

LINEAR OPERATORS

BRANKO ČURĀUS

Throughout this note \mathcal{V} is a vector space over a scalar field \mathbb{F} . \mathbb{N} denotes the set of positive integers and $i, j, k, l, m, n, p \in \mathbb{N}$.

1. FUNCTIONS

First we review formal definitions related to functions. In this section A and B are nonempty sets.

The formal definition of function identifies a function and its graph. A justification for this is the fact that if you know the graph of a function, then you know the function, and conversely, if you know a function you know its graph. Simply stated the definition below says that a function from a set A to a set B is a subset f of the Cartesian product $A \times B$ such that for each $x \in A$ there exists unique $y \in B$ such that $(x, y) \in f$.

A *function* from A into B is a subset f of the Cartesian product $A \times B$ such that

- (a) $\forall x \in A \exists y \in B (x, y) \in f$,
- (b) $\forall x \in A \forall y \in B \forall z \in B (x, y) \in f \wedge (x, z) \in f \Rightarrow y = z$.

The relationship $(x, y) \in f$ is commonly written as $y = f(x)$. The symbol $f : A \rightarrow B$ denotes a function from A to B .

The set A is the *domain* of $f : A \rightarrow B$. The set B is the *codomain* of $f : A \rightarrow B$. The set

$$\{y \in B : \exists x \in A y = f(x)\}$$

is called the *range* of $f : A \rightarrow B$. It is denoted by $\text{ran } f$.

A function $f : A \rightarrow B$ is a *surjection* if for every $y \in B$ there exists $x \in A$ such that $y = f(x)$.

A function $f : A \rightarrow B$ is an *injection* if for every $x_1, x_2 \in A$ the following implication holds: $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$.

A function $f : A \rightarrow B$ is a *bijection* if it is both: a surjection and an injection.

Next we give a formal definition of a composition of two functions. However, before giving a definition we need to prove a proposition.

Proposition 1.1. *Let $f : A \rightarrow B$ and $g : C \rightarrow D$ be functions. If $\text{ran } f \subseteq C$, then*

$$\{(x, z) \in A \times D : \exists y \in B (x, y) \in f \wedge (y, z) \in g\} \quad (1.1)$$

is a function from A to D .

Proof. A proof is a nice exercise. \square

The function defined by (1.1) is called the *composition* of functions f and g . It is denoted by $f \circ g$.

The function

$$\{(x, x) \in A \times A : x \in A\}$$

is called the *identity function* on A . It is denoted by id_A . In the standard notation id_A is the function $\text{id}_A : A \rightarrow A$ such that $\text{id}_A(x) = x$ for all $x \in A$.

A function $f : A \rightarrow B$ is *invertible* if there exist functions $g : B \rightarrow A$ and $h : B \rightarrow A$ such that $f \circ g = \text{id}_B$ and $h \circ f = \text{id}_A$.

Theorem 1.2. *Let $f : A \rightarrow B$ be a function. The following statements are equivalent.*

- (a) *The function f is invertible.*
- (b) *The function f is a bijection.*
- (c) *There exists a unique function $g : B \rightarrow A$ such that $f \circ g = \text{id}_B$ and $g \circ f = \text{id}_A$.*

If f is invertible, then the unique g whose existence is proved in Theorem 1.2 (c) is called the *inverse* of f ; it is denoted by f^{-1} .

Below are few exercises about functions from my Math 312 notes.

Exercise 1.3. Let A , B and C be nonempty sets. Let $f : A \rightarrow B$ and $g : B \rightarrow C$ be injections. Prove that $g \circ f : A \rightarrow C$ is an injection.

Exercise 1.4. Let A , B and C be nonempty sets. Let $f : A \rightarrow B$ and $g : B \rightarrow C$ be surjections. Prove that $g \circ f : A \rightarrow C$ is a surjection.

Exercise 1.5. Let A , B and C be nonempty sets. Let $f : A \rightarrow B$ and $g : B \rightarrow C$ be bijections. Prove that $g \circ f : A \rightarrow C$ is a bijection. Prove that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Exercise 1.6. Let A , B and C be nonempty sets. Let $f : A \rightarrow B$, $g : B \rightarrow C$. Prove that if $g \circ f$ is an injection, then f is an injection.

Exercise 1.7. Let A , B and C be nonempty sets and let $f : A \rightarrow B$, $g : B \rightarrow C$. Prove that if $g \circ f$ is a surjection, then g is a surjection.

Exercise 1.8. Let A , B and C be nonempty sets and let $f : A \rightarrow B$, $g : B \rightarrow C$ and $h : C \rightarrow A$ be three functions. Prove that if any two of the functions $h \circ g \circ f$, $g \circ f \circ h$, $f \circ h \circ g$ are injections and the third is a surjection, or if any two of them are surjections and the third is an injection, then f , g , and h are bijections.

2. LINEAR OPERATORS

In this section \mathcal{U} , \mathcal{V} and \mathcal{W} are vector spaces over a scalar field \mathbb{F} .

2.1. The definition and the vector space of all linear operators. A function $T : \mathcal{V} \rightarrow \mathcal{W}$ is said to be a *linear operator* if the following it satisfies the following conditions:

$$\forall u \in \mathcal{V} \quad \forall v \in \mathcal{V} \quad T(u + v) = T(u) + T(v), \quad (2.1)$$

$$\forall \alpha \in \mathbb{F} \quad \forall v \in \mathcal{V} \quad T(\alpha v) = \alpha T(v). \quad (2.2)$$

The property (2.1) is called *additivity*, while the property (2.2) is called *homogeneity*. Together additivity and homogeneity are called *linearity*.

Denote by $\mathcal{L}(\mathcal{V}, \mathcal{W})$ the set of all linear operators from \mathcal{V} to \mathcal{W} . Define the addition and scaling in $\mathcal{L}(\mathcal{V}, \mathcal{W})$. For $S, T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ and $\alpha \in \mathbb{F}$ we define

$$(S + T)(v) = S(v) + T(v), \quad \forall v \in \mathcal{V}, \quad (2.3)$$

$$(\alpha T)(v) = \alpha T(v), \quad \forall v \in \mathcal{V}. \quad (2.4)$$

$$(2.5)$$

Notice that two plus signs which appear in (2.3) have different meanings. The plus sign on the left-hand side stands for the addition of linear operators that is just being defined, while the plus sign on the right-hand side stands for the addition in \mathcal{W} . Notice the analogous difference in empty spaces between α and T in (2.4). Define the zero mapping in $\mathcal{L}(\mathcal{V}, \mathcal{W})$ to be

$$0_{\mathcal{L}(\mathcal{V}, \mathcal{W})}(v) = 0_{\mathcal{W}}, \quad \forall v \in \mathcal{V}.$$

For $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ we define its opposite operator by

$$(-T)(v) = -T(v), \quad \forall v \in \mathcal{V}.$$

Proposition 2.1. *The set $\mathcal{L}(\mathcal{V}, \mathcal{W})$ with the operations defined in (2.3), and (2.4) is a vector space over \mathbb{F} .*

For $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ and $v \in \mathcal{V}$ it is customary to write Tv instead of $T(v)$.

Example 2.2. Assume that a vector space \mathcal{V} is a direct sum of its subspaces \mathcal{U} and \mathcal{W} , that is $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$. Define the function $P : \mathcal{V} \rightarrow \mathcal{V}$ by

$$Pv = w \quad \Leftrightarrow \quad v = u + w, \quad u \in \mathcal{U}, \quad w \in \mathcal{W}.$$

Then P is a linear operator. It is called the *projection* of \mathcal{V} onto \mathcal{W} parallel to \mathcal{U} ; it is denoted by $P_{\mathcal{W} \parallel \mathcal{U}}$.

2.2. Composition, inverse, isomorphism. In the next two propositions we prove that the linearity is preserved under composition of linear operators and under taking the inverse of a linear operator.

Proposition 2.3. *Let $S : \mathcal{U} \rightarrow \mathcal{V}$ and $T : \mathcal{V} \rightarrow \mathcal{W}$ be linear operators. The composition $T \circ S : \mathcal{U} \rightarrow \mathcal{W}$ is a linear operator.*

Proof. Prove this as an exercise. □

When composing linear operators it is customary to write simply TS instead of $T \circ S$.

The identity function on \mathcal{V} is denoted by $I_{\mathcal{V}}$. It is defined by $I_{\mathcal{V}}(v) = v$ for all $v \in \mathcal{V}$. It is clearly a linear operator.

Proposition 2.4. *Let $T : \mathcal{V} \rightarrow \mathcal{W}$ be a linear operator which is invertible. Then the inverse $T^{-1} : \mathcal{W} \rightarrow \mathcal{V}$ of T is a linear operator.*

Proof. Since T is invertible, by Theorem 1.2 there exists a function $S : \mathcal{W} \rightarrow \mathcal{V}$ such that $ST = I_{\mathcal{V}}$ and $TS = I_{\mathcal{W}}$. Since T is linear and $TS = I_{\mathcal{W}}$ we have

$$T(\alpha Sx + \beta Sy) = \alpha T(Sx) + \beta T(Sy) = \alpha(TS)x + \beta(TS)y = \alpha x + \beta y$$

for all $\alpha, \beta \in \mathbb{F}$ and all $x, y \in \mathcal{W}$. Applying S to both sides of

$$T(\alpha Sx + \beta Sy) = \alpha x + \beta y$$

we get

$$(ST)(\alpha Sx + \beta Sy) = S(\alpha x + \beta y) \quad \forall \alpha, \beta \in \mathbb{F} \quad \forall x, y \in \mathcal{W}.$$

Since $ST = I_{\mathcal{V}}$, we get

$$\alpha Sx + \beta Sy = S(\alpha x + \beta y) \quad \forall \alpha, \beta \in \mathbb{F} \quad \forall x, y \in \mathcal{W},$$

thus proving the linearity of S . Since by definition $S = T^{-1}$ the proposition is proved. \square

A linear operator $T : \mathcal{V} \rightarrow \mathcal{W}$ which is a bijection is called an *isomorphism* between vector spaces \mathcal{V} and \mathcal{W} .

By Theorem 1.2 and Proposition 2.4 each isomorphism is invertible and its inverse is also an isomorphism.

In the next proposition we introduce the most important isomorphism $C_{\mathcal{B}}$.

Proposition 2.5. *Let \mathcal{V} be a finite dimensional vector space over \mathbb{F} and $n = \dim \mathcal{V}$. Let $\mathcal{B} = \{v_1, \dots, v_n\}$ be a basis for \mathcal{V} . There exists a function $C_{\mathcal{B}} : \mathcal{V} \rightarrow \mathbb{F}^n$ such that*

$$C_{\mathcal{B}}(v) = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \quad \Leftrightarrow \quad v = \sum_{j=1}^n \alpha_j v_j.$$

The function $C_{\mathcal{B}}$ is an isomorphism between \mathcal{V} and \mathbb{F}^n .

Proof. Since \mathcal{B} spans \mathcal{V} , for every $v \in \mathcal{V}$ there exist $\alpha_1, \dots, \alpha_n \in \mathbb{F}$ such that $v = \sum_{j=1}^n \alpha_j v_j$. Thus $C_{\mathcal{B}}$ is defined for every $v \in \mathcal{V}$. To prove that $C_{\mathcal{B}}$ is a function assume

$$C_{\mathcal{B}}(v) = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \quad \text{and} \quad C_{\mathcal{B}}(v) = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}.$$

Then

$$v = \sum_{j=1}^n \alpha_j v_j \quad \text{and} \quad v = \sum_{j=1}^n \beta_j v_j.$$

Therefore $0_{\mathcal{V}} = \sum_{j=1}^n (\alpha_j - \beta_j) v_j$. Since \mathcal{B} is linearly independent, $\alpha_j = \beta_j$ for all $j \in \{1, \dots, n\}$. Thus

$$C_{\mathcal{B}}(v) = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \quad \text{and} \quad C_{\mathcal{B}}(v) = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}.$$

This proves that $C_{\mathcal{B}}$ is a function.

The linearity of $C_{\mathcal{B}}$ is easy to verify.

The injectivity of $C_{\mathcal{B}}$ follows from the linear independence of \mathcal{B} .

The surjectivity of $C_{\mathcal{B}}$ follows from the fact that for arbitrary $\alpha_1, \dots, \alpha_n \in \mathbb{F}$ we have $v = \sum_{j=1}^n \alpha_j v_j \in \mathcal{V}$ and therefore

$$C_{\mathcal{B}}(v) = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}. \quad \square$$

In the last part of the proof of Proposition 2.5 we showed that the formula for the inverse $(C_{\mathcal{B}})^{-1} : \mathbb{F}^n \rightarrow \mathcal{V}$ of $C_{\mathcal{B}}$ is given by

$$(C_{\mathcal{B}})^{-1} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \sum_{j=1}^n \alpha_j v_j, \quad \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in \mathbb{F}^n. \quad (2.6)$$

Notice that (2.6) defines a function from \mathbb{F}^n to \mathcal{V} even if \mathcal{B} is not a basis of \mathcal{V} .

Example 2.6. Inspired by the definition of $C_{\mathcal{B}}$ and (2.6) we define a general operator of this kind. Let \mathcal{V} and \mathcal{W} be vector spaces over \mathbb{F} . Let \mathcal{V} be finite dimensional, $n = \dim \mathcal{V}$ and let \mathcal{B} be a basis for \mathcal{V} . Let $\mathcal{C} = (w_1, \dots, w_n)$ be any n -tuple of vectors in \mathcal{W} . The entries of an n -tuple can be repeated, they can all be equal, for example to $0_{\mathcal{W}}$. We define the linear operator $L_{\mathcal{C}}^{\mathcal{B}} : \mathcal{V} \rightarrow \mathcal{W}$ by

$$L_{\mathcal{C}}^{\mathcal{B}}(v) = \sum_{j=1}^n \alpha_j w_j \quad \text{where} \quad \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = C_{\mathcal{B}}(v). \quad (2.7)$$

In fact, $L_{\mathcal{C}}^{\mathcal{B}} : \mathcal{V} \rightarrow \mathcal{W}$ is a composition of $C_{\mathcal{B}} : \mathcal{V} \rightarrow \mathbb{F}^n$ and the operator $\mathbb{F}^n \rightarrow \mathcal{W}$ defined by

$$\begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \mapsto \sum_{j=1}^n \alpha_j w_j \quad \text{for arbitrary} \quad \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \in \mathbb{F}^n. \quad (2.8)$$

It is easy to verify that (2.8) defines a linear operator.

Denote by \mathcal{E} the standard basis of \mathbb{F}^n , that is the basis which consists of the columns of the identity matrix. Then $C_{\mathcal{B}} = L_{\mathcal{E}}^{\mathcal{B}}$ and $(C_{\mathcal{B}})^{-1} = L_{\mathcal{B}}^{\mathcal{E}}$.

Exercise 2.7. Let \mathcal{V} and \mathcal{W} be vector spaces over \mathbb{F} . Let \mathcal{V} be finite dimensional, $n = \dim \mathcal{V}$ and let \mathcal{B} be a basis for \mathcal{V} . Let $\mathcal{C} = (w_1, \dots, w_n)$ be a list of vectors in \mathcal{W} with n entries.

- (a) Characterize the injectivity of $L_{\mathcal{C}}^{\mathcal{B}} : \mathcal{V} \rightarrow \mathcal{W}$.
- (b) Characterize the surjectivity of $L_{\mathcal{C}}^{\mathcal{B}} : \mathcal{V} \rightarrow \mathcal{W}$.
- (c) Characterize the bijectivity of $L_{\mathcal{C}}^{\mathcal{B}} : \mathcal{V} \rightarrow \mathcal{W}$.
- (d) If $L_{\mathcal{C}}^{\mathcal{B}} : \mathcal{V} \rightarrow \mathcal{W}$ is an isomorphism, find a simple formula for $(L_{\mathcal{C}}^{\mathcal{B}})^{-1}$.

2.3. The rank-nullity theorem. Let $T : \mathcal{V} \rightarrow \mathcal{W}$ be a linear operator. The linearity of T implies that the set

$$\text{nul } T = \{v \in \mathcal{V} : Tv = 0_{\mathcal{W}}\}$$

is a subspace of \mathcal{V} . This subspace is called the *null space* of T . Similarly, the linearity of T implies that the range of T is a subspace of \mathcal{W} . Recall that

$$\text{ran } T = \{w \in \mathcal{W} : \exists v \in \mathcal{V} \ w = Tv\}.$$

Proposition 2.8. *A linear operator $T : \mathcal{V} \rightarrow \mathcal{W}$ is an injection if and only if $\text{nul } T = \{0_{\mathcal{V}}\}$.*

Proof. We first prove the “if” part of the proposition. Assume that $\text{nul } T = \{0_{\mathcal{V}}\}$. Let $u, v \in \mathcal{V}$ be arbitrary and assume that $Tu = Tv$. Since T is linear, $Tu = Tv$ implies $T(u-v) = 0_{\mathcal{W}}$. Consequently $u-v \in \text{nul } T = \{0_{\mathcal{V}}\}$. Hence, $u-v = 0_{\mathcal{V}}$, that is $u = v$. This proves that T is an injection.

To prove the “only if” part assume that $T : \mathcal{V} \rightarrow \mathcal{W}$ is an injection. Let $v \in \text{nul } T$ be arbitrary. Then $Tv = 0_{\mathcal{W}} = T0_{\mathcal{V}}$. Since T is injective, $Tv = T0_{\mathcal{V}}$ implies $v = 0_{\mathcal{V}}$. Thus we have proved that $\text{nul } T \subseteq \{0_{\mathcal{V}}\}$. Since the converse inclusion is trivial, we have $\text{nul } T = \{0_{\mathcal{V}}\}$. \square

Theorem 2.9 (Rank-nullity theorem). *Let \mathcal{V} and \mathcal{W} be vector spaces over a scalar field \mathbb{F} and let $T : \mathcal{V} \rightarrow \mathcal{W}$ be a linear operator. If \mathcal{V} is finite dimensional, then $\text{nul } T$ and $\text{ran } T$ are finite dimensional and*

$$\dim(\text{nul } T) + \dim(\text{ran } T) = \dim \mathcal{V}. \quad (2.9)$$

Proof. Assume that \mathcal{V} is finite dimensional. We proved earlier that for an arbitrary subspace \mathcal{U} of \mathcal{V} there exists a subspace \mathcal{X} of \mathcal{V} such that

$$\mathcal{U} \oplus \mathcal{X} = \mathcal{V} \quad \text{and} \quad \dim \mathcal{U} + \dim \mathcal{X} = \dim \mathcal{V}.$$

Thus, there exists a subspace \mathcal{X} of \mathcal{V} such that

$$(\text{nul } T) \oplus \mathcal{X} = \mathcal{V} \quad \text{and} \quad \dim(\text{nul } T) + \dim \mathcal{X} = \dim \mathcal{V}. \quad (2.10)$$

Since $\dim(\text{nul } T) + \dim \mathcal{X}$, to prove the theorem we only need to prove that $\dim \mathcal{X} = \dim(\text{ran } T)$. To this end, let $m = \dim \mathcal{X}$ and let x_1, \dots, x_m be

a basis for \mathcal{X} . We will prove that vectors Tx_1, \dots, Tx_m form a basis for $\text{ran } T$. We first prove

$$\text{span}\{Tx_1, \dots, Tx_m\} = \text{ran } T. \quad (2.11)$$

Clearly $\{Tx_1, \dots, Tx_m\} \subseteq \text{ran } T$. Consequently, since $\text{ran } T$ is a subspace of \mathcal{W} , we have $\text{span}\{Tx_1, \dots, Tx_m\} \subseteq \text{ran } T$. To prove the converse inclusion, let $w \in \text{ran } T$ be arbitrary. Then, there exists $v \in \mathcal{V}$ such that $Tv = w$. Since $\mathcal{V} = (\text{nul } T) + \mathcal{X}$, there exist $u \in \text{nul } T$ and $x \in \mathcal{X}$ such that $v = u + x$. Then $Tv = T(u + x) = Tu + Tx = Tx$. As $x \in \mathcal{X}$, there exist $\xi_1, \dots, \xi_m \in \mathbb{F}$ such that $x = \sum_{j=1}^m \xi_j x_j$. Now we use linearity of T to deduce

$$w = Tv = Tx = \sum_{j=1}^m \xi_j Tx_j.$$

This proves that $w \in \text{span}\{Tx_1, \dots, Tx_m\}$. Since w was arbitrary in $\text{ran } T$ this completes a proof of (2.11).

Next we prove that the vectors Tx_1, \dots, Tx_m are linearly independent. Let $\alpha_1, \dots, \alpha_m \in \mathbb{F}$ be arbitrary and assume that

$$\alpha_1 Tx_1 + \dots + \alpha_m Tx_m = 0_{\mathcal{W}}. \quad (2.12)$$

Since T is linear (2.12) implies that

$$\alpha_1 x_1 + \dots + \alpha_m x_m \in \text{nul } T. \quad (2.13)$$

Recall that $x_1, \dots, x_m \in cX$ and \mathcal{X} is a subspace of \mathcal{V} , so

$$\alpha_1 x_1 + \dots + \alpha_m x_m \in \mathcal{X}. \quad (2.14)$$

Now (2.13), (2.14) and the fact that $(\text{nul } T) \cap \mathcal{X} = \{0_{\mathcal{V}}\}$ imply

$$\alpha_1 x_1 + \dots + \alpha_m x_m = 0_{\mathcal{V}}. \quad (2.15)$$

Since x_1, \dots, x_m are linearly independent (2.15) yields $\alpha_1 = \dots = \alpha_m = 0$. This completes a proof of the linear independence of Tx_1, \dots, Tx_m .

Thus $\{Tx_1, \dots, Tx_m\}$ is a basis for $\text{ran } T$. Consequently $\dim(\text{ran } T) = m$. Since $m = \dim \mathcal{X}$, (2.10) implies (2.9). This completes the proof. \square

The nonnegative integer $\dim(\text{nul } T)$ is called the *nullity* of T ; the nonnegative integer $\dim(\text{ran } T)$ is called the *rank* of T .

The rank-nullity theorem in English reads: If a linear operator is defined on a finite dimensional vector space, then its nullity and its rank are finite and they add up to the dimension of the domain.

Proposition 2.10. *Let \mathcal{V} and \mathcal{W} be vector spaces over \mathbb{F} . Assume that \mathcal{V} is finite dimensional. The following statements are equivalent*

- (a) *There exists a surjection $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$.*
- (b) *\mathcal{W} is finite dimensional and $\dim \mathcal{V} \geq \dim \mathcal{W}$.*

Proposition 2.11. *Let \mathcal{V} and \mathcal{W} be vector spaces over \mathbb{F} . Assume that \mathcal{V} is finite dimensional. The following statements are equivalent*

- (a) *There exists an injection $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$.*

(b) *Either \mathcal{W} is infinite dimensional or $\dim \mathcal{V} \leq \dim \mathcal{W}$.*

Proposition 2.12. *Let \mathcal{V} and \mathcal{W} be vector spaces over \mathbb{F} . Assume that \mathcal{V} is finite dimensional. The following statements are equivalent*

- (a) *There exists an isomorphism $T : \mathcal{V} \rightarrow \mathcal{W}$.*
 (b) *\mathcal{W} is finite dimensional and $\dim \mathcal{W} = \dim \mathcal{V}$.*

2.4. Isomorphism between $\mathcal{L}(\mathcal{V}, \mathcal{W})$ and $\mathbb{F}^{n \times m}$. Let \mathcal{V} and \mathcal{W} be finite dimensional vector spaces over \mathbb{F} , $m = \dim \mathcal{V}$, $n = \dim \mathcal{W}$, let $\mathcal{B} = \{v_1, \dots, v_m\}$ be a basis for \mathcal{V} and let $\mathcal{C} = \{w_1, \dots, w_n\}$ be a basis for \mathcal{W} . The mapping $C_{\mathcal{B}}$ provides an isomorphism between \mathcal{V} and \mathbb{F}^m and $C_{\mathcal{C}}$ provides an isomorphism between \mathcal{W} and \mathbb{F}^n .

Recall that the simplest way to define a linear operator from \mathbb{F}^m to \mathbb{F}^n is to use an $n \times m$ matrix B . It is convenient to consider an $n \times m$ matrix to be an m -tuple of its columns, which are vectors in \mathbb{F}^n . For example, let $\mathbf{b}_1, \dots, \mathbf{b}_m \in \mathbb{F}^n$ be columns of an $n \times m$ matrix B . Then we write

$$B = [\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_m].$$

This notation is convenient since it allows us to write a multiplication of a vector $\mathbf{x} \in \mathbb{F}^m$ by a matrix B as

$$B\mathbf{x} = \sum_{j=1}^m \xi_j \mathbf{b}_j \quad \text{where} \quad \mathbf{x} = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_m \end{bmatrix}. \quad (2.16)$$

Notice the similarity of the definition in (2.16) to the definition (2.7) of the operator $L_{\mathcal{C}}^{\mathcal{B}}$ in Example 2.6. Taking \mathcal{B} to be the standard basis of \mathbb{F}^m and taking \mathcal{C} to be the m -tuple given by B , we have $L_{\mathcal{C}}^{\mathcal{B}}(\mathbf{x}) = B\mathbf{x}$.

Let $T : \mathcal{V} \rightarrow \mathcal{W}$ be a linear operator. Our next goal is to connect T in a natural way to a certain $n \times m$ matrix B . That “natural way” is suggested by following diagram:

$$\begin{array}{ccc} \mathcal{V} & \xrightarrow{T} & \mathcal{W} \\ \downarrow C_{\mathcal{B}} & & \downarrow C_{\mathcal{C}} \\ \mathbb{F}^m & \xrightarrow{B} & \mathbb{F}^n \end{array}$$

We seek an $n \times m$ matrix B such that the action of T between \mathcal{V} and \mathcal{W} is in some sense replicated by the action of B between \mathbb{F}^m and \mathbb{F}^n . Precisely, we seek B such that

$$C_{\mathcal{C}}(Tv) = B(C_{\mathcal{B}}(v)) \quad \forall v \in \mathcal{V}. \quad (2.17)$$

In English: multiplying the vector of coordinates of v by B we get exactly the coordinates of Tv .

Using the basis vectors $v_1, \dots, v_m \in \mathcal{B}$ in (2.17) we see that the matrix

$$B = [C_{\mathcal{C}}(Tv_1) \quad \cdots \quad C_{\mathcal{C}}(Tv_m)] \quad (2.18)$$

has the desired property (2.17).

For an arbitrary $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ the formula (2.18) associates the matrix $B \in \mathbb{F}^{n \times m}$ with T . In other words (2.18) defines a function from $\mathcal{L}(\mathcal{V}, \mathcal{W})$ to $\mathbb{F}^{n \times m}$.

Theorem 2.13. *Let \mathcal{V} and \mathcal{W} be finite dimensional vector spaces over \mathbb{F} , $m = \dim \mathcal{V}$, $n = \dim \mathcal{W}$, let $\mathcal{B} = \{v_1, \dots, v_m\}$ be a basis for \mathcal{V} and let $\mathcal{C} = \{w_1, \dots, w_n\}$ be a basis for \mathcal{W} . The function*

$$M_{\mathcal{C}}^{\mathcal{B}} : \mathcal{L}(\mathcal{V}, \mathcal{W}) \rightarrow \mathbb{F}^{n \times m}$$

defined by

$$M_{\mathcal{C}}^{\mathcal{B}}(T) = [C_{\mathcal{C}}(Tv_1) \ \cdots \ C_{\mathcal{C}}(Tv_m)], \quad T \in \mathcal{L}(\mathcal{V}, \mathcal{W}) \quad (2.19)$$

is an isomorphism.

Proof. It is easy to verify that $M_{\mathcal{C}}^{\mathcal{B}}$ is a linear operator.

Since the definition of $M_{\mathcal{C}}^{\mathcal{B}}(T)$ coincides with (2.18), equality (2.17) yields

$$C_{\mathcal{C}}(Tv) = (M_{\mathcal{C}}^{\mathcal{B}}(T))C_{\mathcal{B}}(v). \quad (2.20)$$

The most direct way to prove that $M_{\mathcal{C}}^{\mathcal{B}}$ is an isomorphism is to construct its inverse. The inverse is suggested by the diagram (2.21).

$$\begin{array}{ccc} \mathcal{V} & \xrightarrow{\text{---} T \text{---}} & \mathcal{W} \\ C_{\mathcal{B}} \downarrow & & \uparrow (C_{\mathcal{C}})^{-1} \\ \mathbb{F}^m & \xrightarrow{B} & \mathbb{F}^n \end{array} \quad (2.21)$$

Define

$$N_{\mathcal{C}}^{\mathcal{B}} : \mathbb{F}^{n \times m} \rightarrow \mathcal{L}(\mathcal{V}, \mathcal{W})$$

by

$$(N_{\mathcal{C}}^{\mathcal{B}}(B))(v) = (C_{\mathcal{C}})^{-1}(B(C_{\mathcal{B}}(v))), \quad B \in \mathbb{F}^{n \times m}. \quad (2.22)$$

Next we prove that

$$N_{\mathcal{C}}^{\mathcal{B}} \circ M_{\mathcal{C}}^{\mathcal{B}} = I_{\mathcal{L}(\mathcal{V}, \mathcal{W})} \quad \text{and} \quad M_{\mathcal{C}}^{\mathcal{B}} \circ N_{\mathcal{C}}^{\mathcal{B}} = I_{\mathbb{F}^{n \times m}}.$$

First for arbitrary $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ and arbitrary $v \in \mathcal{V}$ we calculate

$$\begin{aligned} \left((N_{\mathcal{C}}^{\mathcal{B}} \circ M_{\mathcal{C}}^{\mathcal{B}})(T) \right)(v) &= (C_{\mathcal{C}})^{-1}((M_{\mathcal{C}}^{\mathcal{B}}(T))(C_{\mathcal{B}}(v))) && \text{by (2.22)} \\ &= (C_{\mathcal{C}})^{-1}(C_{\mathcal{C}}(Tv)) && \text{by (2.20)} \\ &= Tv. \end{aligned}$$

Thus $(N_{\mathcal{C}}^{\mathcal{B}} \circ M_{\mathcal{C}}^{\mathcal{B}})(T) = T$ and thus, since $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ was arbitrary, $N_{\mathcal{C}}^{\mathcal{B}} \circ M_{\mathcal{C}}^{\mathcal{B}} = I_{\mathcal{L}(\mathcal{V}, \mathcal{W})}$.

Let now $B \in \mathbb{F}^{n \times m}$ be arbitrary and calculate

$$(M_{\mathcal{C}}^{\mathcal{B}} \circ N_{\mathcal{C}}^{\mathcal{B}})(B) = M_{\mathcal{C}}^{\mathcal{B}}(N_{\mathcal{C}}^{\mathcal{B}}(B))$$

$$\begin{aligned}
&= \left[C_{\mathcal{C}}((N_{\mathcal{C}}^{\mathcal{B}}(B))(v_1)) \cdots C_{\mathcal{C}}((N_{\mathcal{C}}^{\mathcal{B}}(B))(v_m)) \right] \text{ by (2.19)} \\
&= \left[B(C_{\mathcal{B}}(v_1)) \cdots B(C_{\mathcal{B}}(v_m)) \right] \text{ by (2.22)} \\
&= B \left[C_{\mathcal{B}}(v_1) \cdots C_{\mathcal{B}}(v_m) \right] \text{ matrix mult.} \\
&= B I_m \text{ def. of } C_{\mathcal{B}} \\
&= B.
\end{aligned}$$

Thus $(M_{\mathcal{C}}^{\mathcal{B}} \circ N_{\mathcal{C}}^{\mathcal{B}})(B) = B$ for all $B \in \mathbb{F}^{n \times m}$, proving that $M_{\mathcal{C}}^{\mathcal{B}} \circ N_{\mathcal{C}}^{\mathcal{B}} = I_{\mathbb{F}^{n \times m}}$.

This completes the proof that $M_{\mathcal{C}}^{\mathcal{B}}$ is a bijection. Since it is linear, $M_{\mathcal{C}}^{\mathcal{B}}$ is an isomorphism. \square

Theorem 2.14. *Let \mathcal{U} , \mathcal{V} and \mathcal{W} be finite dimensional vector spaces over \mathbb{F} , $k = \dim \mathcal{U}$, $m = \dim \mathcal{V}$, $n = \dim \mathcal{W}$, let \mathcal{A} be a basis for \mathcal{U} , let \mathcal{B} be a basis for \mathcal{V} , and let \mathcal{C} be a basis for \mathcal{W} . Let $S \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ and $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$. Let $M_{\mathcal{B}}^{\mathcal{A}}(S) \in \mathbb{F}^{m \times k}$, $M_{\mathcal{C}}^{\mathcal{B}}(T) \in \mathbb{F}^{n \times m}$ and $M_{\mathcal{C}}^{\mathcal{A}}(TS) \in \mathbb{F}^{n \times k}$ be as defined in Theorem 2.13. Then*

$$M_{\mathcal{C}}^{\mathcal{A}}(TS) = M_{\mathcal{C}}^{\mathcal{B}}(T)M_{\mathcal{B}}^{\mathcal{A}}(S).$$

Proof. Let $\mathcal{A} = \{u, \dots, u_k\}$ and calculate

$$\begin{aligned}
M_{\mathcal{C}}^{\mathcal{A}}(TS) &= \left[C_{\mathcal{C}}(TSu_1) \cdots C_{\mathcal{C}}(TSu_k) \right] \text{ by (2.19)} \\
&= \left[M_{\mathcal{C}}^{\mathcal{B}}(T)(C_{\mathcal{B}}(Su_1)) \cdots M_{\mathcal{C}}^{\mathcal{B}}(T)(C_{\mathcal{B}}(Su_k)) \right] \text{ by (2.20)} \\
&= M_{\mathcal{C}}^{\mathcal{B}}(T) \left[C_{\mathcal{B}}(Su_1) \cdots C_{\mathcal{B}}(Su_k) \right] \text{ matrix mult.} \\
&= M_{\mathcal{C}}^{\mathcal{B}}(T)M_{\mathcal{B}}^{\mathcal{A}}(S). \text{ by (2.19)}
\end{aligned}$$

\square

The following diagram illustrates the content of Theorem 2.14.

$$\begin{array}{ccccc}
& & \mathcal{V} & & \\
& \nearrow S & \downarrow & \searrow T & \\
\mathcal{U} & \xrightarrow{TS} & & \xrightarrow{} & \mathcal{W} \\
& \downarrow C_{\mathcal{A}} & \downarrow C_{\mathcal{B}} & & \downarrow C_{\mathcal{C}} \\
& & \mathbb{F}^m & & \\
& \nearrow M_{\mathcal{B}}^{\mathcal{A}}(S) & & \searrow M_{\mathcal{C}}^{\mathcal{B}}(T) & \\
\mathbb{F}^k & \xrightarrow{M_{\mathcal{C}}^{\mathcal{A}}(TS) = M_{\mathcal{C}}^{\mathcal{B}}(T)M_{\mathcal{B}}^{\mathcal{A}}(S)} & & \xrightarrow{} & \mathbb{F}^n
\end{array}$$