rangle = \|\boldsymbol{x}\|^2.$$

Hence $\boldsymbol{x} = 0$.
(c) Set $S := T^*$ from (b) we deduce

$$\ker T^* = \ker S = \boldsymbol{R}(S^*)^{\perp}.$$

From (a) we deduce $S^* = (T^*)^* = T$ and (c) is now obvious.

Part (d) follows from (c) and Corollary 3.18, while (e) follows from (b) and Corollary 3.18. □

Corollary 3.31. *Suppose that $\boldsymbol{U}$ is a finite dimensional Euclidean vector space, and $T : \boldsymbol{U} \to \boldsymbol{U}$ is a selfadjoint operator. Then*

$$\ker T = \boldsymbol{R}(T)^{\perp}, \quad \boldsymbol{R}(T) = (\ker T)^{\perp}, \quad \boldsymbol{U} = \ker T \oplus \boldsymbol{R}(T).$$ □

Proposition 3.32. *(a) Suppose that $\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{W}$ are finite dimensional Euclidean $\mathbb{F}$-spaces. If*

$$T : \boldsymbol{U} \to \boldsymbol{V}, \quad S : \boldsymbol{V} \to \boldsymbol{W}$$

are linear operators, then

$$(ST)^* = T^*S^*.$$

(b) Suppose that $T : \boldsymbol{U} \to \boldsymbol{V}$ is a linear operator between two finite dimensional Euclidean $\mathbb{F}$-spaces. Then T is invertible if and only if the adjoint $T^* : \boldsymbol{V} \to \boldsymbol{U}$ is invertible. Moreover, if T is invertible, then

$$(T^{-1})^* = (T^*)^{-1}.$$

(c)Suppose that $S, T : \boldsymbol{U} \to \boldsymbol{V}$ are linear operators between two finite dimensional Euclidean $\mathbb{F}$-spaces. Then

$$(S+T)^* = S^* + T^*, \quad (zS)^* = \bar{z}S^*, \quad \forall z \in \mathbb{F}.$$ □

The proof is left to you as Exercise 3.15.

Proposition 3.33. *Suppose that $\boldsymbol{U}$ is a finite dimensional Euclidean $\mathbb{F}$-space and $S, T : \boldsymbol{U} \to \boldsymbol{U}$ are two selfadjoint operators. Then*

$$S = T \Longleftrightarrow \langle S\boldsymbol{u}, \boldsymbol{u}\rangle = \langle T\boldsymbol{u}, \boldsymbol{u}\rangle, \quad \forall \boldsymbol{u} \in \boldsymbol{U}.$$

Proof. The implication " $\Rightarrow$ " is obvious so it suffices to prove that

$$\langle S\boldsymbol{u},\boldsymbol{u}\rangle = \langle T\boldsymbol{u},\boldsymbol{u}\rangle,\ \ \forall \boldsymbol{u}\in \boldsymbol{U} \Rightarrow S = T.$$

We set $A := S - T$. Then A is a selfadjoint operator and it suffices to show that

$$\langle A\boldsymbol{u},\boldsymbol{u}\rangle = 0,\ \ \forall \boldsymbol{u}\in \boldsymbol{U} \Rightarrow A = 0.$$

We distinguish two cases.

(a) $\mathbb{F}=\mathbb{R}$. For any $\boldsymbol{u},\boldsymbol{v}\in\boldsymbol{U}$ we have

$$\begin{aligned}0 = \langle A(\boldsymbol{u}+\boldsymbol{v}),\boldsymbol{u}+\boldsymbol{v}\rangle &= \langle A\boldsymbol{u},\boldsymbol{u}+\boldsymbol{v}\rangle + \langle A\boldsymbol{v},\boldsymbol{u}+_{\boldsymbol{v}}\rangle\\ &= \underbrace{\langle A\boldsymbol{u},\boldsymbol{u}\rangle}_{=0} + \langle A\boldsymbol{u},\boldsymbol{v}\rangle + \langle A\boldsymbol{v},\boldsymbol{u}\rangle + \underbrace{\langle A\boldsymbol{v},\boldsymbol{v}\rangle}_{=0}\\ &= \langle A\boldsymbol{u},\boldsymbol{v}\rangle + \langle \boldsymbol{v}, A\boldsymbol{u}\rangle = 2\langle A\boldsymbol{u},\boldsymbol{v}\rangle.\end{aligned}$$

Hence

$$\langle A\boldsymbol{u},\boldsymbol{v}\rangle = 0,\ \ \forall \boldsymbol{u},\boldsymbol{v}\in\boldsymbol{U}.$$

If in the above equality we let $\boldsymbol{v} = A\boldsymbol{u}$ we deduce

$$\|A\boldsymbol{u}\|^2 = 0,\ \ \forall \boldsymbol{u}\in\boldsymbol{U},$$

i.e., $A = 0$.

(b) $\mathbb{F}=\mathbb{C}$. For any $\boldsymbol{u},\boldsymbol{v}\in\boldsymbol{U}$ we have

$$\begin{aligned}0 = \langle A(\boldsymbol{u}+\boldsymbol{v}),\boldsymbol{u}+\boldsymbol{v}\rangle &= \langle A\boldsymbol{u},\boldsymbol{u}+\boldsymbol{v}\rangle + \langle A\boldsymbol{v},\boldsymbol{u}+\boldsymbol{v}\rangle\\ &= \underbrace{\langle A\boldsymbol{u},\boldsymbol{u}\rangle}_{=0} + \langle A\boldsymbol{u},\boldsymbol{v}\rangle + \langle A\boldsymbol{v},\boldsymbol{u}\rangle + \underbrace{\langle A\boldsymbol{v},\boldsymbol{v}\rangle}_{=0}\\ = \langle A\boldsymbol{u},\boldsymbol{v}\rangle + \langle \boldsymbol{v}, A\boldsymbol{u}\rangle &= \langle A\boldsymbol{u},\boldsymbol{v}\rangle + \overline{\langle A\boldsymbol{u},\boldsymbol{v}\rangle} = 2\,\mathbf{Re}\langle A\boldsymbol{u},\boldsymbol{v}\rangle.\end{aligned}$$

Hence

$$\mathbf{Re}\langle A\boldsymbol{u},\boldsymbol{v}\rangle = 0,\ \ \forall \boldsymbol{u},\boldsymbol{v}\in\boldsymbol{U}. \tag{3.12}$$

Similarly for any $\boldsymbol{u},\boldsymbol{v}\in\boldsymbol{U}$ we have

$$\begin{aligned}0 = \langle A(\boldsymbol{u}+\boldsymbol{i}\boldsymbol{v}),\boldsymbol{u}+\boldsymbol{i}\boldsymbol{v}\rangle &= \langle A\boldsymbol{u},\boldsymbol{u}+\boldsymbol{i}\boldsymbol{v}\rangle + \boldsymbol{i}\langle A\boldsymbol{v},\boldsymbol{u}+\boldsymbol{i}\boldsymbol{v}\rangle\\ &= \underbrace{\langle A\boldsymbol{u},\boldsymbol{u}\rangle}_{=0} - \boldsymbol{i}\langle A\boldsymbol{u},\boldsymbol{v}\rangle + \boldsymbol{i}\langle A\boldsymbol{v},\boldsymbol{u}\rangle - \boldsymbol{i}^2\underbrace{\langle A\boldsymbol{v},\boldsymbol{v}\rangle}_{=0}\\ = -\boldsymbol{i}\langle A\boldsymbol{u},\boldsymbol{v}\rangle + \boldsymbol{i}\langle \boldsymbol{v}, A\boldsymbol{u}\rangle &= -\boldsymbol{i}\Big(\langle A\boldsymbol{u},\boldsymbol{v}\rangle - \overline{\langle A\boldsymbol{u},\boldsymbol{v}\rangle}\Big) = 2\,\mathbf{Im}\langle A\boldsymbol{u},\boldsymbol{v}\rangle.\end{aligned}$$

Hence

$$\mathbf{Im}\langle A\boldsymbol{u},\boldsymbol{v}\rangle = 0,\ \ \forall \boldsymbol{u},\boldsymbol{v}\in\boldsymbol{U}. \tag{3.13}$$

Putting together (3.12) and (3.13) we deduce that

$$\langle A\boldsymbol{u},\boldsymbol{v}\rangle = 0,\ \ \forall \boldsymbol{u},\boldsymbol{v}\in\boldsymbol{U}.$$

If we now let $\boldsymbol{v} = A\boldsymbol{u}$ in the above equality we deduce as in the real case that $A\boldsymbol{u} = 0$, $\forall \boldsymbol{u}\in\boldsymbol{U}$. □

Definition 3.34. Let $\boldsymbol{U},\boldsymbol{V}$ be two finite dimensional Euclidean $\mathbb{F}$-vector spaces. A linear operator $T:\boldsymbol{U}\to\boldsymbol{V}$ is called an *isometry* if for any $\boldsymbol{u}\in\boldsymbol{U}$ we have

$$\|T\boldsymbol{u}\|_{\boldsymbol{V}} = \|\boldsymbol{u}\|_{\boldsymbol{U}}.$$

□

Proposition 3.35. *A linear operator $T:\boldsymbol{U}\to\boldsymbol{V}$ between two finite dimensional Euclidean vector spaces is an isometry if and only if*

$$T^*T=\mathbb{1}_{\boldsymbol{U}}.$$ □

The proof is left to you as Exercise 3.16.

Definition 3.36. Let $\boldsymbol{U}$ be a finite dimensional Euclidean $\mathbb{F}$-space. A linear operator $T:\boldsymbol{U}\to\boldsymbol{U}$ is called an *orthogonal* operator if T is an isometry. We denote by $O(\boldsymbol{U})$ the space of orthogonal operators on $\boldsymbol{U}$. □

Proposition 3.37. *Let $\boldsymbol{U}$ be a finite dimensional Euclidean $\mathbb{F}$-space. Then*

$$T\in O(\boldsymbol{U})\Longleftrightarrow T^*T=TT^*=\mathbb{1}_{\boldsymbol{U}}.$$

Proof. The implication " $\Leftarrow$ " follows from Proposition 3.35. To prove the opposite implication, assume that T is an orthogonal operator. Hence

$$\|T\boldsymbol{u}\|=\|\boldsymbol{u}\|,\ \ \forall\boldsymbol{u}\in\boldsymbol{U}.$$

This implies in particular that $\ker T=0$, so that T is invertible. If we let $\boldsymbol{u}=T^{-1}\boldsymbol{v}$ in the above equality we deduce

$$\|T^{-1}\boldsymbol{v}\|=\|\boldsymbol{v}\|,\ \ \forall\boldsymbol{v}\in\boldsymbol{U}.$$

Hence T^{-1} is also an isometry so that

$$(T^{-1})^*T^{-1}=\mathbb{1}_{\boldsymbol{U}}.$$

Using Proposition 3.32 we deduce $(T^{-1})^*=(T^*)^{-1}$. Hence

$$(T^*)^{-1}T^{-1}=\mathbb{1}_{\boldsymbol{U}}.$$

Taking the inverses of both sides of the above equality we deduce

$$\mathbb{1}_{\boldsymbol{U}}=\big((T^*)^{-1}T^{-1}\big)^{-1}=(T^{-1})^{-1}((T^*)^{-1})^{-1}=TT^*.$$

□

3.6. **Exercises.**

Exercise 3.1. Prove the claims made in Example 3.2 (a), (b), (c). □

Exercise 3.2. Let $\big(\boldsymbol{U}, \langle -,-\rangle\big)$ be an Euclidean space.
(a) Show that for any $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{U}$ we have

$$\|\boldsymbol{u}+\boldsymbol{v}\|^2 + \|\boldsymbol{u}-\boldsymbol{v}\|^2 = 2\big(\,\|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2\,\big).$$

(b) Let $\boldsymbol{u}_0, \dots, \boldsymbol{u}_n \in \boldsymbol{U}$. Prove that

$$\|\boldsymbol{u}_0 - \boldsymbol{u}_n\| \leq \|\boldsymbol{u}_0 - \boldsymbol{u}_1\| + \|\boldsymbol{u}_1 - \boldsymbol{u}_2\| + \cdots + \|\boldsymbol{u}_{n-1} - \boldsymbol{u}_n\|.$$ □

Exercise 3.3. Show that for any complex numbers $z_1, \dots, z_n$ we have

$$\big(\,|z_1| + \cdots + |z_n|\,\big)^2 \leq n\big(|z_1|^2 + \cdots + |z_n|^2\,\big)$$

and

$$\big(|z_1|^2 + \cdots + |z_n|^2\,\big) \leq \big(\,|z_1| + \cdots + |z_n|\,\big)^2.$$

Exercise 3.4. Consider the space $\mathcal{P}_3$ of polynomials with real coefficients and degrees ≤ 3 equipped with the inner product

$$\langle P, Q\rangle = \int_0^1 P(x)Q(x)dx,\ \ \forall P, Q \in \mathcal{P}_3.$$

Construct an orthonormal basis of $\mathcal{P}_3$ by using the Gramm-Schmidt procedure applied to the basis of $\mathcal{P}_3$ given by

$$E_0(x) = 1,\ \ E_1(x) = x,\ \ E_2(x) = x^2,\ \ E_3(x) = x^3.$$ □

Exercise 3.5. Fill in the missing details in the proof of Corollary 3.15. □

Exercise 3.6. Suppose that $T : \boldsymbol{U} \to \boldsymbol{U}$ is a linear operator on a finite dimensional complex Euclidean vector space. Prove that there exists an *orthonormal* basis of U, such that, in this basis T is represented by an upper triangular matrix. □

Exercise 3.7. Prove Proposition 3.16. □

Exercise 3.8. Consider the standard Euclidean space $\mathbb{R}^3$. Denote by $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ the canonical orthonormal basis of $\mathbb{R}^3$ and by $\boldsymbol{V}$ the subspace generated by the vectors

$$\boldsymbol{v}_1 = 12\boldsymbol{e}_1 + 5\boldsymbol{e}_3,\ \ \boldsymbol{v}_2 = \boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3.$$

Find the matrix representing the orthogonal projection $P_{\boldsymbol{V}} : \mathbb{R}^3 \to \mathbb{R}^3$ in the canonical basis $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$. □

Exercise 3.9. Consider the space $\mathcal{P}_3$ of polynomials with real coefficients and degrees ≤ 3. Find $P \in \mathcal{P}_3$ such that $p(0) = p'(0) = 0$ and

$$\int_0^1 \big|\,2 + 3x - p(x)\,\big|^2 dx$$

is as small as possible.
Hint Observe that the set

$$\boldsymbol{V} = \big\{p \in \mathcal{P}_3;\ \ p(0) = p'(0) = 0\,\big\}$$

is a subspace of $\mathcal{P}_3$. Then compute $P_{\boldsymbol{V}}$, the orthogonal projection onto $\boldsymbol{V}$ with respect to the inner product

$$\langle p, q\rangle = \int_0^1 p(x)q(x)dx, \;\; p, q \in \mathcal{P}_3.$$

The answer will be $P_{\boldsymbol{V}}q$, $q = 2 + 3x \in \mathcal{P}_3$. □

Exercise 3.10. Suppose that $(\boldsymbol{U}, \langle -, -\rangle)$ is a finite dimensional real Euclidean space and $P : \boldsymbol{U} \to \boldsymbol{U}$ is a linear operator such that

$$P^2 = P, \;\; \|P\boldsymbol{u}\| \leq \|\boldsymbol{u}\|, \;\; \forall \boldsymbol{u} \in \boldsymbol{U}. \tag{P}$$

Show that there exists a subspace $\boldsymbol{V} \subset \boldsymbol{U}$ such that $P = P_{\boldsymbol{V}}$ = the orthogonal projection onto $\boldsymbol{V}$.
Hint: Let $\boldsymbol{V} = \boldsymbol{R}(P) = P(\boldsymbol{U})$, $\boldsymbol{W} := \ker P$. Using (P) argue by contradiction that $\boldsymbol{V} \subset \boldsymbol{W}^\perp$ and then conclude that $P = P_{\boldsymbol{V}}$. □

Exercise 3.11. Let $\mathcal{P}_2$ denote the space of polynomials with real coefficients and degree ≤ 2. Describe the polynomial $p_0 \in \mathcal{P}_2$ uniquely determined by the equalities

$$\int_{-\pi}^{\pi} \cos x q(x)dx = \int_{-\pi}^{\pi} p_0(x)q(x)dx, \;\; \forall q \in \mathcal{P}_2.$$ □

Exercise 3.12. Let k be a positive integer, and denote by $\mathcal{P}_k$ denote the space of polynomials with real coefficients and degree $\leq k$. For $k = 2, 3, 4$, describe the polynomial $p_k \in \mathcal{P}_k$ uniquely determined by the equalities

$$q(0) = \int_{-1}^{1} p_k(x)q(x)dx, \;\; \forall q \in \mathcal{P}_k.$$ □

Exercise 3.13. Finish the proof of Theorem 3.24. (Pay special attention to the case when $\mathbb{F} = \mathbb{C}$.) □

Exercise 3.14. Let $\mathcal{P}_2$ denote the space of polynomials with real coefficients and degree ≤ 2. We equip it with the inner product

$$\langle p, q\rangle = \int_0^1 p(x)q(x)dx.$$

Consider the linear operator $T : \mathcal{P}_2 \to \mathcal{P}_2$ defined by

$$Tp = \frac{dp}{dx}, \;\; \forall p \in \mathcal{P}_2.$$

Describe the adjoint of T. □

Exercise 3.15. Prove proposition 3.32. □

Exercise 3.16. Suppose that $\boldsymbol{U}$ is a finite dimensional Euclidean $\mathbb{F}$-spaces and $T : \boldsymbol{U} \to \boldsymbol{V}$ is an orthogonal operator. Prove that $\lambda \in \operatorname{spec}(T) \Rightarrow |\lambda| = 1$. □

Exercise 3.17. Let $\boldsymbol{U}$ be a finite dimensional Euclidean $\mathbb{F}$-vector space and $T : \boldsymbol{U} \to \boldsymbol{U}$ is a linear operator. Prove that the following statements are equivalent.

(i) The operator $T : \boldsymbol{U} \to \boldsymbol{U}$ is orthogonal.
(ii) $\langle T\boldsymbol{u}, T\boldsymbol{v}\rangle = \langle \boldsymbol{u}, \boldsymbol{v}\rangle$, $\forall \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{U}$.
(iii) For any orthonormal basis $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ of $\boldsymbol{U}$, the collection $T\boldsymbol{e}_1, \dots, T\boldsymbol{e}_n$ is also an orthonormal basis of $\boldsymbol{U}$.

□

4. SPECTRAL THEORY OF NORMAL OPERATORS

4.1. **Normal operators.** Let $\boldsymbol{U}$ be a finite dimensional complex Euclidean space. A linear operator $T : \boldsymbol{U} \to \boldsymbol{U}$ is called *normal* if

$$T^*T = TT^*.$$

Example 4.1. (a) A selfadjoint operator $T : \boldsymbol{U} \to \boldsymbol{U}$ is a normal operator. Indeed, we have $T = T^*$ so that

$$T^*T = T^2 = TT^*.$$

(b) An orthogonal operator $T : \boldsymbol{U} \to \boldsymbol{U}$ is a normal operator. Indeed Proposition 3.37 implies that

$$T^*T = TT^* = \mathbb{1}_{\boldsymbol{U}}. \qquad \square$$

(c) If $T : \boldsymbol{U} \to \boldsymbol{U}$ is a normal operator and $\lambda \in \mathbb{C}$ then $\lambda \mathbb{1}_{\boldsymbol{U}} - T$ is also a normal operator. Indeed

$$(\lambda \mathbb{1}_{\boldsymbol{U}} - T)^* = (\lambda \mathbb{1}_{\boldsymbol{U}})^* - (T^*) = \bar{\lambda} \mathbb{1}_{\boldsymbol{U}} - T^*$$

and we have

$$(\lambda \mathbb{1}_{\boldsymbol{U}} - T)^*(\lambda \mathbb{1}_{\boldsymbol{U}} - T) = (\bar{\lambda} \mathbb{1}_{\boldsymbol{U}} - T^*) \cdot (\lambda \mathbb{1}_{\boldsymbol{U}} - T). \qquad \square$$

Proposition 4.2. *If $T : \boldsymbol{U} \to \boldsymbol{U}$ is a normal operator then so are any of its powers T^k, $k > 0$.*

Proof. To see this we invoke Proposition 3.32 and we deduce

$$(T^k)^* = (T^*)^k$$

Then

$$(T^k)^* T^k = \underbrace{(T^* \cdots T^*)}_{k} \cdot \underbrace{(T \cdots T)}_{k} = \underbrace{(T^* \cdots T^*)}_{k-1} \cdot \underbrace{(T \cdots T)}_{k} \cdot T^*$$

$$= \underbrace{(T^* \cdots T^*)}_{k-2} \cdot \underbrace{(T \cdots T)}_{k} \cdot (T^*)^2 = \cdots = \underbrace{(T \cdots T)}_{k} (T^*)^k = T^k (T^*)^k = T^k (T^k)^*.$$

$\square$

Definition 4.3. Let $\boldsymbol{U}$ be a finite dimensional Euclidean $\boldsymbol{F}$-space. A linear operator $T : \boldsymbol{U} \to \boldsymbol{U}$ is called *orthogonally diagonalizable* if there exists an orthonormal basis of $\boldsymbol{U}$ such that, in this basis the operator T is represented by a diagonal matrix. $\square$

We can unravel a bit the above definition and observe that a linear operator T on an n-dimensional Euclidean $\mathbb{F}$-space is orthogonally diagonalizable if and only if there exists an orthonormal basis $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ and numbers $a_1, \dots, a_n \in \mathbb{F}$ such that

$$T\boldsymbol{e}_k = a_k \boldsymbol{e}_k, \quad \forall k.$$

Thus the above basis is rather special: it is orthonormal, and it consists of eigenvectors of T. The numbers a_k are eigenvalues of T.

Note that the converse is also true. If $\boldsymbol{U}$ admits a n orthonormal basis consisting of eigenvectors of T, then T is orthogonally diagonalizable.

Proposition 4.4. *Suppose that $\boldsymbol{U}$ is a $\overline{\text{complex}}$ Euclidean space of dimension n and $T : \boldsymbol{U} \to \boldsymbol{U}$ is orthogonally diagonalizable. Then T is a normal operator.*

Proof. Fix an orthonormal basis

$$\underline{\boldsymbol{e}} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$$

of $\boldsymbol{U}$ such that, in this basis, the operator T is represented by the diagonal matrix

$$D = \begin{bmatrix} a_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & a_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & a_n \end{bmatrix}, \quad a_1, \dots, a_n \in \mathbb{C}.$$

The computations in Example 3.25 show that the operator T^* is represented in the basis $\underline{\boldsymbol{e}}$ by the matrix D^*. Clearly

$$DD^* = D^*D = \begin{bmatrix} |a_1|^2 & 0 & 0 & \cdots & 0 & 0 \\ 0 & |a_2|^2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & |a_n|^2 \end{bmatrix}.$$

□

4.2. **The spectral decomposition of a normal operator.** We want to show that the converse of Proposition 4.4 is also true. This is a nontrivial and fundamental result of linear algebra.

Theorem 4.5 (Spectral Theorem for Normal Operators)**.** *Let $\boldsymbol{U}$ be an n-dimensional complex Euclidean space and $T : \boldsymbol{U} \to \boldsymbol{U}$ a* normal *operator. Then T is orthogonally diagonalizable, i.e., there exists an orthonormal basis of $\boldsymbol{U}$ consisting of eigenvectors of T.*

Proof. The key fact behind the Spectral Theorem is contained in the following auxiliary result.

Lemma 4.6. *Let $\lambda \in \operatorname{spec}(T)$. Then*

$$\ker(\lambda \mathbb{1}_{\boldsymbol{U}} - T)^2 = \ker(\lambda \mathbb{1}_{\boldsymbol{U}} - T).$$

We first complete the proof of the Spectral Theorem assuming the validity of the above result. Invoking Lemma 2.9 we deduce that

$$\ker(\lambda \mathbb{1}_{\boldsymbol{U}} - T) = \ker(\lambda \mathbb{1}_{\boldsymbol{U}} - T)^2 = \ker(\lambda \mathbb{1}_{\boldsymbol{U}} - T)^2 = \cdots$$

so that the generalized eigenspace of T corresponding to an eigenvalue λ coincides with the eigenspace $\ker(\lambda \mathbb{1}_{\boldsymbol{U}} - T)$,

$$E_\lambda(T) = \ker(\lambda \mathbb{1}_{\boldsymbol{U}} - T).$$

Suppose that

$$\operatorname{spec}(T) = \{ \lambda_1, \dots, \lambda_\ell \}.$$

From Proposition 2.18 we deduce that

$$\boldsymbol{U} = \ker(\lambda_1 \mathbb{1}_{\boldsymbol{U}} - T) \oplus \cdots \oplus \ker(\lambda_\ell \mathbb{1}_{\boldsymbol{U}} - T). \tag{4.1}$$

The next crucial observation is contained in the following elementary result.

Lemma 4.7. *Suppose that λ, μ are two* distinct *eigenvalues of T, and $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{U}$ are eigenvectors*

$$T\boldsymbol{u} = \lambda \boldsymbol{u}, \quad T\boldsymbol{v} = \mu \boldsymbol{v}.$$

Then

$$T^*\boldsymbol{u} = \bar{\lambda}\nu, \quad T^*\boldsymbol{v} = \bar{\mu}\boldsymbol{v},$$

and

$$\boldsymbol{u} \perp \boldsymbol{v}.$$

Proof. Let $S_\lambda = T - \lambda \mathbb{1}_{\boldsymbol{U}}$ so that $S_\lambda \boldsymbol{u} = 0$. Note that $S_\lambda^* = T^* - \bar{\lambda}\mathbb{1}_{\boldsymbol{U}}$ so that we have to show that $S_\lambda^* \boldsymbol{u} = 0$. As explained in Example 4.1(c), the operator S_λ is normal. We deduce that

$$0 = S_\lambda^* S_\lambda \boldsymbol{u} = S_\lambda S_\lambda^* \boldsymbol{u}.$$

Hence

$$0 = \langle S_\lambda S_\lambda^* \boldsymbol{u}, \boldsymbol{u} \rangle = \langle S_\lambda^* \boldsymbol{u}, S_\lambda^* \boldsymbol{u} \rangle = \| S_\lambda^* \boldsymbol{u} \|^2.$$

This proves that $T^* \boldsymbol{u} = \bar{\lambda} \boldsymbol{u}$. A similar argument shows that $T^* \boldsymbol{v} = \bar{\mu} \boldsymbol{v}$.

From the equality $T\boldsymbol{u} = \lambda \boldsymbol{u}$ we deduce

$$\lambda \langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle T\boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{u}, T^* \boldsymbol{v} \rangle = \langle \boldsymbol{u}, \bar{\mu} \boldsymbol{v} \rangle = \mu \langle \boldsymbol{u}, \boldsymbol{v} \rangle.$$

Hence

$$(\lambda - \mu) \langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0.$$

Since $\lambda \neq \mu$ we deduce $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$. □

From the above result we conclude that the direct summands in (4.1) are mutually orthogonal. Set $d_k = \dim \ker(\lambda_k \mathbb{1}_U - T)$. We fix an orthonormal basis

$$\underline{\boldsymbol{e}}(k) = \boldsymbol{e}_1(k), \dots, \boldsymbol{e}_{d_k}(k)$$

of $\ker(\lambda_k \mathbb{1}_U - T)$. By construction, the vectors in this basis are eigenvectors of T. Since the spaces $\ker(\lambda_k \mathbb{1}_{\boldsymbol{U}} - T)$ are mutually orthogonal we deduce from (4.1) that the union of the orthonormal bases $\underline{\boldsymbol{e}}_k$ is an orthonormal basis of $\boldsymbol{U}$ consisting of eigenvectors of T. This completes the proof of the Spectral Theorem, modulo Lemma 4.6. □

Proof of Lemma 4.6. The operator $S = \lambda \mathbb{1}_{\boldsymbol{U}} - T$ is normal so that the conclusion of the lemma follows if we prove that for any normal operator S we have

$$\ker S^2 = \ker S.$$

Note that $\ker S \subset \ker S^2$ so that it suffices to show that $\ker S^2 \subset \ker S$.

Let $\boldsymbol{u} \in \boldsymbol{U}$ such that $S^2 \boldsymbol{u} = 0$. We have to show that $S\boldsymbol{u} = 0$. Note that

$$0 = (S^*)^2 S^2 \boldsymbol{u} = S^* S^* S S \boldsymbol{u} = S^* S S^* S \boldsymbol{u}.$$

Set $A := S^* S$. Note that A is selfadjoint, $A = A^*$ and we can rewrite the above equality as $0 = A^2 \boldsymbol{u}$. Hence

$$0 = \langle A^2 \boldsymbol{u}, \boldsymbol{u} \rangle = \langle A\boldsymbol{u}, A\boldsymbol{u} \rangle = \| A\boldsymbol{u} \|^2.$$

The equality $A\boldsymbol{u} = 0$ now implies

$$0 = \langle A\boldsymbol{u}, \boldsymbol{u} \rangle = \langle S^* S\boldsymbol{u}, \boldsymbol{u} \rangle = \langle S\boldsymbol{u}, S\boldsymbol{u} \rangle = \| S\boldsymbol{u} \|^2.$$

This completes the proof of Lemma 4.6. □

4.3. **The spectral decomposition of a real symmetric operator.** We begin with the real counterpart of Proposition 4.4

Proposition 4.8. *Suppose that $\boldsymbol{U}$ is a real Euclidean space of dimension n and $T : \boldsymbol{U} \to \boldsymbol{U}$ is orthogonally diagonalizable. Then T is a symmetric operator.*

Proof. Fix an orthonormal basis

$$\underline{\boldsymbol{e}} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$$

of $\boldsymbol{U}$ such that, in this basis, the operator T is represented by the diagonal matrix

$$D = \begin{bmatrix} a_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & a_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & a_n \end{bmatrix}, \quad a_1, \dots, a_n \in \mathbb{R}.$$

Clearly $D = D^*$ and the computations in Example 3.25 show that T is a selfadjoint operator. □

We can now state and prove the real counterpart of Theorem 4.5

Theorem 4.9 (Spectral Theorem for Real Symmetric Operators)**.** *Let $\boldsymbol{U}$ be an n-dimensional real Euclidean space and $T : \boldsymbol{U} \to \boldsymbol{U}$ a* symmetric *operator. Then T is orthogonally diagonalizable, i.e., there exists an orthonormal basis of $\boldsymbol{U}$ consisting of eigenvectors of T.*

Proof. We argue by induction on $n = \dim \boldsymbol{U}$. For $n = 1$ the result is trivially true. We assume that the result is true for real symmetric operators action on Euclidean spaces of dimension $< n$ and we prove that it holds for a symmetric operator T on a real n-dimenasional Euclidean space $\boldsymbol{U}$.

To begin with let us observe that T has at least one real eigenvalue. Indeed, if we fix an orthonormal basis $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ of $\boldsymbol{U}$, then in this basis the operator T is represented by a symmetric $n \times n$ real matrix A. As explained in Corolllary 3.29, all the roots of the characteristic polynomial $\det(\lambda \mathbb{1} - A)$ are real, and they coincide with the eigenvalues of T.

Fix one such eigenvalue $\lambda \in \operatorname{spec}(T) \subset \lambda \in \mathbb{R}$ and denote by E_λ the corresponding eigenspace

$$E_\lambda := \ker(\lambda \mathbb{1} - T) \subset \boldsymbol{U}.$$

Lemma 4.10. *The orthogonal complement $E_\lambda^\perp$ is an invariant subspace of T, i.e.,*

$$\boldsymbol{u} \perp E_\lambda \Rightarrow T\boldsymbol{u} \perp E_\lambda.$$

Proof. Let $\boldsymbol{u} \in E_\lambda^\perp$. We have to show that $T\boldsymbol{u} \perp E_\lambda$, i.e., $T\boldsymbol{u} \perp \boldsymbol{v}$, $\forall \boldsymbol{v} \in E_\lambda$. Given such a $\boldsymbol{v}$ we have

$$T\boldsymbol{v} = \lambda \boldsymbol{v}.$$

Next observe that $\boldsymbol{u} \perp \boldsymbol{v}$ since $\boldsymbol{u} \in E_\lambda^\perp$. Hence

$$\langle T\boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{u}, T\boldsymbol{v} \rangle = \lambda \langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0.$$

□

The restriction S_λ of T to $E_\lambda^\perp$ is a symmetric operator $E_\lambda^\perp \to E_\lambda^\perp$ and the induction hypothesis implies that we can find and orthonormal basis of $E_\lambda^\perp$ such that, in this basis, the operator S_λ is represented by a diagonal matrix D_λ. Fix an arbitrary basis $\underline{\boldsymbol{e}}$ of $E_\lambda^\perp$. The union of $\underline{\boldsymbol{e}}$ with is an orthonormal basis of $\boldsymbol{U}$. In this basis T is represented by the block matrix

$$\begin{bmatrix} \mathbb{1}_{E_\lambda} & 0 \\ 0 & D_\lambda \end{bmatrix}.$$

The above matrix is clearly diagonal. □

4.4. **Exercises.**

Exercise 4.1. Let $\boldsymbol{U}$ be a finite dimensional complex Euclidean vector space and $T : \boldsymbol{U} \to \boldsymbol{U}$ a normal operator. Prove that for any complex numbers $a_0, \dots, a_k$ the operator

$$a_0 \mathbb{1}_{\boldsymbol{U}} + a_1 T + \cdots + a_k T^k$$

is a normal operator. □

Exercise 4.2. Let $\boldsymbol{U}$ be a finite dimensional complex vector space and $\boldsymbol{T} : \boldsymbol{U} \to \boldsymbol{U}$ a *normal* operator. Show that the following statements are equivalent.

(i) The operator T is orthogonal.
(ii) If λ is an eigenvalue of T, then $|\lambda| = 1$. □

Exercise 4.3. (a) Prove that the product of two orthogonal operators on a finite dimensional Euclidean space is an orthogonal operator.
(b) Is it true that the product of two selfadjoint operators on a finite dimensional Euclidean space is also a selfadjoint operator? □

Exercise 4.4. Suppose that $\boldsymbol{U}$ is a finite dimensional Euclidean space and $P : \boldsymbol{U} \to \boldsymbol{U}$ is a linear operator such that $P^2 = P$. Show that the following statements are equivalent.

(i) P is the orthogonal projection onto a subspace $\boldsymbol{V} \subset \boldsymbol{U}$.
(ii) $P^* = P$.

□

Exercise 4.5. Suppose that $\boldsymbol{U}$ is a finite dimensional complex Euclidean space and $T : \boldsymbol{U} \to \boldsymbol{U}$ is a normal operator. Show that

$$\boldsymbol{R}(T) = \boldsymbol{R}(T^*).$$ □

Exercise 4.6. Does there exists a symmetric operator $T : \mathbb{R}^3 \to \mathbb{R}^3$ such that

$$T\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad T\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}?$$ □

Exercise 4.7. Show that a normal operator on a complex Euclidean space is selfadjoint if and only if all its eigenvalues are real. □

Exercise 4.8. Suppose that T is a normal operator on a complex Euclidean space $\boldsymbol{U}$ such that $T^7 = T^5$. Prove that T is selfadjoint and $T^3 = T$. □

Exercise 4.9. Let $\boldsymbol{U}$ be a finite dimensional complex Euclidean space and $T : \boldsymbol{U} \to \boldsymbol{U}$ be a selfadjoint operator. Suppose that there exists a vector $\boldsymbol{u}$, a complex number μ, and a number $\varepsilon > 0$ such that

$$\|T\boldsymbol{u} - \mu\boldsymbol{u}\| < \varepsilon\|\boldsymbol{u}\|.$$

Prove that there exists an eigenvalue λ of T such that $|\lambda - \mu| < \varepsilon$. □

5. APPLICATIONS

5.1. **Symmetric bilinear forms.** Suppose that $\boldsymbol{U}$ is a finite dimensional *real* vector space. Recall that a symmetric bilinear form on $\boldsymbol{U}$ is a bilinear map

$$Q : \boldsymbol{U} \times \boldsymbol{U} \to \mathbb{R}$$

such that

$$Q(\boldsymbol{u}_1, \boldsymbol{u}_2) = Q(\boldsymbol{u}_1, \boldsymbol{u}_2).$$

We denote by $\mathrm{Sym}(\boldsymbol{U})$ the space of symmetric bilinear forms on $\boldsymbol{U}$.

Suppose that $\underline{\boldsymbol{e}} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ is a basis of the real vector space $\boldsymbol{U}$. This basis associated to any symmetric bilinear form $Q \in \mathrm{Sym}(\boldsymbol{U})$ a symmetric matrix

$$A = (a_{ij})_{1\leq i,j\leq n}, \;\; a_{ij} = Q(\boldsymbol{e}_i, \boldsymbol{e}_j) = Q(\boldsymbol{e}_j, \boldsymbol{e}_i) = a_{ji}.$$

Note that the form Q is completely determined by the matrix A. Indeed if

$$\boldsymbol{u} = \sum_{i=1}^n u_i \boldsymbol{e}_i, \;\; \boldsymbol{v} = \sum_{j=1}^n v_j \boldsymbol{e}_j,$$

then

$$Q(\boldsymbol{u}, \boldsymbol{v}) = Q\left(\sum_{i=1}^n u_i \boldsymbol{e}_i, \sum_{j=1}^n v_j \boldsymbol{e}_j \right) = \sum_{i,j=1}^n u_i v_j Q(\boldsymbol{e}_i, \boldsymbol{e}_j) = \sum_{i,j=1}^m a_{ij} u_i v_j.$$

The matrix A is called *the symmetric matrix associated to the symmetric form Q in the basis* $\underline{\boldsymbol{e}}$.

Conversely any symmetric $n \times n$ matrix A defines a symmetric bilinear form $Q_A \in \mathrm{Sym}(\mathbb{R}^n)$ defined by

$$Q(\boldsymbol{u}, \boldsymbol{v}) = \sum_{i,j=1}^m a_{ij} u_i v_j, \;\; \forall \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n.$$

If $\langle -, - \rangle$ denotes the canonical inner product on $\mathbb{R}^n$, then we can rewrite the above equality in the more compact form

$$Q_A(\boldsymbol{u}, \boldsymbol{v}) = \langle \boldsymbol{u}, A\boldsymbol{v} \rangle.$$

Definition 5.1. Let $Q \in \mathrm{Sym}(\boldsymbol{U})$ be a symmetric bilinear form on the finite dimensional real space $\boldsymbol{U}$. The *quadratic* form associated to Q is the function

$$\Phi_Q : \boldsymbol{U} \to \mathbb{R}, \;\; \Phi_Q(\boldsymbol{u}) = Q(\boldsymbol{u}, , \boldsymbol{u}), \;\; \forall \boldsymbol{u} \in \boldsymbol{U}. \qquad \square$$

Observe that if $Q \in \boldsymbol{U}$, then we have the *polarization identity*

$$Q(\boldsymbol{u}, \boldsymbol{v}) = \frac{1}{4}\Big(\Phi_Q(\boldsymbol{u} + \boldsymbol{v}) - \Phi_Q(\boldsymbol{u} - \boldsymbol{v}) \Big).$$

This shows that a symmetric bilinear form is completely determined by its associated quadratic form.

Proposition 5.2. *Suppose that* $Q \in \mathrm{Sym}(\boldsymbol{U})$ *and*

$$\underline{\boldsymbol{e}} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}, \;\; \underline{\boldsymbol{f}} = \{\boldsymbol{f}_1, \dots, \boldsymbol{f}_n\}.$$

Denote by S the matrix describing the transition from the basis $\underline{\boldsymbol{e}}$ to the basis $\underline{\boldsymbol{f}}$. In other words, the j-th column of S describes the coordinates of $\boldsymbol{f}_j$ in the basis $\underline{\boldsymbol{e}}$, i.e.,

$$\boldsymbol{f}_j = \sum_{i=1}^n s_{ij} \boldsymbol{e}_i.$$

Denote by A (respectively B) the matrix associated to Q by the basis $\underline{e}$ (respectively $\underline{f}$). Then

$$B = S^{\dagger}AS. \tag{5.1}$$

Proof. We have

$$\begin{aligned} b_{ij} = Q(\boldsymbol{f}_i, \boldsymbol{f}_j) &= Q\left(\sum_{k=1}^{n} s_{ki}\boldsymbol{e}_k, \sum_{\ell=1}^{n} s_{\ell j}\boldsymbol{e}_\ell\right) \\ &= \sum_{k,\ell=1}^{n} s_{ki}s_{\ell j}Q(\boldsymbol{e}_k\boldsymbol{e}_\ell) = \sum_{k,\ell=1}^{n} s_{ki}a_{k\ell}s_{\ell j} \end{aligned}$$

If we denote by $s^{\dagger}_{ij}$ the entries of the transpose matrix $S^{\dagger}$, $s^{\dagger}_{ij} = s_{ji}$, we deduce

$$b_{ij} = \sum_{k,\ell=1}^{n} s^{\dagger}_{ik}a_{k\ell}s_{\ell j}.$$

The last equality shows that b_{ij} is the (i,j)-entry of the matrix $S^{\dagger}AS$. □

☞ *We strongly recommend the reader to compare the change of base formula (5.1) with the change of base formula (2.1).*

Theorem 5.3. *Suppose that Q is a symmetric bilinear form on a finite dimensional real vector space $\boldsymbol{U}$. Then there exist at least one basis of $\boldsymbol{U}$ such that the matrix associated to Q by this basis is a diagonal matrix.*

Proof. We will employ the spectral theory of real symmetric operators. For this reason with fix an Euclidean inner product $\langle -, - \rangle$ on $\boldsymbol{U}$ and the choose a basis $\underline{e}$ of $\boldsymbol{U}$ which is *orthonormal* with respect to the above inner product. We denote by A the symmetric matrix associated to Q by this basis, i.e.,

$$a_{ij} = Q(\boldsymbol{e}_i, \boldsymbol{e}_j).$$

This matrix defines a symmetric operator $T_A : \boldsymbol{U} \to \boldsymbol{U}$ by the formula

$$T_A\boldsymbol{e}_j = \sum_{i=1}^{n} a_{ij}\boldsymbol{e}_i.$$

Let us observe that

$$Q(\boldsymbol{u}, \boldsymbol{v}) = \langle \boldsymbol{u}, T_A\boldsymbol{v} \rangle, \quad \forall \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{U}. \tag{5.2}$$

To verify the above equality we first notice that both sides of the above equalities are bilinear on $\boldsymbol{u}, \boldsymbol{v}$ so that it suffices to check the equality in the special case when the vectors $\boldsymbol{u}, \boldsymbol{v}$ belong to the basis $\underline{e}$. We have

$$\langle \boldsymbol{e}_i, T_A\boldsymbol{e}_j \rangle = \left\langle \boldsymbol{e}_i, \sum_{k} a_{kj}\boldsymbol{e}_k \right\rangle = a_{ij} = Q(\boldsymbol{e}_i, \boldsymbol{e}_j).$$

The spectral theorem for real symmetric operators implies that there exists an orthonormal basis $\underline{f}$ of $\boldsymbol{U}$ such that the matrix B representing T_A in this basis is diagonal. If S denotes the matrix describing the transition to the basis $\underline{e}$ to the basis $\underline{f}$ then the equality (2.1) implies that

$$B = S^{-1}AS.$$

Since the biases $\underline{\boldsymbol{e}}$ and $\underline{\boldsymbol{f}}$ are orthonormal we deduce from Exercise 3.17 that the matrix S is orthogonal, i.e., $S^\dagger S = SS^\dagger = \mathbb{1}$. Hence $S^{-1} = S^\dagger$ and we deduce that

$$B = S^* AS.$$

The above equality and Proposition 5.2 imply that the diagonal matrix B is also the matrix associated to Q by the basis $\underline{\boldsymbol{f}}$. □

Theorem 5.4 (The law of inertia). *Let $\boldsymbol{U}$ be a real vector space of dimension n and Q a symmetric bilinear form on $\boldsymbol{U}$. Suppose that $\underline{\boldsymbol{e}}$, $\underline{\boldsymbol{f}}$ are bases of $\boldsymbol{U}$ such that the matrices associated to Q by these bases are diagonal.*[3] *Then these matrices have the same number of positive (respectively negative, respectively zero) entries on their diagonal.*

Proof. We will show that these two matrices have the same number of positive elements and the same number of negative entries on their diagonals. Automatically then they must have the same number of trivial entries on their diagonals.

We take care of the positive entries first. Denote by A the matrix associated to Q by $\underline{\boldsymbol{e}}$ and by B the matrix associated by $\underline{\boldsymbol{f}}$. We denote by p the number of positive entries on the diagonal of A and by q the number of positive entries on the diagonal of B. We have to show that $p = q$. We argue by contradiction and we assume that $p \neq q$, say $p > q$.

We can label the elements in the basis $\underline{\boldsymbol{e}}$ so that

$$a_{ii} = Q(\boldsymbol{e}_i, \boldsymbol{e}_i) > 0, \;\; \forall i \leq p, \;\; a_{jj} = Q(\boldsymbol{e}_j, \boldsymbol{e}_j) \leq 0, \;\; \forall j > p. \tag{5.3}$$

Observe that since A is diagonal we have

$$Q(\boldsymbol{e}_i, \boldsymbol{e}_j) = 0, \;\; \forall i \neq j. \tag{5.4}$$

Similarly we can label the elements in the basis $\underline{\boldsymbol{f}}$ so that

$$b_{kk} = Q(\boldsymbol{f}_k, \boldsymbol{f}_k) > 0, \;\; \forall k \leq q, \;\; b_{\ell\ell} = Q(\boldsymbol{f}_\ell, \boldsymbol{f}_\ell) \leq 0, \;\; \forall \ell > q. \tag{5.5}$$

Since B is diagonal we have

$$Q(\boldsymbol{f}_i, \boldsymbol{f}_j) = 0, \;\; \forall i \neq j. \tag{5.6}$$

Denote $\boldsymbol{V}$ the subspace spanned by the vectors $\boldsymbol{e}_i$, $i = 1, \dots, p$, and by $\boldsymbol{W}$ the subspace spanned by the vectors $\boldsymbol{f}_{q+1}, \dots, \boldsymbol{f}_n$. From the equalities (5.3), (5.4), (5.5), (5.6) we deduce that

$$Q(\boldsymbol{v}, \boldsymbol{v}) > 0, \;\; \forall \boldsymbol{v} \in \boldsymbol{V} \setminus 0, \tag{5.7a}$$

$$Q(\boldsymbol{w}, \boldsymbol{w}) \leq 0, \;\; \forall \boldsymbol{w} \in \boldsymbol{W} \setminus 0. \tag{5.7b}$$

On the other hand, we observe that $\dim \boldsymbol{V} = p$, $\dim \boldsymbol{W} = n - q$. Hence

$$\dim \boldsymbol{V} + \dim \boldsymbol{W} = n + p - q > n > \dim \boldsymbol{U}$$

so that there exists a vector

$$\boldsymbol{u} \in (\boldsymbol{V} \cap \boldsymbol{W}) \setminus 0.$$

The vector $\boldsymbol{u}$ cannot simultaneously satisfy both inequalities (5.7a) and (5.7b). This contradiction implies that $p = q$.

Using the above argument for the form $-Q$ we deduce that A and B have the same number of negative elements on their diagonals. □

[3] Such a bases are called *diagonalizing* bases of Q.

The above theorem shows that no matter what diagonalizing basis of Q we choose, the diagonal matrix representing Q in that basis will have the same number of positive negative and zero elements on its diagonal. We will denote these common numbers by $\mu_+(Q)$, $\mu_-(Q)$ and respectively $\mu_0(Q)$. These numbers are called the *indices of inertia* of the symmetric form Q. The integer $\mu_-(Q)$ is called *Morse index* of the symmetric form Q and the difference

$$\sigma(Q) = \mu_+(Q) - \mu_-(Q).$$

is called the *signature* of the form Q.

Definition 5.5. A symmetric bilinear form $Q :\in \mathrm{Sym}(U)$ is called *positive definite* is

$$Q(\boldsymbol{u},\boldsymbol{u}) > 0, \;\; \forall \boldsymbol{u} \in \boldsymbol{U} \setminus 0.$$

It is called *positive semidefinite* if

$$Q(\boldsymbol{u},\boldsymbol{u}) \geq 0, \;\; \forall \boldsymbol{u} \in \boldsymbol{U}.$$

It is called *negative (semi)definite* if $-Q$ is positive (semi)definite. □

We observe that a symmetric bilinear form on an n-dimensional real space $\boldsymbol{U}$ is positive definite if and only if $\mu_+(Q) = n = \dim \boldsymbol{U}$.

Definition 5.6. A real symmetric $n\times n$ matrix A is called *positive definite* if and only if the associated symmetric bilinear form $Q_A \in \mathrm{Sym}(\mathbb{R}^n)$ is positive definite. □

5.2. **Nonnegative operators.** Suppose that $\boldsymbol{U}$ is a finite dimensional Euclidean $\mathbb{F}$-vector space.

Definition 5.7. A linear operator $T : \boldsymbol{U} \to \boldsymbol{U}$ is called *nonnegative* if the following hold.

(i) T is selfadjoint, $T^* = T$.
(ii) $\langle T\boldsymbol{u}, \boldsymbol{u}\rangle \geq 0$, for all $\boldsymbol{u} \in \boldsymbol{U}$.

The operator is called *positive* if it is nonnegative and $\langle T\boldsymbol{u}, \boldsymbol{u}\rangle = 0 \Longleftrightarrow \boldsymbol{u} = 0$. □

Example 5.8. Suppose that $T : \boldsymbol{U} \to \boldsymbol{U}$ is a linear operator. Then the operator $S := T^*T$ is nonnegative. Indeed, it is selfadjoint and

$$\langle S\boldsymbol{u}, \boldsymbol{u}\rangle = \langle T^*T\boldsymbol{u}, \boldsymbol{u}\rangle = \langle T\boldsymbol{u}, T\boldsymbol{u}\rangle = \|T\boldsymbol{u}\|^2 \geq 0.$$

Note that S is positive if and only $\ker S = 0$ so that S is injective. □

Definition 5.9. Suppose that $T : \boldsymbol{U} \to \boldsymbol{U}$ is a linear operator on the finite dimensional $\mathbb{F}$-space $\boldsymbol{U}$. A *square root* of T is a linear operator $S : \boldsymbol{U} \to \boldsymbol{U}$ such that $S^2 = T$. □

Theorem 5.10. *Let $T : \boldsymbol{U} \to \boldsymbol{U}$ be a linear operator on the n-dimensional Euclidean $\mathbb{F}$-space. Then the following statements are equivalent.*

(i) *The operator T is nonnegative.*
(ii) *The operator T is selfadjoint and all its eigenvalues are nonnegative.*
(iii) *The operator T admits a nonnegativesquare root.*
(iv) *The operator T admits a selfadjoint root.*
(v) *there exists an operator $S : \boldsymbol{U} \to \boldsymbol{U}$ such that $T = S^*S$.*

Proof. (i) ⇒ The operator T being nonnegative is also selfadjoint. Hence all its eigenvalues are real. If λ is an eigenvalue of T and $\boldsymbol{u} \in \ker(\lambda\mathbb{1}_{\boldsymbol{U}} - T) \setminus 0$. then

$$\lambda\|\boldsymbol{u}\|^2 = \langle T\boldsymbol{u}, \boldsymbol{u}\rangle \geq 0.$$

This implies $\lambda \geq 0$.

(ii) $\Rightarrow$ (iii) Since T is selfadjoint, there exists an orthonormal basis $\underline{\boldsymbol{e}} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ of $\boldsymbol{U}$ such that, in this basis, the operator T is represented by the diagonal matrix

$$A = \mathrm{Diag}(\lambda_1, \dots, .\lambda_n),$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of T. They are all nonnegative so we can form a new diagonal matrix

$$B = \mathrm{Diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}).$$

The matrix B defines a selfadjoint linear operator S on $\boldsymbol{U}$ which is represented by the matrix B in the basis $\underline{\boldsymbol{e}}$. More precisely

$$S\boldsymbol{e}_i = \sqrt{\lambda_i}\boldsymbol{e}_i, \;\; \forall i = 1, \dots, n.$$

If $\boldsymbol{u} = \sum_{i=1}^n u_i \boldsymbol{e}_i$, then

$$S\boldsymbol{u} = \sum_{i=1}^n \sqrt{\lambda_i} u_i \boldsymbol{e}_i, \;\; \langle S\boldsymbol{u}, \boldsymbol{u}\rangle = \sum_{i=1}^n \sqrt{\lambda_i} |u_i|^2 \geq 0.$$

Hence S is nonnegative. From the obvious equality $A = B^2$ we deduce $T = S^2$ so that S is nonnegative square root of T.

The implication (iii) $\Rightarrow$ (iv) is obvious because any nonnegative square root of T is automatically a selfadjoint square root. To prove the implication (iv) $\Rightarrow$ (v) observe that if S is a selfadjoint square root of T then

$$T = S^2 = S^* S.$$

The implication (v) $\Rightarrow$ (i) was proved in Example 5.8. □

Proposition 5.11. *Let $\boldsymbol{U}$ be a finite dimension Euclidean $\mathbb{F}$-space. Then any nonnegative operator $T : \boldsymbol{U} \to \boldsymbol{U}$ admits a* unique *nonnegative square root.*

Proof. We have an orthogonal decomposition

$$\boldsymbol{U} = \bigoplus_{\lambda \in \mathrm{spec}(T)} \ker(\lambda \mathbb{1}_{\boldsymbol{U}} - T)$$

so that any vector $\boldsymbol{u} \in \boldsymbol{U}$ can be written uniquely as

$$\boldsymbol{u} = \sum_{\lambda \in \mathrm{spec}(T)} \boldsymbol{u}_\lambda, \;\; \boldsymbol{u}_\lambda \in \ker(\lambda \mathbb{1}_{\boldsymbol{U}} - T). \tag{5.8}$$

Moreover

$$T\boldsymbol{u} = \sum_{\lambda \in \mathrm{spec}(T)} \lambda \boldsymbol{u}_\lambda.$$

Suppose that S is a nonnegative square root of T. If $\mu \in \mathrm{spec}(S)$ and $\boldsymbol{u} \in \ker(\mu \mathbb{1}_{\boldsymbol{U}} - S)$, then

$$T\boldsymbol{u} = S^2 \boldsymbol{u} = S(S\boldsymbol{u}) = S(\mu \boldsymbol{u}) = \mu^2 \boldsymbol{u}.$$

Hence

$$\mu^2 \in \mathrm{spec}(T)$$

and

$$\ker(\mu \mathbb{1}_{\boldsymbol{U}} - S) \subset \ker(\mu^2 \mathbb{1}_{\boldsymbol{U}} - T).$$

We have a similar orthogonal decomposition

$$\boldsymbol{U} = \bigoplus_{\mu \in \mathrm{spec}(S)} \ker(\mu \mathbb{1}_{\boldsymbol{U}} - S) \subset \bigoplus_{\mu \in \mathrm{spec}(S)} \ker(\mu^2 \mathbb{1}_{\boldsymbol{U}} - T) \subset \bigoplus_{\lambda \in \mathrm{spec}(T)} \ker(\lambda \mathbb{1}_{\boldsymbol{U}} - T) = \boldsymbol{U}.$$

This implies that

$$\operatorname{spec}(T) = \{\mu^2;\ \mu \in \operatorname{spec}(S)\},\ \ \ker(\mu \mathbb{1}_{\boldsymbol{U}} - S) = \ker(\mu^2 \mathbb{1}_{\boldsymbol{U}} - T),\ \ \forall \mu \in \operatorname{spec}(S).$$

Since all the eigenvalues of S are nonnegative we deduce that for any $\lambda \in \operatorname{spec}(T)$ we have $\sqrt{\lambda} \in \operatorname{spec}(S)$. Thus if $\boldsymbol{u}$ is decomposed as in (5.8),

$$\boldsymbol{u} = \sum_{\lambda \in \operatorname{spec}(T)} \boldsymbol{u}_\lambda,$$

then

$$S\boldsymbol{u} = \sum_{\lambda \in \operatorname{spec}(T)} \sqrt{\lambda}\boldsymbol{u}_\lambda.$$

The last equality determines S uniquely. □

Definition 5.12. If T is a nonnegative operator, then its unique nonnegative square root is denoted by $\sqrt{T}$. □

5.3. **Exercises.**

Exercise 5.1. Let Q be a symmetric bilinear form on an n-dimensional real vector space. Prove that there exists a basis of $\boldsymbol{U}$ such that the matrix associated to Q by this basis is diagonal and all the entries belong to $\{-1, 0, 1\}$. □

Exercise 5.2. Let Q be a symmetric bilinear form on an n-dimensional real vector space $\boldsymbol{U}$ with indices of inertia μ_+, μ_-, μ_0. We say Q is positive definite on a subspace $\boldsymbol{V} \subset \boldsymbol{U}$ if

$$Q(\boldsymbol{v}, \boldsymbol{v}) > 0, \;\; \forall \boldsymbol{v} \in \boldsymbol{V} \setminus 0.$$

(a) Prove that if Q is positive definite on a subspace $\boldsymbol{V}$, then $\dim \boldsymbol{V} \leq \mu_+$.
(b) Show that there exists a subspace of dimension μ_+ on which Q is positive definite.

Hint: (a) Choose a diagonalizing basis $\underline{\boldsymbol{e}} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ of Q. Assume that $Q(\boldsymbol{e}_i, \boldsymbol{e}_i) > 0$, for $i = 1, \dots, \mu_+$ and $Q(\boldsymbol{e}_j, \boldsymbol{e}_j) \leq 0$ for $j > \mu_+$. Argue by contradiction that

$$\dim \boldsymbol{V} + \dim \operatorname{span}\{\boldsymbol{e}_j; \;\; j > \mu_+\} \leq \dim \boldsymbol{U} = n.$$ □

Exercise 5.3. Let Q be a symmetric bilinear form on an n-dimensional real vector space $\boldsymbol{U}$. with indices of inertia μ_+, μ_-, μ_0. Define

$$\operatorname{Null}(Q) := \big\{\, \boldsymbol{u} \in \boldsymbol{U}; \;\; Q(\boldsymbol{u}, \boldsymbol{v}) = 0, \;\; \forall \boldsymbol{v} \in \boldsymbol{U}. \,\big\}.$$

Show that $\operatorname{Null}(Q)$ is a vector subspace of $\boldsymbol{U}$ of dimension μ_0. □

Exercise 5.4 (Jacobi)**.** For any $n \times n$ matrix M we denote by M_i the $i \times i$ matrix determined by the first i rows and columns of M.

Suppose that Q is a symmetric bilinear form on the real vector space $\boldsymbol{U}$ of dimension n. Fix a basis $\underline{\boldsymbol{e}} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ of $\boldsymbol{U}$. Denote by A the matrix associated to Q by the basis $\underline{\boldsymbol{e}}$ and assume that

$$\Delta_i := \det A_i \neq 0, \;\; \forall i = 1, \dots, n.$$

(a) Prove that there exists a basis $\underline{\boldsymbol{f}} = (\boldsymbol{f}_1, \dots, \boldsymbol{f}_n)$ of $\boldsymbol{U}$ with the following properties.

(i) $\operatorname{span}\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_i\} = \operatorname{span}\{\boldsymbol{f}_1, \dots, \boldsymbol{f}_i\}$, $\forall i = 1, \dots, n$.
(ii) $Q(\boldsymbol{f}_k, \boldsymbol{e}_i) = 0$, $\forall 1 \leq i < k \leq n$, $Q(\boldsymbol{f}_k, \boldsymbol{e}_k) = 1$, $\forall k = 1, \dots, n$.

Hint: For fixed k, express the vector $\boldsymbol{f}_k$ in terms of the vectors $\boldsymbol{e}_i$,

$$\boldsymbol{f}_k = \sum_{i=1}^{n} s_{ik} \boldsymbol{e}_i$$

and then show that the conditions (i) and (ii) above uniquely determine the coefficients s_{ik}, again with k fixed.
(b) If $\underline{\boldsymbol{f}}$ is the basis found above, show that

$$Q(\boldsymbol{f}_k, \boldsymbol{f}_i) = 0, \;\; \forall i \neq k,$$
$$Q(\boldsymbol{f}_k, \boldsymbol{f}_k) = \frac{\Delta_{k-1}}{\Delta_k}, \; \forall k = 1, \dots, n, \;\; \Delta_0 := 1.$$

(c) Show that the Morse index $\mu_-(Q)$ is the number of sign changes in the sequence

$$1, \Delta_1, \Delta_2, \dots, \Delta_n.$$ □

Exercise 5.5 (Sylvester)**.** For any $n \times n$ matrix M we denote by M_i the $i \times i$ matrix determined by the first i rows and columns of M.

Suppose that Q is a symmetric bilinear form on the real vector space $\boldsymbol{U}$ of dimension n. Fix a basis $\underline{\boldsymbol{e}} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ of $\boldsymbol{U}$ and denote by A the matrix associated to Q by the basis $\underline{\boldsymbol{e}}$. Prove that the following statements are equivalent.

(i) Q is positive definite.
(ii) $\det A_i > 0, \forall i = 1, \dots, n$. □

Exercise 5.6. (a) Let and $f : [0,1] \to [0,\infty)$ a continuous function which is not identically zero. For any $k = 01, 2, \dots$ we define the *k-th momentum* of f to be the real number

$$\mu_k := \mu_k(f) = \int_0^1 x^k f(x)dx.$$

Prove that the symmetric $(n+1) \times (n+1)$-matrix symmetric matrix

$$A = (a_{ij})_{0 \le i,j \le n}, \;\; a_{ij} = \mu_{i+j}.$$

is positive definite.
Hint: Associate to any vector

$$\boldsymbol{u} = \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_n \end{bmatrix}$$

the polynomial $P_{\boldsymbol{u}}(x) = u_0 + u_1 x + \cdots + u_n x^n$ and then express the integral

$$\int_0^1 P_{\boldsymbol{u}}(x) P_{\boldsymbol{v}}(x) f(x)$$

in terms of A.
(b) Prove that the symmetric $n \times n$ symmetric matrix

$$B = (b_{ij})_{1 \le i,j \le n}, \;\; b_{ij} = \frac{1}{i+j}$$

is positive definite.
(c) Prove that the symmetric $(n+1) \times (n+1)$ symmetric matrix

$$C = (c_{ij})_{0 \le i,j \le n}, \;\; c_{ij} = (i+j)!,$$

where $0! := 1$, $n! = 1 \cdot 2 \cdots n$.
Hint: Show that for any $k = 0, 1, 2, \dots$ we have

$$\int_0^\infty x^k e^{-x} dx = k!,$$

and then due the trick in (a). □

Exercise 5.7. (a) Show that the 2×2-symmetric matrix

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

is positive definite.
(b) Denote by Q_A the symmetric bilinear form on $\mathbb{R}^2$ defined by then above matrix A. Since Q_A is positive definite, it defines an inner product $\langle -, - \rangle_A$ on $\mathbb{R}^2$,

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle_A = Q_A(\boldsymbol{u}, \boldsymbol{v}) = \langle \boldsymbol{u}, A\boldsymbol{v} \rangle$$

where $\langle -, - \rangle$ denotes the canonical inner product on $\mathbb{R}^2$. Denote by $T^\#$ the adjoint of T with respect to the inner product $\langle -, - \rangle_A$, i.e.,

$$\langle T\boldsymbol{u}, \boldsymbol{v} \rangle_A = \langle \boldsymbol{u}, T^\# \boldsymbol{v} \rangle_A, \;\; \forall \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^2. \tag{\#}$$

Show that

$$T^{\#} = A^{-1}T^{*}A,$$

where T^* is the adjoint of T with respect to the canonical inner product $\langle -, - \rangle$. What does this formula tell you in the special case when T is described by the symmetric matrix

$$\left[\begin{array}{cc} 1 & 2 \\ 2 & 3 \end{array}\right]?$$

Hint: In (#) use the equalities $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_A = \langle \boldsymbol{x}, A\boldsymbol{y} \rangle$, $\langle T\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, T^*\boldsymbol{y} \rangle$, $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$. □

Exercise 5.8. Suppose that $\boldsymbol{U}$ is a finite dimensional Euclidean $\mathbb{F}$-space and $T : \boldsymbol{U} \to \boldsymbol{U}$ is an invertible operator. Prove that $\sqrt{T^*T}$ is invertible and the operator $S = T(\sqrt{T^*T})^{-1}$ is orthogonal. □

Exercise 5.9. (a) Suppose that $\boldsymbol{U}$ is a finite dimensional real Euclidean space and $Q \in \mathrm{Sym}(\boldsymbol{U})$ is a positive definite symmetric bilinear form. Prove that there exists a unique positive operator

$$T : \boldsymbol{U} \to \boldsymbol{U}$$

such that

$$Q(\boldsymbol{u}, \boldsymbol{v}) = \langle T\boldsymbol{u}, T\boldsymbol{v} \rangle, \quad \forall \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{U}. \qquad \square$$

6. ELEMENTS OF LINEAR TOPOLOGY

6.1. **Normed vector spaces.** As usual, $\mathbb{F}$ will denote one of the fields $\mathbb{R}$ or $\mathbb{C}$.

Definition 6.1. A norm on the $\mathbb{F}$-vector space $\boldsymbol{U}$ is a function

$$\|-\|:\boldsymbol{U}\to[0,\infty),$$

satisfying the following conditions.

(i) (Nondegeneracy)

$$\|\boldsymbol{u}\|=0\Longleftrightarrow\boldsymbol{u}=0.$$

(ii) (Homogeneity)

$$\|c\boldsymbol{u}\|=|c|\cdot\|\boldsymbol{u}\|,\ \ \forall\boldsymbol{u}\in\boldsymbol{U},\ \ c\in\boldsymbol{F}.$$

(iii) (Triangle inequality)

$$\|\boldsymbol{u}+\boldsymbol{v}\|\leq\|\boldsymbol{u}\|+\|\boldsymbol{v}\|,\ \ \forall\boldsymbol{u},\boldsymbol{v}\in\boldsymbol{U}.$$

A *normed* $\mathbb{F}$-*space* is a pair $(\boldsymbol{U},\|-\|)$ where $\boldsymbol{U}$ is a $\mathbb{F}$-vector space and $\|-\|$ is a norm on $\boldsymbol{U}$. □

The distance between two vectors $\boldsymbol{u},\boldsymbol{v}$ in a normed space $(\boldsymbol{U},\|-\|)$ is the quantity

$$\operatorname{dist}(\boldsymbol{u},\boldsymbol{v}):=\|\boldsymbol{u}-\boldsymbol{v}\|.$$

Note that the triangle inequality implies that

$$\operatorname{dist}(\boldsymbol{u},\boldsymbol{w})\leq\operatorname{dist}(\boldsymbol{u},\boldsymbol{v})+\operatorname{dist}(\boldsymbol{v},\boldsymbol{w}),\ \ \forall\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in\boldsymbol{U}\tag{6.1a}$$

$$\big|\,\|\boldsymbol{u}\|-\|\boldsymbol{v}\|\,\big|\leq\|\boldsymbol{u},\boldsymbol{v}\|,\ \ \forall\boldsymbol{u},\boldsymbol{v}\in\boldsymbol{U}.\tag{6.1b}$$

Example 6.2. (a) The absolute value $|-|:\mathbb{R}\to[0,\infty)$ is a norm on the one-dimensional real vector space $\mathbb{R}$.
(b) The absolute value $|-|:\mathbb{C}\to[0,\infty)$, $|x+\boldsymbol{i}y|=\sqrt{x^2+y^2}$ is a norm on the one dimensional complex vector space $\mathbb{C}$.
(c) If $\boldsymbol{U}$ is an Euclidean $\mathbb{F}$-space with inner product $\langle -,-\rangle$ then the function

$$|-|:\boldsymbol{U}\to[0,\infty),\ \ |u|=\sqrt{\langle\boldsymbol{u},\boldsymbol{u}\rangle}$$

is a norm on $\boldsymbol{U}$ called the *norm determined by the inner product*.
(d) Consider the space $\mathbb{R}^N$ with canonical basis $\boldsymbol{e}_1,\dots,\boldsymbol{e}_N$. The function

$$|-|_2:\mathbb{R}^N\to[0,\infty),\ \ |u_1\boldsymbol{e}_1+\cdots+u_N\boldsymbol{e}_N|_2=\sqrt{|u_1|^2+\cdots+|u_N|^2}$$

is the norm associated to the canonical inner product on $\mathbb{R}^n$.

The functions

$$|-|_1,\ |-|_\infty:\mathbb{R}^N\to[0,\infty)$$

defined by

$$|u_1\boldsymbol{e}_1+\cdots+u_N\boldsymbol{e}_N|_\infty:=\max_{1\leq k\leq N}|u_k|,\ \ |\boldsymbol{u}|_1=\sum_{k=1}^N|u_k|,$$

are also a norm on $\mathbb{R}^N$. The proof of this fact is left as an exercise. The norm $|-{}_1$ is sometimes known as the *taxicab norm*. One can define similarly norms $|-|_2$ and $|-|_\infty$ on $\mathbb{C}^n$.
(e) Let $\boldsymbol{U}$ denote the space $C([0,1])$ of continuous functions $\boldsymbol{u}:[0,1]\to\mathbb{R}$. Then the function

$$\|-\|_\infty:C([0,1])\to[0,\infty),\ \ \|\boldsymbol{u}\|_\infty:=\sup_{x\in[0,1]}|\boldsymbol{u}(x)|$$

is a norm on $\boldsymbol{U}$. The proof is left as an exercise.

(f) If $(\boldsymbol{U}_0, \|-\|_0)$ and $(\boldsymbol{U}_1, \|-\|_1)$ are two normed $\mathbb{F}$-spaces, then their cartesian product $\boldsymbol{U}_0 \times \boldsymbol{U}_1$ is equipped with a natural norm

$$\|(\boldsymbol{u}_0, \boldsymbol{u}_1)\| := \|\boldsymbol{u}_0\|_0 + \|\boldsymbol{u}_1\|_1, \;\; \forall \boldsymbol{u}_0 \in \boldsymbol{U}_0, \;\; \boldsymbol{u}_1 \in \boldsymbol{U}_1. \qquad \square$$

Theorem 6.3 (Minkowski)**.** *Consider the space $\mathbb{R}^N$ with canonical basis $\boldsymbol{e}_1, \dots, \boldsymbol{e}_N$. Let $p \in (1,\infty)$. Then the function*

$$|-|_p : \mathbb{R}^N \to [0,\infty), \;\; |u_1\boldsymbol{e}_1 + \cdots + u_n\boldsymbol{e}_n|_p = \left(\sum_{k=1}^{p} |u_k|^p\right)^{\frac{1}{p}}$$

is a norm on $\mathbb{R}^N$.

Proof. The non degeneracy and homogeneity of $|-|_p$ are obvious. The tricky part is the triangle inequality. We will carry the proof of this inequality in several steps. Define $q \in (1,\infty)$ by the equality

$$1 = \frac{1}{p} + \frac{1}{q} \Longleftrightarrow q = \frac{p}{p-1}.$$

Step 1. Young's inequality. We will prove that if a, b are nonnegative real numbers then

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}. \tag{6.2}$$

The inequality is obviously true if $ab = 0$ so we assume that $ab \neq 0$. Set $\alpha := \frac{1}{p}$ and define

$$f : (0,\infty) \to \mathbb{R}, \;\; f(x) = x^\alpha - \alpha x + \alpha - 1.$$

Observe that

$$f'(x) = \alpha(x^{\alpha-1} - 1).$$

Hence $f'(x) = 0 \Longleftrightarrow x = 1$. Moreover, since $\alpha < 1$ we deduce

$$f''(1) = \alpha(\alpha - 1) < 0$$

so 1 is a global max of $f(x)$, i.e.,

$$f(x) \leq f(1) = 0, \;\; \forall x > 0.$$

Now let $x = \frac{a^p}{b^q}$. We deduce

$$\left(\frac{a^p}{b^q}\right)^{\frac{1}{p}} - \frac{1}{p}\frac{a^p}{b^q} \leq 1 - \frac{1}{p} = \frac{1}{q} \Rightarrow b^q\left(\frac{a^p}{b^q}\right)^{\frac{1}{p}} - \frac{a^p}{p} \leq \frac{b^q}{q}$$

$$\Rightarrow ab^{q-\frac{q}{p}} \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

This is precisely (6.2) since $q - \frac{q}{p} = q(1 - \frac{1}{p}) = 1$.

Step 2. Hölder's inequality. We prove that for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$ we have

$$|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leq |\boldsymbol{x}|_p |\boldsymbol{y}|_q, \tag{6.3}$$

where $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ is the canonical inner product of the vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$,

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1 y_1 + \cdots + x_N x_N.$$

Clearly it suffices to prove the inequality only in the case $\boldsymbol{x}, \boldsymbol{y} \neq 0$. Using Young's inequality with

$$a = \frac{|x_k|}{|\boldsymbol{x}|_p}, \;\; b = \frac{|y_k|}{|\boldsymbol{y}|_q}$$

we deduce

$$\frac{|x_k y_k|}{|\boldsymbol{x}|_p \cdot |\boldsymbol{y}|_q} \leq \frac{1}{p}\frac{|x_k|^p}{|\boldsymbol{x}|_p^p} + \frac{1}{q}\frac{|y_k|^q}{|\boldsymbol{y}|_q^q}, \ \ \forall k = 1, \dots, N.$$

Summing over k we deduce

$$\frac{1}{|\boldsymbol{x}|_p \cdot |\boldsymbol{y}|_q} \sum_{k=1} |x_k y_k| \leq \frac{1}{p|\boldsymbol{x}|_p^p} \underbrace{\sum_{k=1}^N |x_k|^p}_{=|\boldsymbol{x}|_p^p} + \frac{1}{q|\boldsymbol{y}|_q^q} \underbrace{\sum_{k=1}^N |y_k|^q}_{=|\boldsymbol{y}|_q^q} = \frac{1}{p} + \frac{1}{q}.$$

Hence

$$\sum_{k=1} |x_k x_k| \leq |\boldsymbol{u}|_p \cdot |\boldsymbol{v}|_q.$$

Now observe that

$$|\langle \boldsymbol{x}, \boldsymbol{y}\rangle| = \left|\sum_{k=1}^N x_k y_k\right| \leq \sum_{k=1} |x_k y_k|.$$

Step 3. Conclusion. We have

$$|\boldsymbol{u}+\boldsymbol{v}|_p^p = \sum_{k=1}^N |u_k + v_k|^p = \sum_{k=1}^N |u_k + v_k| \cdot |u_k + v_k|^{p-1}$$

$$\leq \sum_{k=1}^N |u_k| \cdot |u_k + v_k|)^{p-1} + \sum_{k=1}^N |v_k| \cdot |u_k + v_k|)^{p-1}.$$

Using Hölder's inequality and the equalities $q(p-1) = p$, $\frac{1}{q} = \frac{p}{q} \cdot \frac{1}{p}$ we deduce

$$\sum_{k=1}^N |u_k| \cdot |u_k + v_k|^{p-1} \leq \left(\sum_{k=1}^N |u_k|^p\right)^{\frac{1}{p}} \cdot \left(\sum_{k=1}^N |u_k + v_k|^{q(p-1)}\right)^{\frac{1}{q}} = |\boldsymbol{u}|_p \cdot |\boldsymbol{u}+\boldsymbol{v}|_p^{\frac{p}{q}}$$

Similarly, we have

$$\sum_{k=1}^N |v_k| \cdot |u_k + v_k|^{p-1} \leq |\boldsymbol{v}|_p \cdot |\boldsymbol{u}+\boldsymbol{v}|_p^{\frac{p}{q}}.$$

Hence

$$|\boldsymbol{u}+\boldsymbol{v}|_p^p \leq |\boldsymbol{u}+\boldsymbol{v}|^{\frac{p}{q}} \big(\, |\boldsymbol{u}|_p + |\boldsymbol{v}|_p\,\big),$$

so that

$$|\boldsymbol{u}+\boldsymbol{v}|_p = |\boldsymbol{u}+\boldsymbol{v}|_p^{p-\frac{p}{q}} \leq |\boldsymbol{u}|_p + |\boldsymbol{v}|_p.$$

□

Definition 6.4. Suppose that $\|-\|_0$ and $\|-\|_1$ are two norms on the same vector space $\boldsymbol{U}$ we say that $\|-\|_1$ is stronger than $\|-\|_0$, and we denote this by $\|-\|_0 \prec \|-\|_1$ if there exists $C > 0$ such that

$$\|\boldsymbol{u}\|_0 \leq C\|\boldsymbol{u}\|_1, \ \ \forall \boldsymbol{u} \in \boldsymbol{U}.$$

□

The relation $\prec$ is a partial order on the set of norms on a vector space $\boldsymbol{U}$. This means that if $\|-\|_0, \|-\|_1, \|-\|_2$ are three norms such that

$$\|-\|_0 \prec \|-\|_1, \ \ \|-\|_1 \prec \|-\|_2,$$

then

$$\|-\|_0 \prec \|-\|_2.$$

Definition 6.5. Two norms $\|-\|_0$ and $\|-\|_1$ on the same vector space $\boldsymbol{U}$ are called *equivalent*, and we denote this by $\|-\|_0 \sim \|-\|_1$, if

$$\|_0 - \|_1 \prec \|-\|_1 \prec \|-\|_0. \qquad \square$$

Observe that two norms $\|-0\|_1$ are equivalent if and only if there exist positive constants c, C such that

$$\|-\|_0 \sim \|-\|_1 \Longleftrightarrow \exists c, C > 0: \quad c\|\boldsymbol{u}\|_1 \leq \|\boldsymbol{u}\|_0 \leq C\|\boldsymbol{u}\|_1, \forall \boldsymbol{u} \in \boldsymbol{U}. \tag{6.4}$$

The relation "$\sim$" is an equivalence relation on the space of norms on a given vector space.

Proposition 6.6. *The norms $|-|_p$, $p \in [1,\infty]$, on $\mathbb{R}^N$ are all equivalent with the norm $|-|_\infty$*

Proof. Indeed for $\boldsymbol{u} \in \mathbb{R}^N$ we have

$$|\boldsymbol{u}|_p = \left(\sum_{k=1}^N |u_k|^p\right)^{\frac{1}{p}} \leq \left(N \max_{1\leq k\leq N} |u_k|^p\right)^{\frac{1}{p}} = N^{\frac{1}{p}} |\boldsymbol{u}|_\infty,$$

so that $|-|_p \prec |-|_\infty$. Conversely

$$|\boldsymbol{u}|_\infty = \left(\max_{1\leq k\leq N} |u_k|^p\right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^N |u_k|^p\right)^{\frac{1}{p}} = |\boldsymbol{u}|_p.$$

$\square$

Proposition 6.7. *The norm $|-|_\infty$ on $\mathbb{R}^N$ is stronger than any other norm $\|-\|$ on $\mathbb{R}^N$.*

Proof. Let

$$C = \max_{k=1}^N \|\boldsymbol{e}_k\|,$$

where $\boldsymbol{e}_1, \dots, \boldsymbol{e}_N$ is the canonical basis in We have

$$\|\boldsymbol{u}\| = \|u_1\boldsymbol{e}_1 + \cdots + u_n\boldsymbol{e}_N\| \leq |u_1|\|\boldsymbol{e}_1\| + \cdots + |u_N|\|\boldsymbol{e}_N\|.$$

$\square$

Corollary 6.8. *For any $p \in [1,\infty]$, the norm $|-|_p$ on $\mathbb{R}^N$ is stronger than any other norm $\|-\|$ on $\mathbb{R}^N$.*

Proof. Indeed the norm $|-|_p$ is stronger than $\_\infty$ which is stronger than $\|-\|$. $\square$

6.2. **Convergent sequences.**

Definition 6.9. Let $\boldsymbol{U}$ be a vector space. A sequence $(\boldsymbol{u}(n))_{n\geq 0}$ of vectors in $\boldsymbol{U}$ is said to converge to $\boldsymbol{u}^*$ in the norm $\|-\|$ if

$$\lim_{n\to\infty} \|\boldsymbol{u}(n) - \boldsymbol{u}*\| = 0,$$

i.e.,

$$\forall \varepsilon > 0 \;\; \exists \nu = \nu(\varepsilon) > 0 \;\text{ such that }\; \|\boldsymbol{u}(n) - \boldsymbol{u}^*\| < \varepsilon, \;\; \forall n \geq \nu.$$

We will use the notation

$$\boldsymbol{u}^* = \lim_{n\to\infty} \boldsymbol{u}(n)$$

to indicate that the sequence $\boldsymbol{u}(n)$ converges to $\boldsymbol{u}^*$. $\square$

Proposition 6.10. *A sequence of vectors*

$$\boldsymbol{u}(n) = \begin{bmatrix} u_1(n) \\ \vdots \\ u_N(n) \end{bmatrix} \in \mathbb{R}^N$$

converges to

$$\boldsymbol{u}^* = \begin{bmatrix} u_1^* \\ \vdots \\ u_N^* \end{bmatrix} \in \mathbb{R}^N$$

in the norm $|-|_\infty$ if and only if

$$\lim_{n\to\infty} u_k(n) = u_k^*, \quad \forall k = 1, \dots, N.$$

Proof. We have

$$\lim_{n\to\infty} |\boldsymbol{u}(n) - \boldsymbol{u}^*|_\infty \Longleftrightarrow \lim_{n\to\infty} \max_{1\leq k\leq N} |u_k(n) - u_k^*| = 0 \Longleftrightarrow \lim_{n\to\infty} |u_k(n) - u_k^*| = 0, \quad \forall k = 1, \dots, N.$$

□

Proposition 6.11. *Let $(\boldsymbol{U}, \|-\|)$ be a normed space. If the sequence $(\,\boldsymbol{u}(n)\,)_{n\geq 0}$ in $\boldsymbol{U}$ converges to $\boldsymbol{u}^* \in \boldsymbol{U}$ in the norm $\|-\|$, then*

$$\lim_{n\to\infty} \|\boldsymbol{u}(n)\| = \|\boldsymbol{u}^*\|.$$

Proof. We have

$$\big|\, \|\boldsymbol{u}(n)\| - \|\boldsymbol{u}^*\| \,\big| \stackrel{(6.1b)}{\leq} \|\boldsymbol{u}(n) - \boldsymbol{u}^*\| \to 0 \;\text{ as } n \to \infty.$$

□

The proof of the next result is left as an exercise.

Proposition 6.12. *Let $(\boldsymbol{U}, \|-\|)$ be a normed $\mathbb{F}$-vector space. If the sequences $(\,\boldsymbol{u}(n)\,)_{n\geq 0}$, $(\boldsymbol{v}(n))_{n\geq 0}$ converge in the norm $\|-\|$ to $\boldsymbol{u}^*$ and respectively $\boldsymbol{v}^*$, then their sum $\boldsymbol{u}(n) + \boldsymbol{v}(n)$ converges in the norm $\|-\|$ to $\boldsymbol{u}^* + \boldsymbol{v}^*$. Moreover, for any scalar $\lambda \in \mathbb{F}$, the sequence $\lambda\boldsymbol{u}(n)$ converges in the norm $\|-\|$ to the vector $\lambda\boldsymbol{u}^*$.* □

Proposition 6.13. *Let $\boldsymbol{U}$ be a vector space and $\|-\|_0$, $\|-\|_1$ be two norms on $\boldsymbol{U}$. The following statements are equivalent.*

(i) $\|-\|_0 \prec \|-\|_1$.
(ii) *If a sequence converges in the norm $\|-\|_1$, then it converges to the same limit in the norm $\|-\|_0$ as well.*

Proof. (i) ⇒ (ii). We know that there exists a constant $C > 0$ such that $\|\boldsymbol{u}\|_0 \leq C\|\boldsymbol{u}\|_1$, $\forall \boldsymbol{u} \in \boldsymbol{U}$. We have to show that if the sequence $(\,\boldsymbol{u}(n)\,)_{n\geq 0}$ converges in the norm $\|-\|_1$ to $\boldsymbol{u}^*$, then it also converges to $\boldsymbol{u}^*$ in the norm $\|-\|_0$. We have

$$\|\boldsymbol{u}(n) - \boldsymbol{u}^*\|_0 \leq C\|\boldsymbol{u}(n) - \boldsymbol{u}^*\|_1 \to 0 \;\text{ as } n \to \infty.$$

(ii) ⇒ (i). We argue by contradiction and we assume that for any $n > 0$ there exists $\boldsymbol{v}(n) \in \boldsymbol{U} \setminus 0$ such that

$$\|\boldsymbol{v}(n)\|_0 \geq n\|\boldsymbol{v}(n)\|_1. \tag{6.5}$$

Set

$$\boldsymbol{u}(n) := \frac{1}{\|\boldsymbol{v}(n)\|_0}\boldsymbol{v}(n), \quad \forall n > 0.$$

Multiplying both sides of (6.5) by $\frac{1}{\|\boldsymbol{v}(n)\|_0}$ we deduce

$$1 = \|\boldsymbol{u}(n)\|_0 \geq n\|\boldsymbol{u}(n)\|_1, \quad \forall n > 0.$$

Hence $\boldsymbol{u}(n) \to 0$ in the norm $\|-\|_1$ and thus $\boldsymbol{u}(n) \to 0$ in the norm $\|-\|_0$. Using Proposition 6.11 we deduce $\|\boldsymbol{u}(n)\|_0 \to 0$. This contradicts the fact that $\|\boldsymbol{u}(n)\|_0 = 1$ for any $n > 0$. □

Corollary 6.14. *Two norms on the same vector space are equivalent if and only if any sequence that converges in one norm converges to the same limit in the other norm as well.* □

Theorem 6.15. *Any norm $\|-\|$ on $\mathbb{R}^N$ is equivalent to the norm $|-|_\infty$.*

Proof. We already know that $\|-\| \leq |-\_\infty$ so it suffices to show that $|-|_\infty \prec \|-\|$. Consider the "sphere"

$$S := \big\{\boldsymbol{u} \in \mathbb{R}^N;\ \ |\boldsymbol{u}|_\infty = 1\big\},$$

We set

$$m := \inf_{\boldsymbol{u}\in S} \|\boldsymbol{u}\|.$$

We claim that

$$m > 0 \tag{6.6}$$

Let us observe that the above inequality implies that $|-|_\infty \prec \|-\|$. Indeed for any $\boldsymbol{u} \in \mathbb{R}^N \setminus 0$ we have

$$\bar{\boldsymbol{u}} := \frac{1}{|\boldsymbol{u}|_\infty}\boldsymbol{u} \in S$$

so that

$$\|\bar{\boldsymbol{u}}\| \geq m.$$

Multiplying both sides of the above inequality with $|\boldsymbol{u}|_\infty$ we deduce

$$\|\boldsymbol{u}\| \geq m|\boldsymbol{u}|_\infty \Rightarrow |\boldsymbol{u}|_\infty \leq \frac{1}{m}\|\boldsymbol{u}\|, \quad \forall \boldsymbol{u} \in \mathbb{R}^N \setminus 0 \Longleftrightarrow |-|_\infty \prec \|-\|.$$

Let us prove the claim (6.6). Choose a sequence $\boldsymbol{u}(n) \in S$ such that

$$\lim_{n\to\infty} \|\boldsymbol{u}(n)\| = m. \tag{6.7}$$

Since $\boldsymbol{u}(n) \in S$, coordinates $u_k(n)$ of $\boldsymbol{u}(n)$ satisfy the inequalities

$$u_k(n) \in [-1, 1], \quad \forall k = 1, \dots, N, \quad \forall n.$$

The Bolzano-Weierstrass theorem implies that we can extract a subsequence $(\,\boldsymbol{u}(\nu)\,)$ of the sequence $(\,\boldsymbol{u}(n)\,)$ such that sequence $(\,u_k(\nu)\,)$ converges to some $u_k^* \in \mathbb{R}$, for any $k = 1, \dots, N$. Let $\boldsymbol{u}^* \in \mathbb{R}^N$ be the vector with coordinates $u_1^*, \dots, u_N^*$. Using Proposition 6.10 we deduce that $\boldsymbol{u}(\nu) \to \boldsymbol{u}^*$ in the norm $|-|_\infty$. Since $|\boldsymbol{u}(\nu)|_\infty = 1, \forall \nu$, we deduce from Proposition 6.11 that $|\boldsymbol{u}^*|_\infty = 1$, i.e., $\boldsymbol{u}^* \in S$.

On the other hand $\|-\| \prec |-|_\infty$ and we deuce that $\boldsymbol{u}(\nu)$ converges to $\boldsymbol{u}^*$ in the norm $\|-\|$. Hence

$$\|\boldsymbol{u}^*\| = \lim_{\nu\to\infty} \|\boldsymbol{u}(\nu)\| \stackrel{(6.7)}{=} m.$$

Since $|\boldsymbol{u}^*| = 1$ we deduce that $\boldsymbol{u}^* \neq 0$ so that $m = \|\boldsymbol{u}^*\| > 0$. □

Corollary 6.16. *On a finite dimensional real vector space $\boldsymbol{U}$ any two norms are equivalent.*

Proof. Let $N = \dim \boldsymbol{U}$. We can then identify $\boldsymbol{U}$ with $\mathbb{R}^N$ and we can invoke Theorem 6.15 which implies that any two norms on $\mathbb{R}^N$ are equivalent. □

Corollary 6.17. *Let $\boldsymbol{U}$ be a finite dimensional real vector space. A sequence in $\boldsymbol{U}$ converges in some norm if and only if it converges to the same limit in any norm.* □

6.3. **Completeness.**

Definition 6.18. Let $(\boldsymbol{U}, \|-\|)$ be a normed space. A sequence $(\,\boldsymbol{u}(n)\,)_{n\geq 0}$ of vectors in $\boldsymbol{U}$ is said to be a *Cauchy sequence* (in the norm $\|-\|$) if

$$\lim_{m,n\to\infty} \|\boldsymbol{u}(m) - \boldsymbol{u}(n)\| = 0,$$

i.e.,

$$\forall \varepsilon > 0 \;\; \exists \nu = \nu(\varepsilon) > 0 \;\; \text{such that} \operatorname{dist}\big(\boldsymbol{u}(m), \boldsymbol{u}(n)\,\big) = \|\boldsymbol{u}(m) - \boldsymbol{u}(n)\| < \varepsilon, \;\; \forall m, n > \nu. \quad \square$$

Proposition 6.19. *Let $(\boldsymbol{U}, \|-\|)$ be a normed space. If the sequence $(\,\boldsymbol{u}(n)\,)$ is convergent in the norm $\|-\|$, then it is also Cauchy in the norm $\|-\|$.*

Proof. Denote by $\boldsymbol{u}^*$ the limit of the sequence $(\,\boldsymbol{u}(n)\,)$. For any $\varepsilon > 0$ we can find $\nu = \nu(\varepsilon) > 0$ such that

$$\|\boldsymbol{u}(n) - \boldsymbol{u}^*\| < \frac{\varepsilon}{2}.$$

If $m, n > \nu(\varepsilon)$ then

$$\|\boldsymbol{u}(m) - \boldsymbol{u}(n)\| \leq \|\boldsymbol{u}(m) - \boldsymbol{u}^*\| + \|\boldsymbol{u}^* - \boldsymbol{u}^n\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Definition 6.20. A normed space $(\boldsymbol{U}, \|-\|)$ is called *complete* if any Cauchy sequence is also convergent. A *Banach space* is a complete normed space. □

Proposition 6.21. *Let $\|-\|_0$ and $\|-\|_1$ be two equivalent norms on the same vector space $\boldsymbol{U}$. Then the following statements are equivalent.*

(i) *The space $\boldsymbol{U}$ is complete in the norm $\|-\|_0$.*
(ii) *The space $\boldsymbol{U}$ is complete in the norm $\|-\|_1$.*

Proof. We prove only that (i) ⇒ (ii). The opposite implication is completely similar. Thus, we know that $\boldsymbol{U}$ is $\|-\|_0$-complete and we have to prove that it is also $\|-\|_0$-complete.

Suppose that $(\,\boldsymbol{u}(n)\,)$ is a Cauchy sequence in the norm $\|-\|_1$. We have to prove that it converges in the norm $\|-\|_1$. Since $\|-\|_0 \prec \|-\|_1$ we deduce that there exists a constant $C > 0$ such that

$$\|\boldsymbol{u}(n) - \boldsymbol{u}(m)\|_0 \leq C\|\boldsymbol{u}(n) - \boldsymbol{u}(m)\|_1 \to 0 \;\; \text{as } m, n \to \infty.$$

Hence the sequence $(\,\boldsymbol{u}(n)\,)$ is also Cauchy in the norm $\|-\|_0$. Since $\boldsymbol{U}$ is $\|-\|_0$-complete, we deduce that the sequence $(\,\boldsymbol{u}(n)\,)$ converges in the norm $\|-\|_0$ to some vector $\boldsymbol{u}^* \in \boldsymbol{U}$.

Using the fact that $\|-\|_1 \prec \|-\|_0$ we deduce from Proposition 6.13 that the sequence converges to $\boldsymbol{u}^*$ in the norm $\|-\|_1$ as well. □

Theorem 6.22. *Any finite dimensional normed space $(\boldsymbol{U}, \|-\|)$ is complete.*

Proof. For simplicity we assume that $\boldsymbol{U}$ is a real vector space of dimension N. Thus we can identify it with $\mathbb{R}^N$. Invoking Proposition 6.21 we see that it suffices to prove that $\boldsymbol{U}$ complete with respect to any norm equivalent to $|-\|$. Invoking Theorem 6.15 we see that it suffices to prove that $\mathbb{R}^N$ is complete with respect to the norm $|-|_\infty$.

Suppose that $(\boldsymbol{u}(n))$ is a sequence in $\mathbb{R}^N$ which is Cauchy with respect to the norm $|-|_\infty$. As usual, we denote by $u_1(n),\dots,u_N(n)$ the coordinates of $\boldsymbol{u}(n)\in\mathbb{R}^N$. From the inequalities

$$|\,u_k(m)-u_k(n)|\leq|\,\boldsymbol{u}(m)-\boldsymbol{u}(n)\,|_\infty,\;\;\forall k=1,\dots,N,$$

we deduce that for any $k=1,\dots,N$ the sequence of coordinates $(\,u_k(n)\,)$ is a Cauchy sequence of real numbers. Therefore it converges to some real number u_k^*. Invoking Proposition 6.10 we deduce that the sequence $(\,\boldsymbol{u}(n)\,)$ converges to the vector $\boldsymbol{u}^*\in\mathbb{R}^N$ with coordinates $u_1^*,\dots,u_N^*$. □

6.4. **Continuous maps.** Let $(\boldsymbol{U},\|-\|_{\boldsymbol{U}})$ and $(\boldsymbol{V},\|-\|_{\boldsymbol{V}})$ be two normed vector spaces and $S\subset\boldsymbol{U}$.

Definition 6.23. A map $f:S\to\mathbb{V}$ is said to be continuous at a point $\boldsymbol{s}^*\in S$ if for any sequence $(\,\boldsymbol{s}(n)\,)$ of vectors in S that converges to $\boldsymbol{s}^*$ in the norm $\|-\|_{\boldsymbol{U}}$ the sequence $f(\boldsymbol{s}(n))$ of vectors in $\boldsymbol{V}$ converges in the norm $\|-\|_{\boldsymbol{V}}$ to the vector $f(\boldsymbol{s}^*)$. The function f is said to be *continuous* (on S) if it is continuous at any point $\boldsymbol{s}^*$ in S. □

Remark 6.24. The notion of continuity depends on the choice of norms on $\boldsymbol{U}$ and $\boldsymbol{V}$ because it relies on the notion of convergence in these norms. Hence, if we replace these norms with equivalent ones, the notion of continuity does not change. □

Example 6.25. (a) Let $(\boldsymbol{U},\|-\|)$ be a normed space. Proposition 6.11 implies that the function $f:\boldsymbol{U}\to\mathbb{R}$, $f(\boldsymbol{u})=\|\boldsymbol{u}\|$ is continuous.
(b) Suppose that $\boldsymbol{U},\boldsymbol{V}$ are two normed spaces, $S\subset\boldsymbol{U}$ and $f:S\to\boldsymbol{V}$ is a continuous map. Then the restriction of f to any subset $S'\subset\boldsymbol{U}$ is also acontinuous map $f|_{S'}:S'\to\boldsymbol{V}$. □

Proposition 6.26. *Let $\boldsymbol{U},\boldsymbol{V}$ and S as above, $f:S\to\boldsymbol{V}$ a map and $\boldsymbol{s}^*\in S$. Then the following statements are equivalent.*

(i) *The map f is continuous at $\boldsymbol{s}^*$.*
(ii) *For any $\varepsilon>0$ there exists $\delta=\delta(\varepsilon)>0$ such that if $\boldsymbol{s}\in S$ and $\|\boldsymbol{s}-\boldsymbol{s}^*\|_{\boldsymbol{U}}<\delta$, then $\|f(\boldsymbol{s})-f(\boldsymbol{s}^*)\|_{\boldsymbol{V}}<\varepsilon$.*

Proof. (i) $\Rightarrow$ (ii) We argue by contradiction and we assume that there exists $\varepsilon_0>0$ such that for any $n>0$ there exists $\boldsymbol{s}_n\in S$ such that

$$\|\boldsymbol{s}_n-\boldsymbol{s}^*\|_{\boldsymbol{U}}<\frac{1}{n}\;\text{ and }\;\|f(\boldsymbol{s}_n)-f(\boldsymbol{s}^*)\|_{\boldsymbol{V}}\geq\varepsilon_0.$$

From the first inequality above we deduce that the sequence $\boldsymbol{s}_n$ converges to $\boldsymbol{s}^*$ in the norm $\|-\|_{\boldsymbol{U}}$. The continuity of f implies that

$$\lim_{n\to\infty}\|f(\boldsymbol{s}_n)-f(\boldsymbol{s}^*)\|_{\boldsymbol{V}}=0.$$

The last equality contradicts the fact that $\|f(\boldsymbol{s}_n)-f(\boldsymbol{s}^*)\|_{\boldsymbol{V}}\geq\varepsilon_0$ for any n.
(ii) $\Rightarrow$ (i) Let $\boldsymbol{s}(n)$ be a sequence in S that converges to $\boldsymbol{s}^*$. We have to show that $f(\,\boldsymbol{s}(n)\,)$ converges to $f(\boldsymbol{s}^*)$.

Let $\varepsilon>0$. By (ii), there exists $\delta(\varepsilon)>0$ such that if $\boldsymbol{s}\in S$ and $\|\boldsymbol{s}-\boldsymbol{s}^*\|_{\boldsymbol{V}}<\delta(\varepsilon)$, then $\|f(\boldsymbol{s})-f(\boldsymbol{s}^*)\|_{\boldsymbol{V}}<\varepsilon$. Since $\boldsymbol{s}(n)$ converges to $\boldsymbol{s}^*$, we can find $\nu=\nu(\varepsilon)$ such that if $n\geq\nu(\varepsilon)$, then

$\|\boldsymbol{s}(n)-\boldsymbol{s}^*\|_{\boldsymbol{V}}<\delta(\varepsilon)$. In particular, for any $n\geq\nu(\varepsilon)$ we have

$$\|f(\,\boldsymbol{s}(n)\,)-f(\boldsymbol{s}^*)\|_{\boldsymbol{V}}<\varepsilon.$$

This proves that $f(\,\boldsymbol{s}(n)\,)$ converges to $f(\boldsymbol{s}^*)$. □

Definition 6.27. A map $f:S\to\boldsymbol{V}$ is said to be *Lipschitz* if there exists $L>0$ such that

$$\|f(\boldsymbol{s}_1)-f(\boldsymbol{s}_2)\|_{\boldsymbol{V}}\leq L\|\boldsymbol{s}_1-\boldsymbol{s}_2\|_{\boldsymbol{U}},\ \ \forall\boldsymbol{s}_1,\boldsymbol{s}_2\in S.$$

Observe that if $f:S\to\boldsymbol{V}$ is a Lipschitz map, then

$$\sup\left\{\frac{|f(\boldsymbol{s}_1)-f(\boldsymbol{s}_2)\|_{\boldsymbol{V}}}{\|\boldsymbol{s}_1-\boldsymbol{s}_2\|_{\boldsymbol{U}}};\ \ \boldsymbol{s}_1,\boldsymbol{s}_2\in S;\ \ \boldsymbol{s}_1\neq\boldsymbol{s}_2\right\}<\infty.$$

The above supremum is called the *Lipschitz* constant of f.

Proposition 6.28. *A Lipschitz map $f:S\to\boldsymbol{V}$ is continuous.*

Proof. Denote by L the Lipschitz constant of f. Let $\boldsymbol{s}^*\in S$ and $\boldsymbol{s}(n)$ a sequence in S that converges to $\boldsymbol{s}^*$.

$$\|f(\boldsymbol{s}(n))-f(\boldsymbol{s}^*)\|_{\boldsymbol{V}}\leq L\|\boldsymbol{s}(n)-\boldsymbol{s}^*\|_{\boldsymbol{U}}\to 0\ \text{ as } n\to\infty.$$

□

Proposition 6.29. *Let $(\boldsymbol{U},\|-\|_{\boldsymbol{U}})$ and $(\boldsymbol{V},\|-\|_{\boldsymbol{V}})$ be two normed vector spaces and $T:\boldsymbol{U}\to\boldsymbol{V}$ a linear map. Then the following statements are equivalent.*

(i) *The map $T:\boldsymbol{U}\to\boldsymbol{V}$ is continuous on $\boldsymbol{U}$.*
(ii) *There exists $C>0$ such that*

$$\|T\boldsymbol{u}\|_{\boldsymbol{V}}\leq C\|\boldsymbol{u}\|_{\boldsymbol{U}},\ \ \forall\boldsymbol{u}\in\boldsymbol{U}. \tag{6.8}$$

(iii) *The map $T:\boldsymbol{U}\to\boldsymbol{V}$ is Lipschitz.*

Proof. (i) ⇒ (ii) We argue by contradiction. We assume that for any positive integer n there exists a vector $\boldsymbol{u}_n\in\boldsymbol{U}$ such that

$$\|T\boldsymbol{u}_n\|_{\boldsymbol{V}}\geq n\|\boldsymbol{u}_n\|_{\boldsymbol{U}}.$$

Set

$$\bar{\boldsymbol{u}}_n:=\frac{1}{n\|\boldsymbol{u}(n)\|_{\boldsymbol{U}}}\boldsymbol{u}(n).$$

Then

$$\|\bar{\boldsymbol{u}}_n\|_{\boldsymbol{U}}=\frac{1}{n}, \tag{6.9a}$$

$$\|T\bar{\boldsymbol{u}}_n\|_{\boldsymbol{V}}\geq n\|\bar{\boldsymbol{u}}_n\|_{\boldsymbol{U}}=1. \tag{6.9b}$$

The equality (6.9a) implies that $\bar{\boldsymbol{u}}_n\to 0$. Since T is continuous we deduce $\|T\bar{\boldsymbol{u}}_n\|_{\boldsymbol{V}}\to 0$. This contradicts (6.9b).
(ii) ⇒ (iii) For any $\boldsymbol{u}_1,\boldsymbol{u}_2\in\boldsymbol{U}$ we have

$$\|T\boldsymbol{u}_1-T\boldsymbol{u}_2\|_{\boldsymbol{V}}=\|T(\boldsymbol{u}_1-\boldsymbol{u}_2)\|_{\boldsymbol{V}}\leq C\|\boldsymbol{u}_1-\boldsymbol{u}_2\|_{\boldsymbol{U}}.$$

This proves that T is Lipschitz. The implication (iii) ⇒ (i) follows from Proposition 6.28. □

Definition 6.30. Let $(\boldsymbol{U}, \|-\|_{\boldsymbol{U}})$ and $(\boldsymbol{V}, \|-\|_{\boldsymbol{V}})$ be two normed vector spaces and $T : \boldsymbol{U} \to \boldsymbol{V}$ a linear map. The *norm* of T, denoted by $\|T\|$ or $\|T\|_{\boldsymbol{U},\boldsymbol{V}}$, is the infimum of all constants $C > 0$ such that (6.8) is satisfied. If there is no such C we set $\|T\| := \infty$. Equivalently

$$\|T\| = \sup_{\boldsymbol{u}\in\boldsymbol{U}\setminus 0} \frac{\|T\boldsymbol{u}\|_{\boldsymbol{V}}}{\|\boldsymbol{u}\|_{\boldsymbol{U}}} = \sup_{\|\boldsymbol{u}\|_{\boldsymbol{U}}=1} \|T\boldsymbol{u}\|_{\boldsymbol{V}}.$$

We denote by $\mathcal{B}(\boldsymbol{U}, \boldsymbol{V})$ the space of linear operators such that that $\|T\| < \infty$, and we will refer to such operators as *bounded*. When $\boldsymbol{U} = \boldsymbol{V}$ and $\|-\|_{\boldsymbol{U}} = \|-\|_{\boldsymbol{V}}$ we set

$$\mathcal{B}(\boldsymbol{U}) := \mathcal{B}(\boldsymbol{U}, \boldsymbol{U}).$$ □

Observe that we can rephrase Proposition 6.29 as follows.

Corollary 6.31. *Let $(\boldsymbol{U}, \|-\|_{\boldsymbol{U}})$ and $(\boldsymbol{V}, \|-\|_{\boldsymbol{V}})$ be two normed vector spaces. A linear map $T : \boldsymbol{U} \to \boldsymbol{V}$ is continuous if and only if it is bounded.* □

Let us observe that if $T : \boldsymbol{U} \to \boldsymbol{V}$ is a continuous operator, then

$$\|T\| \leq C \Longleftrightarrow \|T\boldsymbol{u}\|_{\boldsymbol{V}} \leq C\|\boldsymbol{u}\|_{\boldsymbol{U}}, \ \ \forall \boldsymbol{u} \in \boldsymbol{U}. \tag{6.10}$$

We have the following useful result whose proof is left as an exercise.

Proposition 6.32. *Let $(\boldsymbol{U}, \|-\|)$ be a normed $\mathbb{F}$-vector space and consider the space $\mathcal{B}(\boldsymbol{U})$ of continuous linear operators $\boldsymbol{U} \to \boldsymbol{U}$. If $S, T \in \mathcal{B}(\boldsymbol{U})$ and $\lambda \in \mathbb{F}$, then*

$$\|\lambda S\| = |\lambda| \cdot \|S\|, \ \ \|S + T\| \leq \|S\| + \|T\|, \ \ \|S \circ T\| \leq \|S\| \cdot \|T\|.$$

In particular, the map $\mathcal{B}(\boldsymbol{U}) \ni T \mapsto \|T\| \in [0, \infty)$ is a norm on the vector space of bounded linear operators $\boldsymbol{U} \to \boldsymbol{U}$. For $T \in \mathcal{B}(\boldsymbol{U})$ the quantity $\|T\|$ is called the operator norm *of T.* □

Corollary 6.33. *If (S_n) and (T_n) are sequences in $\mathcal{B}(\boldsymbol{U})$ which converge in the operator norm to $S \in \mathcal{B}(\boldsymbol{U})$ and respectively $T \in \mathcal{B}(U)$, then the sequence (S_nT_n) converges in the operator norm to ST.*

Proof. We have

$$\begin{aligned}\|S_nT_n - ST\| &= \|(S_nT_n - ST_n) + (ST_n - ST)\| \leq \|(S_n - S)T_n\| + \|S(T_n - T)\| \\ &\leq \|(S_n - S)\| \cdot \|T_n\| + \|S\| \cdot \|(T_n - T)\|.\end{aligned}$$

Observe that

$$\|(S_n - S)\| \to 0, \ \ \|T_n\| \to \|T\|, \|S\| \cdot \|(T_n - T)\| \to 0.$$

Hence

$$\|(S_n - S)\| \cdot \|T_n\| + \|S\| \cdot \|(T_n - T)\| \to 0,$$

and thus $\|S_nT_n - ST\| \to 0$. □

Theorem 6.34. *Let $(\boldsymbol{U}, \|-\|_{\boldsymbol{U}})$ and $(\boldsymbol{V}, \|-\|_{\boldsymbol{V}})$ be two* finite dimensional *normed vector spaces. Then any linear map $T : \boldsymbol{U} \to \boldsymbol{V}$ is continuous.*

Proof. For simplicity we assume that the vector spaces are real vector spaces, $N = \dim \boldsymbol{U}$, $M = \dim \boldsymbol{V}$. By choosing bases in $\boldsymbol{U}$ and $\boldsymbol{V}$ we can identify them with the standard spaces $\boldsymbol{U} \cong \mathbb{R}^N$, $\boldsymbol{V} \cong \mathbb{R}^M$.

The notion of continuity is only based on the notion of convergence of sequences and as we know, on finite dimensional spaces the notion of convergence of sequences is independent of the norm used. Thus we can assume that the norm on $\boldsymbol{U}$ and $\boldsymbol{V}$ is $|-|_\infty$.

The operator T can be represented by a $M \times N$ matrix

$$A = (a_{ij})_{1 \leq i \leq m,\, 1 \leq j \leq N}.$$

Set

$$C := \sum_{j,k} |a_{kj}|.$$

In other words, C is the sum of the absolute values of all the entries of A. If

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix} \in \mathbb{R}^N$$

then the vector $\boldsymbol{v} = T\boldsymbol{u} \in \mathbb{R}^M$ has coordinates

$$v_k = \sum_{j=1}^{N} a_{kj} u_j, \quad k = 1, \dots, M.$$

We deduce that that for any $k = 1, \dots, M$ we have

$$|v_k| \leq \sum_{j=1}^{N} |a_{kj}| \cdot |u_j| \leq \max_{1 \leq j \leq N} |u_j| \sum_{j=1}^{N} |a_{kj}| = |\boldsymbol{u}|_\infty \sum_{j=1}^{N} |a_{kj}| \leq C|\boldsymbol{u}|_\infty.$$

Hence, for any $\boldsymbol{u} \in \mathbb{R}^N$ we

$$|T\boldsymbol{u}|_\infty = |\boldsymbol{v}|_\infty = \max_{1 \leq k \leq M} |v_k| \leq C|\boldsymbol{u}|_\infty.$$

This proves that T is continuous. □

6.5. **Series in normed spaces.** Let $(\boldsymbol{U}, \| - \|)$ be a normed space.

Definition 6.35. A *series* in $\boldsymbol{U}$ is a *formal* infinite sum of the form

$$\sum_{n \geq 0} \boldsymbol{u}(n) = \boldsymbol{u}(0) + \boldsymbol{u}(1) + \boldsymbol{u}(2) + \cdots,$$

where $\boldsymbol{u}(n)$ is a vector in $\boldsymbol{U}$ for any $n \geq 0$. The n-th *partial sum* of the series is the finite sum

$$S_n = \boldsymbol{u}(0) + \boldsymbol{u}(1) + \cdots + \boldsymbol{u}(n).$$

The series is called *convergent* if the sequence S_n is convergent. The limit of this sequence is called the *sum* of the series. The series is called *absolutely* convergent if the series of nonnegative real numbers

$$\sum_{n \geq 0} \|\boldsymbol{u}(n)\|$$

is convergent. □

Proposition 6.36. *Let $(\boldsymbol{U}, \| - \|)$ be a Banach space, i.e., a* complete *normed space. If the series in $\boldsymbol{U}$*

$$\sum_{n \geq 0} \boldsymbol{u}(n) \tag{6.11}$$

is absolutely convergent, then it is also convergent. Moreover

$$\| \sum_{n \geq 0} \boldsymbol{u}(n) \| \leq \sum_{\geq 0} \|\boldsymbol{u}(n)\|. \tag{6.12}$$

Proof. Denote by S_n the n-th partial sum of the series (6.11), and by S_n^+ the n-th partial sum of the series

$$\sum_{n\geq 0}\|\boldsymbol{u}(n)\|.$$

Since $\boldsymbol{U}$ is complete, it suffices to show that the sequence (S_n) is Cauchy. For $m<n$ we have

$$\|S_n - S_m\| = \|\boldsymbol{u}(m+1)+\cdots+\boldsymbol{u}(n)\| \leq \|\boldsymbol{u}(m+1)\|+\cdots+\|\boldsymbol{u}(n)\| = S_n^+ - S_m^+.$$

Since the sequence (S_n^+) is convergent, we deduce that it also Cauchy so that

$$\lim_{m,n\to\infty}(S_n^+ - S_m^+) = 0.$$

This forces

$$\lim_{m,n\to\infty}\|S_n - S_m\| = 0.$$

The inequality (6.12) is obtained by letting $n\to\infty$ in the inequality

$$\|\boldsymbol{u}(0)+\cdots+\boldsymbol{u}(n)\| \leq \|\boldsymbol{u}(0)\|+\cdots+\|\boldsymbol{u}(n)\|.$$

□

Corollary 6.37 (Comparison trick)**.** *Let*

$$\sum_{n\geq 0}\boldsymbol{u}(n)$$

be a series in the Banach space $(\boldsymbol{U}, \|-\|)$*. If there exists a convergent series of positive real numbers*

$$\sum_{n\geq 0}c_n,$$

such that

$$\|\boldsymbol{u}(n)\| \leq c_n,\ \ \forall n\geq 0, \tag{6.13}$$

then the series $\sum_{n\geq 0}\boldsymbol{u}(n)$ *is absolutely convergent and thus convergent.*

Proof. The inequality (6.13) implies that the series of nonnegative real numbers $\sum_{n\geq 0}\|\boldsymbol{u}(n)\|$ is convergent. □

6.6. **The exponential of a matrix.** Suppose that A is an $N\times N$ complex matrix. We regard it as a linear operator $A:\mathbb{C}^n\to\mathbb{C}^n$. As such, it is a continuous map, in any norm we choose on $\mathbb{C}^n$. For simplicity, we will work with the Euclidean norm

$$|\boldsymbol{u}| = |\boldsymbol{u}|_2 = \sqrt{|u_1|^2+\cdots+|u_N|^2}.$$

The operator norm of A is then the quantity

$$\|A\| = \sup_{\boldsymbol{u}\in\mathbb{C}^N\setminus 0}\frac{|A\boldsymbol{u}|}{|\boldsymbol{u}|}.$$

We denote by $\mathcal{B}_N$ the vector space of continuous linear operators $\mathbb{C}^N\to\mathbb{C}^N$ equipped with the operator norm defined above. The space is a finite dimensional complex vector space ($\dim_{\mathbb{C}}\mathcal{B}_N = N^2$) and thus it is complete. Consider the series in $\mathcal{B}_N$

$$\sum_{n\geq 0}\frac{1}{n!}A^n = \mathbb{1} + \frac{1}{1!}A + \frac{1}{2!}A^2 + \cdots .$$

Observe that

$$\|A^n\| \leq \|A\|^n$$

so that

$$\frac{1}{n!}\|A^n\| \leq \frac{1}{n!}\|A\|^n.$$

The series of nonnegative numbers

$$\sum_{n\geq 0}\frac{1}{n!}\|A\|^n = 1 + \frac{1}{1!}\|A\| + \frac{1}{2!}\|A\|^2 + \cdots ,$$

is $e^{\|A\|}$ Thus the series $\sum_{n\geq 0}\frac{1}{n!}A^n$ is absolutely convergent and hence convergent.

Definition 6.38. For any complex $N \times N$ matrix A we denote by e^A the sum of the convergent (and absolutely convergent) series

$$\mathbb{1} + \frac{1}{1!}A + \frac{1}{2!}A^2 + \cdots .$$

The $N \times N$ matrix A is called the *exponential of* A. □

Note that

$$\|e^A\| \leq e^{\|A\|}.$$

Example 6.39. Suppose that A is a diagonal matrix,

$$A := \mathrm{Diag}(\lambda_1, \dots, \lambda_N).$$

Then

$$A^n = \mathrm{Diag}(\lambda_1^n, \dots, \lambda_N^n)$$

and thus

$$e^A = \mathrm{Diag}\left(\sum_{n\geq 0}\frac{1}{n!}\lambda_1^n, \dots, \sum_{n\geq 0}\frac{1}{n!}\lambda_N^n\right) = \mathrm{Diag}\big(e^{\lambda_1}, \dots, e^{\lambda_N}\big).$$ □

Before we explain the uses of the exponential of of a matrix, we describe a simple strategy for computing it.

Proposition 6.40. *Let A be a complex $N \times N$ matrix and S and invertible $N \times N$ matrix. Then*

$$S^{-1}e^AS = e^{S^{-1}AS}. \tag{6.14}$$

Proof. Set $A_S := S^{-1}AS$. We have to prove that

$$S^{-1}e^AS = e^{A_S}.$$

Set

$$S_n := \mathbb{1} + \frac{1}{1!}A + \cdots + \frac{1}{n!}A^n,\ \ S_n' = \mathbb{1} + \frac{1}{1!}A_S + \cdots + \frac{1}{n!}A_S^n.$$

Observe that that for any positive integer k

$$A_S^k = \underbrace{(S^{-1}AS)(S^{-1}AS)\cdots(S^{-1}AS)}_{k} = S^{-1}A^kS.$$

Moreover, for any two $N \times N$ matrices B, C we have

$$S^{-1}(B+C)S = A^{-1}BS + S^{-1}CS.$$

We deduce that

$$S^{-1}S_nS = S_n',\ \ \forall n \geq 0.$$

Since $S_n \to e^A$, we deduce from Corollary 6.33 that

$$S_n' = S^{-1}S_nS \to S^{-1}e^AS.$$

On the other hand $S_n' \to e^{A_S}$ and this implies (6.16). $\square$

We can rewrite (6.16) in the form

$$e^A = Se^{S^{-1}AS}S^{-1}. \tag{6.15}$$

If we can choose S cleverly so that $S^{-1}AS$ is not too complicated, then we have a good shot at computing $e^{S^{-1}AS}$, which then leads via (6.15) to a description of e^A.

Here is one such instance. Suppose that A is normal, i.e., $A^*A = AA^*$. The spectral theorem for normal operators then implies that there exists an invertible matrix S such that

$$S^{-1}AS = \mathrm{Diag}(\lambda_1, \dots, \lambda_N),$$

where $\lambda_1, \dots, \lambda_N \in \mathbb{C}$ are the eigenvalues of A. Moreover the matrix S also has an explicit description: its j-th column is an eigenvector of Euclidean norm 1 of A corresponding to the eigenvalue λ_j. The computations in Example 6.39) then lead to the formula

$$e^A = S^{-1}\,\mathrm{Diag}(e^{\lambda_1}, \dots, e^{\lambda_N})S. \tag{6.16}$$

6.7. **The exponential of a matrix and systems of linear differential equations.** The usual exponential function e^x satisfies the very important property

$$e^{x+y} = e^xe^y, \ \ \forall x, y.$$

If A, B are two $N \times N$ matrices, then the equality

$$e^{A+B} = e^Ae^B$$

no longer holds because the multiplication of matrices is non commutative. Something weaker does hold.

Theorem 6.41. *Suppose that A, B are complex $N \times N$ matrices that commute, i.e., $AB = BA$. Then*

$$e^{A+B} = e^Ae^B.$$

Proof. The proof relies on the following generalization of Newton's binomial formula whose proof is left as an exercise.

Lemma 6.42 (Newton's binomial formula: the matrix case)**.** *If A, B are two commuting $N \times N$ matrices, then for any positive integer n we have*

$$(A+B)^n = \sum_{k=0}^n \binom{n}{k}A^{n-k}B^k = \binom{n}{0}A^n + \binom{n}{1}A^{n-1}B + \binom{n}{2}A^{n-2}B^2 + \cdots,$$

where $\binom{n}{k}$ is the binomial coefficient

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}, \ \ \forall 0 \le k \le n.$$ $\square$

We set

$$S_n(A) = \mathbb{1} + \frac{1}{1!}A + \cdots + \frac{1}{n!}A^n, \;\; S_n(B) = \mathbb{1} + \frac{1}{1!}B + \cdots + \frac{1}{n!}B^n,$$
$$S_n(A+B) = \mathbb{1} + \frac{1}{1!}(A+B) + \cdots + \frac{1}{n!}(A+B)^n.$$

We have

$$S_n(A) \to e^A, \;\; S_n(B) \to e^B, \;\; S_n(A+B) \to e^{A+B} \;\; \text{as } n \to \infty.$$

and we will prove that

$$\lim_{n\to\infty} \|S_n(A)S_n(B) - S_n(A+B)\| = 0,$$

which then implies immediately the desired conclusion. Observe that

$$S_n(A)S_n(B) = \left(\mathbb{1} + \frac{1}{1!}A + \cdots + \frac{1}{n!}A^n \right) \left(\mathbb{1} + \frac{1}{1!}B + \cdots + \frac{1}{n!}B^n \right)$$
$$= \sum_{k=0}^{n} \left(\sum_{i=0}^{k} \frac{1}{i!}A^i \frac{1}{(k-i)!}A^iB^{k-i} \right) + \sum_{k=n+1}^{2n} \left(\sum_{i=0}^{n} \frac{1}{i!}A^i \frac{1}{(k-i)!}A^iB^{k-i} \right)$$
$$= \sum_{k=0}^{n} \frac{1}{k!} \underbrace{\left(\sum_{i=1}^{k} \frac{k!}{i!(k-i)!}A^iB^{k-i} \right)}_{=(A+B)^k} + \sum_{k=n+1}^{2n} \frac{1}{k!} \left(\sum_{i=1}^{n} \frac{k!}{i!(k-i)!}A^iB^{k-i} \right)$$
$$= \sum_{k=0}^{n} \frac{1}{k!}(A+B)^k + \sum_{k=n+1}^{2n} \frac{1}{k!} \left(\sum_{i=1}^{n} \frac{k!}{i!(k-i)!}A^iB^{k-i} \right)$$

Hence

$$\underbrace{S_n(A)S_n(B) - S_n(A+B)}_{=:R_n} = \sum_{k=n+1}^{2n} \frac{1}{k!} \left(\sum_{i=1}^{n} \frac{k!}{i!(k-i)!}A^iB^{k-i} \right)$$

and we deduce

$$\|R_n\| \le \sum_{k=n+1}^{2n} \frac{1}{k!} \left(\sum_{i=1}^{n} \frac{k!}{i!(k-i)!}\|A^iB^{k-i}\| \right) \le \sum_{k=n+1}^{2n} \frac{1}{k!} \left(\sum_{i=1}^{n} \frac{k!}{i!(k-i)!}\|A\|^i\|B\|^{k-i}\| \right)$$
$$\le \sum_{k=n+1}^{2n} \frac{1}{k!} \underbrace{\left(\sum_{i=1}^{k} \frac{k!}{i!(k-i)!}\|A\|^i\|B\|^{k-i}\| \right)}_{(\|A\|+\|B\|)^k} = \sum_{k=n+1}^{2n} \frac{1}{k!}(\|A\|+\|B\|)^k.$$

Set $C := \|A\| + \|B\|$ so that

$$\|R_n\| \le \sum_{k=n+1}^{2n} \frac{1}{k!}C^k. \tag{6.17}$$

Fix an positive integer n_0 such that

$$n_0 > 2C.$$

For $k > n_0$ we have

$$k! = 1 \cdot 2 \cdots n_0 \cdot (n_0+1) \cdots k \ge n_0^{k-n_0}$$

so that

$$\frac{1}{k!} \le \frac{1}{n_0^{k-n_0}}, \;\; \forall k > n_0.$$

We deduce that if $n > n_0$, then

$$\sum_{k=n+1}^{2n}\frac{1}{k!}C^k = C^{n_0}\sum_{k=n+1}^{2n}\frac{1}{k!}C^{k-n_0} \leq C_0^n\sum_{k=n+1}^{2n}\left(\frac{C}{n_0}\right)^{k-n_0}$$

$$\leq C^{n_0}\sum_{k=n+1}^{2n}\frac{1}{2^{k-n_0}} = (2C)^{n_0}\sum_{k=n+1}^{2n}\frac{1}{2^k} = (2C)^{n_0}\left(\frac{1}{2^{n+1}}+\cdots+\frac{1}{2^{2n}}\right)$$

$$= \frac{(2C)^{n_0}}{2^{n+1}}\underbrace{\left(1+\frac{1}{2}+\cdots+\frac{1}{2^{n-1}}\right)}_{\leq 2} \leq 2\frac{(2C)^{n_0}}{2^{n+1}} = \frac{(2C)^{n_0}}{2^n}$$

Using (6.17) we deduce

$$\|R_n\| \leq \frac{(2C)^{n_0}}{2^n} \to 0 \quad \text{as } n\to\infty.$$

□

6.8. **Closed and open subsets.**

Definition 6.43. (a) Let $\boldsymbol{U}$ be a vector space. A subset $\mathscr{C}\subset\boldsymbol{U}$ is said to be *closed* in the norm $\|-\|$ on $\boldsymbol{U}$ if for any sequence $(\,\boldsymbol{u}(n)\,)$ of vectors in $\mathscr{C}$ which converges in the norm $\|-\|$ to a vector $\boldsymbol{u}^*\in\boldsymbol{U}$ we have $\boldsymbol{u}^*\in\mathscr{C}$.
(b) A subset $\mathcal{O}$ in $\boldsymbol{U}$ is said to be *open* in the norm $\|-\|$ if the complement $U\setminus\mathcal{O}$ is closed in the norm $\|-\|$. □

Remark 6.44. The notion of closed set with respect to a norm is based only on the notion of convergence with respect to a norm. Thus, if a set is closed in a given norm, it is closed in any equivalent. In particular, the notion of closed set in a finite dimensional normed space is independent of the norm used. A similar observation applies to open sets. □

Proposition 6.45. *Suppose that $(\boldsymbol{U},\|-\|)$ is a normed vector space and $\mathcal{O}\subset\boldsymbol{U}$. Then the following statements are equivalent.*

(i) *The set $\mathcal{O}$ is open in the norm $\|-\|$.*
(ii) *For any $\boldsymbol{u}\in\mathcal{O}$ there exists $\varepsilon>0$ such that any $\boldsymbol{v}\in\boldsymbol{U}$ satisfying $\|\boldsymbol{v}-\boldsymbol{u}\|<\varepsilon$ belongs to $\mathcal{O}$.*

□

Proof. (i) ⇒ (ii) We argue by contradiction. We assume that there exists $\boldsymbol{u}^*\in\mathcal{O}$ so that for any $n>0$ there exists $\boldsymbol{v}_n\in\boldsymbol{U}$ satisfying

$$\|\boldsymbol{v}_n-\boldsymbol{u}^*\|<\frac{1}{n} \ \text{ but } \ \boldsymbol{v}_n\notin\mathcal{O}.$$

Set $\mathscr{C}$ so that $\mathscr{C}$ is closed and $\boldsymbol{v}_n\in\mathscr{C}$ for any $n>0$. The inequality $\|\boldsymbol{v}_n-\boldsymbol{u}^*\|<\frac{1}{n}$ implies $\boldsymbol{v}_n\to\boldsymbol{u}^*$. Since $\mathscr{C}$ is closed we deduce that $\boldsymbol{u}^*\in\mathscr{C}$. This contradicts the initial assumption that $\boldsymbol{u}^*\in\mathcal{O}=$
$bsU\setminus\mathscr{C}$.

(ii) ⇒ (i) We have to prove that $\mathscr{C}=\boldsymbol{U}\setminus\mathcal{O}$ is closed, given that $\mathcal{O}$ satisfies (ii). Again we argue by contradiction. Suppose that there exists sequence $(\,\boldsymbol{u}(n)\,)$ in $\mathscr{C}$ which converges to a vector $\boldsymbol{u}^*\in\boldsymbol{U}\setminus\mathscr{C}=\mathcal{O}$. Thus there exists $\varepsilon>0$ such that any $\boldsymbol{v}\in\boldsymbol{U}$ satisfying $\|\boldsymbol{v}-\boldsymbol{u}^*\|$ does not belong to $\mathscr{C}$. This contradicts the fact that $\boldsymbol{u}(n)\to\boldsymbol{u}^*$ because at least one of the terms $\boldsymbol{u}(n)$ satisfies $\|\boldsymbol{u}(n)-\boldsymbol{u}^*\|<\varepsilon$, and yet it belongs to $\mathscr{C}$. □

Definition 6.46. Let $(\boldsymbol{U}, \|-\|)$ be an normed space. The *open ball* of center $\boldsymbol{u}$ and radius r is the set

$$B(\boldsymbol{u}, r) := \{\boldsymbol{v} \in \boldsymbol{U};\ \ \|\boldsymbol{v} - \boldsymbol{u}\| < r\}. \qquad \square$$

We can rephrase Proposition 6.45 as follows.

Corollary 6.47. *Suppose that $(\boldsymbol{U}, \|-\|)$ is a normed vector space. A subset $\mathcal{O} \subset \boldsymbol{U}$ is open in the norm $\|-\|$ if and only if for any $\boldsymbol{u} \in \mathcal{O}$ there exists $\varepsilon > 0$ such that $B(\boldsymbol{u}, \varepsilon) \subset \mathcal{O}$.* $\square$

The proof of the following result is left as an exercise.

Theorem 6.48. *Suppose that $(\boldsymbol{U}, \|-\|_{\boldsymbol{U}})$, $(\boldsymbol{V}, \|-\|_{\boldsymbol{V}})$ are normed spaces and $f : \boldsymbol{U} \to \boldsymbol{V}$ is a map. Then the following statements are equivalent.*

(i) *The map f is continuous.*
(ii) *For any subset $\mathcal{O} \subset \boldsymbol{V}$ that is open in the norm $\|-\|_{\boldsymbol{V}}$ the pre image $f^{-1}(\mathcal{O}) \subset \boldsymbol{U}$ is open in the norm $\|-\|_{\boldsymbol{V}}$.*
(iii) *For any subset $\mathscr{C} \subset \boldsymbol{V}$ that is closed in the norm $\|-\|_{\boldsymbol{V}}$ the pre image $f^{-1}(\mathscr{C}) \subset \boldsymbol{U}$ is closed in the norm $\|-\|_{\boldsymbol{V}}$.* $\square$

6.9. **Compactness.** We want to discuss a central concept in modern mathematics which is often used proving existence results.

Definition 6.49. Let $(\boldsymbol{U}, \|-\|)$ be a normed space. A subset $K \subset \boldsymbol{U}$ is called *compact* if any sequence of vectors in K has a subsequence that converges to some vector in K. $\square$

Remark 6.50. Since the notion of compactness is expressed only in terms of convergence of sequences we deduce that if a set is compact in some norm, it is compact in any equivalent norm. In particular, on finite dimensional vector spaces the notion of compactness is independent of the norm. $\square$

Example 6.51 (Fundamental Example)**.** For any real numbers $a < b$ the closed interval $[a, b]$ is a compact subset of $\mathbb{R}$. $\square$

The next existence result illustrates our claim of the usefulness of compactness in establishing existence results.

Theorem 6.52 (Weierstrass)**.** *Let $(\boldsymbol{U}, \|-\|)$ be a normed space and $f : K \to \mathbb{R}$ a continuous function defined on the* compact *subset $K \subset \boldsymbol{U}$. Then* there exist *$\boldsymbol{u}_*, \boldsymbol{u}^* \in K$ such that*

$$f(\boldsymbol{u}_*) \leq f(\boldsymbol{u}) \leq f(\boldsymbol{u}^*),\ \ \forall \boldsymbol{u} \in K,$$

i.e.,

$$f(\boldsymbol{u}_*) = \inf_{\boldsymbol{u} \in K} f(\boldsymbol{u}),\ \ f(\boldsymbol{u}^*) = \sup_{\boldsymbol{u} \in K} f(\boldsymbol{u}).$$

In particular, f is bounded both from above and from below.

Proof. Let

$$m := \inf_{\boldsymbol{u} \in K} f(\boldsymbol{u}) \in [-\infty, \infty).$$

There exists a sequence $(\boldsymbol{u}(n))$ in K such that[4]

$$\lim_{n \to \infty} f(\boldsymbol{u}(n)) = m.$$

[4] Such sequence is called a *minimizing sequence* of f.

Since K is compact, there exists a subsequence $(\,\boldsymbol{u}(\nu)\,)$ of $(\,\boldsymbol{u}(n)\,)$ that converges to some $\boldsymbol{u}_* \in K$. Observing that

$$\lim_{\nu\to\infty} f(\,\boldsymbol{u}(\nu)\,) = \lim_{n\to\infty} f(\,\boldsymbol{u}(n)\,) = m$$

we deduce from the continuity of f that

$$f(\boldsymbol{u}_*) = \lim_{\nu\to\infty} f(\,\boldsymbol{u}(\nu)\,) = m.$$

The existence of a $\boldsymbol{u}^*$ such that $f(\boldsymbol{u}^*) = \sup_{\boldsymbol{u}\in K} f(\boldsymbol{u})$ is proved in a similar fashion. □

Definition 6.53. Let $(\boldsymbol{U}, \|-\|)$ be a normed space. A set $S \subset \boldsymbol{U}$ is called *bounded* in the norm $\|-\|$ if there exists $M > 0$ such that

$$\|\boldsymbol{s}\| \leq M, \;\; \forall \boldsymbol{s} \in S.$$

In other words, S is bounded if and only if

$$\sup_{\boldsymbol{s}\in S} \|\boldsymbol{s}\| < \infty$$ □

Let us observe again that if a set S is bounded in a norm $\|-\|$, it is also bounded in any other equivalent norm.

Proposition 6.54. *Let $(\boldsymbol{U}, \|-\|)$ be a normed space and $K \subset \boldsymbol{U}$ a compact subset. Then K is both bounded and closed.*

Proof. We first prove that K is bounded. Indeed, Example 6.25 shows that the function $f : K \to \mathbb{R}$, $f(\boldsymbol{u}) = \|\boldsymbol{u}\|$ is continuous and Theorem 6.52 implies that $\sup_{\boldsymbol{u}\in K} f(\boldsymbol{u}) < \infty$. This proves that K is bounded.

To prove that K is also closed consider a seequence $(\boldsymbol{u}(n)\,)$ in K which converges to some $\boldsymbol{u}^*$ in $\boldsymbol{U}$. We have to show that in fact $\boldsymbol{u}^* \in K$. Indded, since K is compact, the sequence $(\,\boldsymbol{u}(n)\,)$ contains a subsequence which converges to some point in K. On the other hand, since $(\,\boldsymbol{u}(n)\,)$ is convergent, limit of this subsequence must coincide with the limit of $\boldsymbol{u}(n)$ which is $\boldsymbol{u}^*$. Hence $\boldsymbol{u}^* \in K$. □

Theorem 6.55. *Let $(\boldsymbol{U}, \|-\|)$ be a* finite dimensional *normed space, and $K \subset \boldsymbol{U}$. Then the following statements are equivalent.*

(i) *The set K is compact.*
(ii) *The set K is bounded and closed.*

Proof. The implication (i) ⇒ (ii) is true for any normed space so we only have to concetrate on the opposite implication. For simplicity we assume that $\boldsymbol{U}$ is a real vector space of dimension N. By fixing a basis of $\boldsymbol{U}$ we can identify it with $\mathbb{R}^N$. Since the notions of compactness, boundedness and closedness are independent of the norm we use on $\mathbb{R}^N$ we can assume that $\|-\|$ is the norm $|-|_\infty$.

Since K is bounded we deduce that there exists $M > 0$ such that

$$|\boldsymbol{u}|_\infty \leq M, \;\; \forall \boldsymbol{u} \in K.$$

In particular we deduce that

$$u_k \in [-M, M], \;\; \forall k = 1, \dots, N, \;\; \forall \boldsymbol{u} \in K,$$

where as usual we denote by $u_1, \dots, u_N$ the coordinates of a vector $\boldsymbol{u} \in \mathbb{R}^N$.

Suppose now that $(\,\boldsymbol{u}(n)\,)$ is a sequence in K. We have to show that it contains a subsequence that converges to a vector in K. We deduce from the above that

$$u_k(n) \in [-M, M], \;\; \forall k = 1, \dots, N, \;\; \forall n.$$

Since the interval $[-M, M]$ is a compact subset of $\mathbb{R}$ we can find a subsequence $(, \boldsymbol{u}(\nu))$ of $(\boldsymbol{u}(n))$ such that each of the sequences $(u_k(\nu))$, $k = 1, \dots, N$ converges to some $u_k^* \in [-M, M]$. denote by $\boldsymbol{u}^*$ the vector in $\mathbb{R}^N$ with coordinates $u_1^*, \dots, u_N^*$. Invoking Proposition 6.10 we deuce that the subsequence $(\boldsymbol{u}(\nu))$ converges in the norm $|-|_\infty$ to $\boldsymbol{u}^*$. Since K is closed and the sequence $(\boldsymbol{u}(\nu))$ is in K we deduce that its limit $\boldsymbol{u}^*$ is also in K. $\square$

6.10. **Exercises.**

Exercise 6.1. (a) Consider the space $\mathbb{R}^N$ with canonical basis $\boldsymbol{e}_1, \dots, \boldsymbol{e}_N$. Prove that the functions $|-|_1,\ |-|_\infty : \mathbb{R}^N \to [0,\infty)$ defined by

$$|u_1\boldsymbol{e}_1 + \cdots + u_N\boldsymbol{e}_N|_\infty := \max_{1\le k\le N} |u_k|, ,\quad |\boldsymbol{u}|_1 = \sum_{k=1}^N |u_k|$$

are norms on $\mathbb{R}^N$.
(b) Let $\boldsymbol{U}$ denote the space $C([0,1])$ of continuous functions $\boldsymbol{u} : [0,1] \to \mathbb{R}$. Then the function

$$\|-\|_\infty : C([0,1]) \to [0,\infty),\quad \|\boldsymbol{u}\|_\infty := \sup_{x\in[0,1]} |\boldsymbol{u}(x)|$$

is a norm on $\boldsymbol{U}$. □

Exercise 6.2. Prove Proposition 6.12. □

Exercise 6.3. (a) Prove that the space $\boldsymbol{U}$ of continuous functions $\boldsymbol{u} : [0,1] \to \mathbb{R}$ equipped with the norm

$$\|\boldsymbol{u}\| = \max_{x\in[0,1]} |\boldsymbol{u}(x)|$$

is a complete normed space.
(b) Prove that the above vector space $\boldsymbol{U}$ is *not finite dimensional*.

For part (a) you need to use the concept of uniform convergence of functions. For part (b) consider the linear functionals $L_n : \boldsymbol{U} \to \mathbb{R}$ $L_n(\boldsymbol{u}) = \boldsymbol{u}(\frac{1}{n})$, $\forall \boldsymbol{u} \in \boldsymbol{U}$, $n > 0$ and then show they are linearly independent. □

Exercise 6.4. Prove Proposition 6.32. □

Exercise 6.5. Prove Theorem 6.48. □

Exercise 6.6. Let $(\boldsymbol{U}, \|-\|)$ be a normed space.

(a) Show that if $\mathscr{C}_1, \dots, \mathscr{C}_n$ are closed subsets of $\boldsymbol{U}$, then so is their union $\mathscr{C}_1 \cup \cdots \cup \mathscr{C}_n$.
(b) Show that if $(\mathscr{C}_i)_{i\in I}$ is an arbitrary family of closed subset of $\boldsymbol{U}$, then their intersection $\cap_{i\in I}\mathscr{C}_i$ is also closed.
(c) Show that if $\mathcal{O}_1, \dots, \mathcal{O}_n$ are open subsets of $\boldsymbol{U}$, then so is their intersection $\mathcal{O}_1 \cap \cdots \cap \mathcal{O}_n$.
(d) Show that if $(\mathcal{O}_i)_{i\in I}$ is an arbitrary family of open subset of $\boldsymbol{U}$, then their union $\cup_{i\in I}\mathcal{O}_i$ is also open. □

Exercise 6.7. Suppose that $(\boldsymbol{U}, \|-\|)$ is a normed space. Prove that for any $\boldsymbol{u} \in \boldsymbol{U}$ and any $r > 0$ the ball $B(\boldsymbol{u}, r)$ is an open subset. □

Exercise 6.8. (a) Show that the subset $[0,\infty) \subset \mathbb{R}$ is closed.
(b) Use (a), Theorem 6.48 and Exercise 6.6 to show that the set

$$\mathcal{P}_N := \big\{\boldsymbol{\mu} \in \mathbb{R}^N;\ \ \mu_1 + \cdots + \mu_N = 1,\ \ \mu_k \ge 0,\ \ \forall k = 1,\dots,N\,\big\}$$

is closed. Draw a picture of $\mathcal{P}_N$ for $n = 1, 2, 3$.
Hint: Use Theorem 6.34 to prove that the maps $f : \mathbb{R}^N \to \mathbb{R}$ and $\ell_k : \mathbb{R}^N \to \mathbb{R}$ given by

$$f(\boldsymbol{u}) = u_1 + \cdots + u_N,\quad \ell_k(\boldsymbol{u}) = u_k$$

are continuous. $\square$

Exercise 6.9. Consider the 2×2-matrix

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

(a) Compute J^2, J^3, J^4, J^5.

(b) Let $t \in \mathbb{R}$. Compute e^{tJ}. Describe the behavior of the point

$$\boldsymbol{u}(t) = e^{tJ} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \in \mathbb{R}^2$$

as t varies in the interval $[0, 2\pi]$. $\square$

Exercise 6.10. Prove Newton's binomial formula, Lemma 6.42, using induction on n. $\square$

REFERENCES

[1] S. Axler: *Linear Algebra Done Right*, Springer Verlag, 2004.
[2] V.V. Prasolov: *Problems and Theorem in Linear Algebra*, in Russian.
[3] S. Treil: *Linear Algebra Done Wrong*, Notes for a Linear algebra course at Brown University.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF NOTRE DAME, NOTRE DAME, IN 46556-4618.
E-mail address: nicolaescu.1@nd.edu
URL: http://www.nd.edu/~lnicolae/