LINEAR OPERATORS

BRANKO ĆURGUS

Throughout these notes, we study vector spaces over a scalar field $\mathbb{F}$, which is either $\mathbb{R}$ or $\mathbb{C}$. The set of positive integers is denoted by $\mathbb{N}$, and its elements are i, j, k, l, m, n, p. For a nonempty finite set A, the number of elements in A is denoted by $\#A \in \mathbb{N}$, with $\#\emptyset = 0$.

Vector spaces and sets of vectors are denoted by capital calligraphic letters, such as $\mathscr{V}, \mathscr{X}, \mathscr{A}$, etc. Vectors in abstract vector spaces are denoted by lowercase Latin letters, such as u, v, x, y, etc. Linear operators are denoted by uppercase Latin letters, such as S, T, etc. Scalars are represented by lowercase Greek letters, such as α, β, etc.

Vectors in Euclidean spaces $\mathbb{F}^n$ are denoted by boldface lowercase letters, such as $\mathbf{a}, \mathbf{b}$, etc. Matrices with entries in $\mathbb{F}$ are denoted by uppercase Latin letters in sans-serif font, such as $\mathsf{A}, \mathsf{B}, \mathsf{M}$, etc. The $n \times n$ identity matrix is denoted by I_n, while $\mathsf{0}$ represents a zero matrix whose size will be specified in context. The transpose of a matrix M is denoted by $\mathsf{M}^\top$.

Pay attention to exceptions to these conventions. If you notice significant deviations, please let me know.

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 $(x, y) \in F$.

Definition 1.1. A **function** from A into B is a subset f of the Cartesian product $A \times B$ such that the following two conditions are satisfied

$$\forall x \in A \quad \exists y \in B \quad \text{such that} \quad (x, y) \in f. \qquad (\textbf{Total})$$

$$\forall x \in A \quad \forall y, z \in B \quad (x, y) \in f \wedge (x, z) \in f \quad \Rightarrow \quad y = z \qquad (\textbf{UniVal})$$

◊

Date: February 3, 2026 at 14:51.

The condition (**UniVal**) in Definition 1.1 is popularly known as the *vertical line test*. Its full form is as follows:

$$\begin{array}{l}\forall x_1, x_2 \in A\\ \forall y_1, y_2 \in B\end{array} \quad (x_1, y_1) \in f \wedge (x_2, y_2) \in f \wedge x_1 = x_2 \;\Rightarrow\; y_1 = y_2. \tag{1.1}$$

The implication in (1.1) is important since its partial contrapositive is often used in proofs. Its partial contrapositive is:

$$\begin{array}{l}\forall x_1, x_2 \in A\\ \forall y_1, y_2 \in B\end{array} \quad (x_1, y_1) \in f \wedge (x_2, y_2) \in f \wedge y_1 \neq y_2 \;\Rightarrow\; x_1 \neq x_2. \tag{1.2}$$

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

The reason you might not recognize the implication in (1.2) as familiar is that in Definition 1.1, (1.1), and (1.2), instead of the standard notation $y = f(x)$, we used the graph notation $(x, y) \in f$. The implication in (1.2) in the standard notation reads: For all $x_1, x_2 \in A$ the following implication holds: $f(x_1) \neq f(x_2) \;\Rightarrow\; x_1 \neq x_2$.

Definition 1.2. Let $f \subset A \times B$ be a function. The set A is said to be the **domain** of $f : A \to B$. The set B is said to be the **codomain** of $f : A \to B$. The set

$$\bigl\{ y \in B : \exists x \in A \text{ such that } (x, y) \in f \bigr\}$$

is called the **range** of $f : A \to B$. It is denoted by $\operatorname{ran}(f)$. ◊

Definition 1.3. Let $f \subset A \times B$ be a function. The function $f : A \to B$ is said to be a **surjection** if the following condition is satisfied

$$\forall y \in B \;\; \exists x \in A \quad \text{such that} \quad (x, y) \in f. \tag{\textbf{Surject}}$$

The function $f : A \to B$ is said to be an **injection** if the following condition is satisfied

$$\begin{array}{l}\forall x_1, x_2 \in A\\ \forall y_1, y_2 \in B\end{array} \quad (x_1, y_1) \in f \wedge (x_2, y_2) \in f \wedge x_1 \neq x_2 \;\Rightarrow\; y_1 \neq y_2. \tag{\textbf{Inject}}$$

◊

Definition 1.4. Let $f \subset A \times B$ be a function. The function $f : A \to B$ is said to be a **bijection** if it is both: a surjection and an injection. That is, $f \subset A \times B$ is a bijection if it satisfies four conditions: (**Total**), (**UniVal**), (**Surject**), and (**Inject**). ◊

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

Proposition 1.5. *Let $f : A \to B$ and $g : C \to D$ be functions. If $\operatorname{ran} f \subseteq C$, then*

$$\bigl\{ (x, z) \in A \times D : \exists\, y \in B \;\; (x, y) \in f \wedge (y, z) \in g \bigr\} \tag{1.3}$$

is a function from A to D.

Proof. A proof is a nice exercise. □

The function defined by (1.3) 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 $\mathrm{id}_A : A \to A$ such that $\mathrm{id}_A(x) = x$ for all $x \in A$.

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

Theorem 1.6. *Let $f : A \to B$ 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 $g : B \to A$ such that $f \circ g = \mathrm{id}_B$ and $g \circ f = \mathrm{id}_A$.*

If f is invertible, then the unique g whose existence is proved in Theorem 1.6 (c) is called the **inverse** of f; it is denoted by f^{-1}.

Let $f : A \to B$ be a function. It is common to extend the notation $f(x)$ for $x \in A$ to subsets of A. For $X \subseteq A$ we introduce the notation

$$f(X) = \{y \in B : \exists\, x \in X \;\; y = f(x)\}.$$

With this notation, the range of f is simply the set $f(A)$. It is also common to extend this notation to describe "inverse" image of a subset in B. For $Y \subseteq B$ we introduce the notation

$$f^{-1}(Y) = \{x \in A : f(x) \in Y\}.$$

Notice that this notation is used for arbitrary function f. It does not imply that f is invertible. Here f^{-1} is just a notational device.

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

Exercise 1.7. Let A, B and C be nonempty sets. Let $f : A \to B$ and $g : B \to C$ be injections. Prove that $g \circ f : A \to C$ is an injection. ♢

Exercise 1.8. Let A, B and C be nonempty sets. Let $f : A \to B$ and $g : B \to C$ be surjections. Prove that $g \circ f : A \to C$ is a surjection. ♢

Exercise 1.9. Let A, B and C be nonempty sets. Let $f : A \to B$ and $g : B \to C$ be bijections. Prove that $g \circ f : A \to C$ is a bijection. Prove that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. ♢

Exercise 1.10. Let A, B and C be nonempty sets. Let $f : A \to B$, $g : B \to C$. Prove that if $g \circ f$ is an injection, then f is an injection. ♢

Exercise 1.11. Let A, B and C be nonempty sets and let $f : A \to B$, $g : B \to C$. Prove that if $g \circ f$ is a surjection, then g is a surjection. ♢

Exercise 1.12. Let A, B and C be nonempty sets and let $f : A \to B$, $g : B \to C$ and $h : C \to A$ 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 $\mathscr{U}$, $\mathscr{V}$ and $\mathscr{W}$ are vector spaces over a scalar field $\mathbb{F}$.

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

$$\forall\, u \in \mathscr{V} \quad \forall\, v \in \mathscr{V} \qquad T(u+v) = T(u) + f(v), \tag{2.1}$$

$$\forall\, \alpha \in \mathbb{F} \quad \forall\, v \in \mathscr{V} \qquad T(\alpha v) = \alpha\, T(v). \tag{2.2}$$

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

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

$$(S+T)(v) = S(v) + T(v), \qquad \forall\, v \in \mathscr{V}, \tag{2.3}$$

$$(\alpha T)(v) = \alpha\, T(v), \qquad \forall\, v \in \mathscr{V}. \tag{2.4}$$

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 $\mathscr{W}$. Notice the analogous difference in empty spaces between α and T in (2.4). Define the zero mapping in $\mathscr{L}(\mathscr{V}, \mathscr{W})$ to be

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

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

$$(-T)(v) = -T(v), \qquad \forall\, v \in \mathscr{V}.$$

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

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

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

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

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

The definition of the linearity of a function between vector spaces is expressed in the standard functional notation. The next proposition states that a function between vector spaces is linear if and only if its graph is a subspace of the direct product of the domain and the codomain of that function.

Proposition 2.3. *Let $\mathscr{V}$ and $\mathscr{W}$ be vector spaces over a scalar field $\mathbb{F}$. Let $f : \mathscr{V} \to \mathscr{W}$ be a function and denote by $\mathscr{G}_f$ the graph of f; that is let*

$$\mathscr{G}_f = \bigl\{ \bigl(v, w\bigr) \in \mathscr{V} \times \mathscr{W} \ : \ v \in \mathscr{V} \ \text{and} \ \ w = f(v) \bigr\} \subseteq \mathscr{V} \times \mathscr{W}.$$

The function f is linear if and only if the set $\mathscr{G}_f$ is a subspace of the vector space $\mathscr{V} \times \mathscr{W}$.

Proposition 2.4. *Let $\mathscr{V}$ and $\mathscr{W}$ be vector spaces over a scalar field $\mathbb{F}$. Let $T \in \mathscr{L}(\mathscr{V}, \mathscr{W})$, let $\mathscr{G}$ be a subspace of $\mathscr{V}$ and let $\mathscr{H}$ be a subspace of $\mathscr{W}$. Then*

$$T(\mathscr{G}) = \bigl\{ w \in \mathscr{W} \ : \ \exists v \in \mathscr{G} \ \text{such that} \ \ w = Tv \bigr\}$$

is a subspace of $\mathscr{W}$ and

$$T^{-1}(\mathscr{H}) = \bigl\{ v \in \mathscr{V} \ : \ Tv \in \mathscr{H} \bigr\}$$

is a subspace of $\mathscr{V}$.

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 2.5. *Let $S : \mathscr{U} \to \mathscr{V}$ and $T : \mathscr{V} \to \mathscr{W}$ be linear operators. The composition $T \circ S : \mathscr{U} \to \mathscr{W}$ 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 $\mathscr{V}$ is denoted by $I_{\mathscr{V}}$. It is defined by $I_{\mathscr{V}}(v) = v$ for all $v \in \mathscr{V}$. It is clearly a linear operator.

Proposition 2.6. *Let $T : \mathscr{V} \to \mathscr{W}$ be a linear operator which is a bijection. Then the inverse $T^{-1} : \mathscr{W} \to \mathscr{V}$ of T is a linear operator.*

Proof. Since T is a bijection, from what we learned about function, there exists a function $S : \mathscr{W} \to \mathscr{V}$ such that $ST = I_{\mathscr{V}}$ and $TS = I_{\mathscr{W}}$. Since T is linear and $TS = I_{\mathscr{W}}$ we have

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

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

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

we get

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

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

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

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

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

By Proposition 2.6 each isomorphism is invertible and its inverse is also an isomorphism.

2.3. **The Coordination Operator.** The following theorem introduces a fundamental isomorphism between a finite-dimensional vector space $\mathscr{V}$ over $\mathbb{F}$ and $\mathbb{F}^m$, where $m = \dim \mathscr{V}$.

Theorem 2.7. *Let $\mathscr{V}$ be a finite-dimensional vector space over $\mathbb{F}$, let $m = \dim \mathscr{V}$, and let $\mathscr{B} = (v_1, \dots, v_m)$ be an ordered basis for $\mathscr{V}$. The function $C_{\mathscr{B}} : \mathscr{V} \to \mathbb{F}^m$ is defined for all $v \in \mathscr{V}$ as follows:*

$$C_{\mathscr{B}}(v) = \mathbf{a} \quad \textit{where} \quad \mathbf{a} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix} \in \mathbb{F}^m \quad \textit{and} \quad v = \alpha_1 v_1 + \cdots + \alpha_m v_m.$$

This function is an isomorphism between $\mathscr{V}$ and $\mathbb{F}^m$.

Proof. First, we express $C_{\mathscr{B}}$ in terms of its graph representation:

$$C_{\mathscr{B}} = \left\{ (v, \mathbf{a}) \in \mathscr{V} \times \mathbb{F}^m : v = \alpha_1 v_1 + \cdots + \alpha_m v_m \wedge \mathbf{a} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix} \right\}$$

To establish that $C_{\mathscr{B}}$ is a bijection, we prove the following four statements: (**Total**), (**UniVal**), (**Surject**), and (**Inject**).

A blueprint of the proof is as follows:

(1) $\mathscr{V} = \operatorname{span} \mathscr{B}$ implies that $C_{\mathscr{B}}$ is (**Total**).
(2) $\mathscr{B}$ is linearly independent implies that $C_{\mathscr{B}}$ is (**UniVal**).
(3) The axioms of a vector space **AE** and **SE** imply that $C_{\mathscr{B}}$ is (**Surject**). (This implication follows from the (**Total**) property of the additions and scaling functions.)
(4) The axioms of a vector space **AE** and **SE** imply that $C_{\mathscr{B}}$ is (**Inject**). (The implication in (**Inject**) follows from the (**UniVal**) properties of the additions and scaling functions.)

To prove that the bijection $C_{\mathscr{B}}$ is linear, we need to show that $C_{\mathscr{B}}$ is a subspace of $\mathscr{V} \times \mathbb{F}^m$. This is a straightforward verification. □

It is important to point out that the formula for the inverse function

$$(C_{\mathscr{B}})^{-1} : \mathbb{F}^m \to \mathscr{V}$$

is given by

$$(C_{\mathscr{B}})^{-1}\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_m\end{bmatrix}=\sum_{j=1}^{m}\alpha_j v_j \quad \text{for all} \quad \begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_m\end{bmatrix}\in\mathbb{F}^m. \tag{2.5}$$

Observe that equation (2.5) defines a function from $\mathbb{F}^m$ to $\mathscr{V}$ regardless of whether $\mathscr{B}$ forms a basis of $\mathscr{V}$.

Definition 2.8. The function $C_{\mathscr{B}}:\mathscr{V}\to\mathbb{F}^m$ introduced in Theorem 2.7 is called the **Coordination Operator**. $\Diamond$

Example 2.9. Inspired by the definition of $C_{\mathscr{B}}$ and (2.5), we define a general operator in this spirit. Let $\mathscr{V}$ and $\mathscr{W}$ be vector spaces over $\mathbb{F}$. Let $\mathscr{V}$ be finite-dimensional, $m=\dim\mathscr{V}$ and let $\mathscr{B}=(v_1,\ldots,v_m)$ be a basis for $\mathscr{V}$. Let $\mathscr{C}=(w_1,\ldots,w_m)$ be *any* m-tuple of vectors in $\mathscr{W}$. The entries of an m-tuple can be repeated, they can all be equal, for example equal to $0_{\mathscr{W}}$. We define the linear operator $L_{\mathscr{C}}^{\mathscr{B}}:\mathscr{V}\to\mathscr{W}$ by

$$L_{\mathscr{C}}^{\mathscr{B}}(v)=\sum_{j=1}^{m}\alpha_j w_j \qquad \text{where} \qquad \begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_m\end{bmatrix}=C_{\mathscr{B}}(v). \tag{2.6}$$

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

$$\begin{bmatrix}\xi_1\\ \vdots\\ \xi_m\end{bmatrix}\mapsto\sum_{j=1}^{m}\xi_j w_j \qquad \text{for all} \qquad \begin{bmatrix}\xi_1\\ \vdots\\ \xi_m\end{bmatrix}\in\mathbb{F}^m. \tag{2.7}$$

It is straightforward to verify that (2.7) defines a linear operator.

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

Exercise 2.10. Let $\mathscr{V}$ and $\mathscr{W}$ be vector spaces over $\mathbb{F}$. Let $\mathscr{V}$ be finite-dimensional, $m=\dim\mathscr{V}$ and let $\mathscr{B}$ be a basis for $\mathscr{V}$. Let $\mathscr{C}=(w_1,\ldots,w_m)$ be a list of vectors in $\mathscr{W}$ with m entries.

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

$\Diamond$

Proposition 2.11. *Let $\mathscr{V}$ and $\mathscr{W}$ be vector spaces over $\mathbb{F}$. Let $\mathscr{V}$ be finite-dimensional, $m=\dim\mathscr{V}$ and let $\mathscr{B}=(v_1,\ldots,v_m)$ be a basis for $\mathscr{V}$. For every $T\in\mathscr{L}(\mathscr{V},\mathscr{W})$ we have $T=L_{\mathscr{C}}^{\mathscr{B}}$ if and only if $\mathscr{C}=(Tv_1,\ldots,Tv_m)$.*

2.4. **The nullity-rank theorem.** Let $T : \mathscr{V} \to \mathscr{W}$ is be a linear operator. The linearity of T implies that the set

$$\operatorname{nul} T = \{v \in \mathscr{V} : Tv = 0_{\mathscr{W}}\}$$

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

$$\operatorname{ran} T = \{w \in \mathscr{W} : \exists\, v \in \mathscr{V} \quad w = Tv\}.$$

Proposition 2.12. *A linear operator $T : \mathscr{V} \to \mathscr{W}$ is an injection if and only if* $\operatorname{nul} T = \{0_{\mathscr{V}}\}$.

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

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

Theorem 2.13 (Nullity-Rank Theorem)**.** *Let $\mathscr{V}$ and $\mathscr{W}$ be vector spaces over a scalar field $\mathbb{F}$ and let $T : \mathscr{V} \to \mathscr{W}$ be a linear operator. If $\mathscr{V}$ is finite-dimensional, then* $\operatorname{nul} T$ *and* $\operatorname{ran} T$ *are finite-dimensional and*

$$\dim(\operatorname{nul} T) + \dim(\operatorname{ran} T) = \dim \mathscr{V}. \tag{2.8}$$

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

$$\mathscr{U} \oplus \mathscr{X} = \mathscr{V} \qquad \text{and} \qquad \dim \mathscr{U} + \dim \mathscr{X} = \dim \mathscr{V}.$$

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

$$(\operatorname{nul} T) \oplus \mathscr{X} = \mathscr{V} \qquad \text{and} \qquad \dim(\operatorname{nul} T) + \dim \mathscr{X} = \dim \mathscr{V}. \tag{2.9}$$

Since $\dim(\operatorname{nul} T) + \dim \mathscr{X} = \dim \mathscr{V}$, to prove the theorem we only need to prove that $\dim \mathscr{X} = \dim(\operatorname{ran} T)$. To this end, we consider the restriction $T|_{\mathscr{X}} : \mathscr{X} \to \operatorname{ran} T$ of T to the subspace $\mathscr{X}$. This operator is defined by

$$T|_{\mathscr{X}}(v) = Tv \quad \forall v \in \mathscr{X}.$$

We will prove that $T|_{\mathscr{X}}$ is an isomorphism. Let $\{x_1, \ldots, x_m\}$ be a basis for $\mathscr{X}$. To prove that $T|_{\mathscr{X}}$ is a surjection, we will prove

$$\operatorname{span}\{Tx_1, \ldots, Tx_m\} = \operatorname{ran} T. \tag{2.10}$$

Clearly $\{Tx_1, \ldots, Tx_m\} \subseteq \operatorname{ran} T$. Consequently, since $\operatorname{ran} T$ is a subspace of $\mathscr{W}$, we have $\operatorname{span}\{Tx_1, \ldots, Tx_m\} \subseteq \operatorname{ran} T$. To prove the converse inclusion, let $w \in \operatorname{ran} T$ be arbitrary. Then, there exists $v \in \mathscr{V}$ such that $Tv = w$. Since $\mathscr{V} = (\operatorname{nul} T) + \mathscr{X}$, there exist $u \in \operatorname{nul} T$ and $x \in \mathscr{X}$ such that $v = u + x$.

Then $Tv = T(u+x) = Tu+Tx = Tx$. As $x \in \mathscr{X}$, there exist $\xi_1, \ldots, \xi_m \in \mathbb{F}$ such that $x = \sum_{j=1}^{m} \xi_j x_j$. Now we use linearity of T to deduce

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

This proves that $w \in \operatorname{span}\{Tx_1, \ldots, Tx_m\}$. Since w was arbitrary in $\operatorname{ran} T$ this completes a proof of (2.10).

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

$$\alpha_1 T x_1 + \cdots + \alpha_m T x_m = 0_{\mathscr{W}}. \tag{2.11}$$

Since T is linear (2.11) implies that

$$\alpha_1 x_1 + \cdots + \alpha_m x_m \in \operatorname{nul} T. \tag{2.12}$$

Recall that $x_1, \ldots, x_m \in \mathscr{X}$ and $\mathscr{X}$ is a subspace of $\mathscr{V}$, so

$$\alpha_1 x_1 + \cdots + \alpha_m x_m \in \mathscr{X}. \tag{2.13}$$

Now (2.12), (2.13) and the fact that $(\operatorname{nul} T) \cap \mathscr{X} = \{0_{\mathscr{V}}\}$ imply

$$\alpha_1 x_1 + \cdots + \alpha_m x_m = 0_{\mathscr{V}}. \tag{2.14}$$

Since $x_1, \ldots, x_m$ are linearly independent, (2.14) yields $\alpha_j = 0$ for all $j \in \{1, \ldots, m\}$. This completes a proof of the linear independence of the vectors $Tx_1, \ldots, Tx_m$.

Thus $\{Tx_1, \ldots, Tx_m\}$ is a basis for $\operatorname{ran} T$. Consequently $\dim(\operatorname{ran} T) = m$. Since $m = \dim \mathscr{X}$, (2.9) implies (2.8). This completes the proof. □

A **direct** proof of the Nullity-Rank Theorem is as follows:

Proof. Since $\operatorname{nul} T$ is a subspace of $\mathscr{V}$ it is finite-dimensional. Set $k = \dim\bigl(\operatorname{nul} T\bigr)$ and let $\mathscr{C} = \bigl\{u_1, \ldots, u_k\bigr\}$ be a basis for $\operatorname{nul} T$.

Since $\mathscr{V}$ is finite-dimensional there exists a finite set $\mathscr{F} \subset \mathscr{V}$ such that $\operatorname{span}(\mathscr{F}) = \mathscr{V}$. Then the set $T\mathscr{F}$ is a finite subset of $\mathscr{W}$ and $\operatorname{ran} T = \operatorname{span}\bigl(T\mathscr{F}\bigr)$. Thus $\operatorname{ran} T$ is finite-dimensional. Let $\dim\bigl(\operatorname{ran} T\bigr) = m$ and let $\mathscr{G} = \bigl\{w_1, \ldots, w_m\bigr\}$ be a basis of $\operatorname{ran} T$.

Since clearly for every $j \in \bigl\{1, \ldots, m\bigr\}$, $w_j \in \operatorname{ran} T$, we have that for every $j \in \bigl\{1, \ldots, m\bigr\}$ there exists $v_j \in \mathscr{V}$ such that $Tv_j = w_j$. Set $\mathscr{D} = \bigl\{v_1, \ldots, v_m\bigr\}$.

Further set $\mathscr{B} = \mathscr{C} \cup \mathscr{D}$.

We will prove the following three facts:

(I) $\mathscr{C} \cap \mathscr{D} = \emptyset$,

(II) $\operatorname{span} \mathscr{B} = \mathscr{V}$,

(III) $\mathscr{B}$ is a linearly independent set.

To prove (I), notice that, since $\mathscr{G}$ is linearly independent, the vectors in $\mathscr{G}$ are nonzero. Therefore, for every $v \in \mathscr{D}$ we have that $Tv \neq 0_{\mathscr{W}}$. Since for every $u \in \mathscr{C}$ we have $Tu = 0_{\mathscr{W}}$ we conclude that $u \in \mathscr{C}$ implies $u \notin \mathscr{D}$. This proves (I).

Next we prove (II). By the definition of $\mathscr{B}$ we have $\mathscr{B} \subset \mathscr{V}$. Since $\mathscr{V}$ is a vector space, we have $\operatorname{span} \mathscr{B} \subseteq \mathscr{V}$.

To prove the converse inclusion, let $v \in \mathscr{V}$ be arbitrary. Then $Tv \in \operatorname{ran} T$. Since $\mathscr{G}$ spans $\operatorname{ran} T$, there exist $\beta_1, \ldots, \beta_m \in \mathbb{F}$ such that

$$Tv = \sum_{j=1}^{m} \beta_j w_j.$$

Set

$$v' = \sum_{j=1}^{m} \beta_j v_j.$$

Then, by linearity of T, we have

$$Tv' = \sum_{j=1}^{m} \beta_j T v_j = \sum_{j=1}^{m} \beta_j w_j = Tv.$$

The last equality and the linearity of T yield $T(v - v') = 0_{\mathscr{W}}$. Consequently, $v - v' \in \operatorname{nul} T$. Since $\mathscr{C}$ spans $\operatorname{nul} T$, there exist $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ such that

$$v - v' = \sum_{j=1}^{k} \alpha_i u_i.$$

Consequently,

$$v = v' + \sum_{j=1}^{k} \alpha_i u_i = \sum_{j=1}^{k} \alpha_i u_i + \sum_{j=1}^{m} \beta_j v_j.$$

This proves that for arbitrary $v \in \mathscr{V}$ we have $v \in \operatorname{span} \mathscr{B}$. Thus $\mathscr{V} \subseteq \operatorname{span} \mathscr{B}$ and (II) is proved.

To prove (III), let $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ and $\beta_1, \ldots, \beta_m \in \mathbb{F}$ be arbitrary and assume that

$$\sum_{j=1}^{k} \alpha_i u_i + \sum_{j=1}^{m} \beta_j v_j = 0_{\mathscr{V}}. \tag{2.15}$$

Applying T to both sides of the last equality, and using the fact that $u_i \in \operatorname{nul} T$ and the definition of v_j we get

$$\sum_{j=1}^{m} \beta_j w_j = 0_{\mathscr{W}}.$$

Since $\mathscr{E}$ is a linearly independent set the last equality implies that $\beta_j = 0$ for all $j \in \{1, \ldots, m\}$. Now substitute these equalities in (2.15) to get

$$\sum_{j=1}^{k} \alpha_i u_i = 0_{\mathscr{V}}.$$

Since $\mathscr{C}$ is a linearly independent set the last equality implies that $\alpha_i = 0$ for all $i \in \{1, \ldots, k\}$. This proves the linear independence of $\mathscr{B}$.

It follows from (II) and (III) that $\mathscr{B}$ is a basis for $\mathscr{V}$. By (I) we have that $\#\mathscr{B} = \#\mathscr{C} + \#\mathscr{D} = k + m$. This completes proof of the theorem. $\square$

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

The nullity-rank 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 2.14. *Let $\mathscr{V}$ and $\mathscr{W}$ be vector spaces over $\mathbb{F}$. Assume that $\mathscr{V}$ is finite-dimensional. The following statements are equivalent*

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

Proposition 2.15. *Let $\mathscr{V}$ and $\mathscr{W}$ be vector spaces over $\mathbb{F}$. Assume that $\mathscr{V}$ is finite-dimensional. The following statements are equivalent*

(a) *There exists an injection $T \in \mathscr{L}(\mathscr{V}, \mathscr{W})$.*
(b) *Either $\mathscr{W}$ is infinite-dimensional or $\dim \mathscr{V} \leq \dim \mathscr{W}$.*

Proposition 2.16. *Let $\mathscr{V}$ and $\mathscr{W}$ be vector spaces over $\mathbb{F}$. Assume that $\mathscr{V}$ is finite-dimensional. The following statements are equivalent*

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

3. Matrix of a linear operator

3.1. **A Natural Isomorphism between $\mathscr{L}(\mathscr{V}, \mathscr{W})$ and $\mathbb{F}^{n\times m}$.** Let $\mathscr{V}$ and $\mathscr{W}$ be nontrivial finite-dimensional vector spaces over $\mathbb{F}$, $m = \dim \mathscr{V}$, $n = \dim \mathscr{W}$, let $\mathscr{B} = \{v_1, \ldots, v_m\}$ be a basis for $\mathscr{V}$, and let $\mathscr{C} = \{w_1, \ldots, w_n\}$ be a basis for $\mathscr{W}$. The mapping $C_{\mathscr{B}}$ provides an isomorphism between $\mathscr{V}$ and $\mathbb{F}^m$ and $C_{\mathscr{C}}$ provides an isomorphism between $\mathscr{W}$ 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 A. 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 $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{F}^n$ be columns of an $n \times m$ matrix A. Then we write

$$\mathsf{A} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_m \end{bmatrix}.$$

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

$$\mathsf{A}\mathbf{x} = \sum_{j=1}^{m} \xi_j \mathbf{a}_j \qquad \text{where} \qquad \mathbf{x} = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_m \end{bmatrix}. \tag{3.1}$$

Notice the similarity of the definition in (3.1) to the definition (2.6) of the operator $L_{\mathscr{C}}^{\mathscr{B}}$ in Example 2.9. Taking $\mathscr{B}$ to be the standard basis $\mathscr{E}_m$ of $\mathbb{F}^m$ and taking $\mathscr{C}$ to be the m-tuple of columns of A, which are vectors in $\mathbb{F}^n$—call this m-tuple $\mathscr{A}$—we have $L_{\mathscr{A}}^{\mathscr{E}_m}(\mathbf{x}) = \mathsf{A}\mathbf{x}$.

In some sense, we identify the vector space $\mathbb{F}^{n\times m}$ with the vector space $\bigl(\mathbb{F}^n\bigr)^m$.

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

$$\begin{array}{ccc} \mathscr{V} & \xrightarrow{\quad T \quad} & \mathscr{W} \\ C_{\mathscr{B}} \Big\downarrow \Big\uparrow C_{\mathscr{B}}^{-1} & & \Big\downarrow C_{\mathscr{C}} \\ \mathbb{F}^m & \underset{\mathsf{A}}{\dashrightarrow} & \mathbb{F}^n \end{array} \tag{3.2}$$

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

$$\mathsf{A}\bigl(C_{\mathscr{B}}(v)\bigr) = C_{\mathscr{C}}(Tv) \qquad \forall\, v \in \mathscr{V}. \tag{3.3}$$

In English: multiplying the vector of coordinates of v by A we get exactly the coordinates of Tv. Or, starting from $\mathbf{x} \in \mathbb{F}^m$, we follow the arrows in the commutative diagram in (3.2), and calculate

$$\mathsf{A}\mathbf{x} = C_{\mathscr{C}}\Bigl(T\bigl(C_{\mathscr{B}}^{-1}(\mathbf{x})\bigr)\Bigr). \tag{3.4}$$

With $\mathbf{x} = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_m \end{bmatrix}$ and $\mathscr{B} = (v_1, \ldots, v_m)$ in (3.4) we have that

$$\begin{aligned} \mathsf{A}\mathbf{x} &= C_{\mathscr{C}}\Bigl(T\bigl(C_{\mathscr{B}}^{-1}(\mathbf{x})\bigr)\Bigr) \\ &= C_{\mathscr{C}}\left(T\left(\sum_{j=1}^{m} \xi_j v_j\right)\right) \\ &= C_{\mathscr{C}}\left(\sum_{j=1}^{m} \xi_j T v_j\right) \\ &= \sum_{j=1}^{m} \xi_j C_{\mathscr{C}}(Tv_j). \end{aligned}$$

The preceding sequence of equalities and the definition of the matrix-vector multiplication shows that

$$\mathsf{A} = \Bigl[C_{\mathscr{C}}(Tv_1) \ \cdots \ C_{\mathscr{C}}(Tv_m) \Bigr]. \tag{3.5}$$

The matrix A defined in (3.5) has the desired property stated in (3.3). With this A the diagram in (3.2) is commutative.

For an arbitrary $T \in \mathscr{L}(\mathscr{V},\mathscr{W})$ the formula (3.5) associates the matrix $\mathsf{A} \in \mathbb{F}^{n\times m}$ with T. In other words (3.5) defines a function from $\mathscr{L}(\mathscr{V},\mathscr{W})$ to $\mathbb{F}^{n\times m}$. We formally define this function in the following theorem.

Theorem 3.1. *Let $\mathscr{V}$ and $\mathscr{W}$ be finite-dimensional vector spaces over $\mathbb{F}$, $m = \dim\mathscr{V}$, $n = \dim\mathscr{W}$, let $\mathscr{B} = \{v_1,\ldots,v_m\}$ be a basis for $\mathscr{V}$ and let $\mathscr{C} = \{w_1,\ldots,w_n\}$ be a basis for $\mathscr{W}$. The function*

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

defined by

$$M_{\mathscr{C}}^{\mathscr{B}}(T) = \Big[C_{\mathscr{C}}(Tv_1) \ \cdots \ C_{\mathscr{C}}(Tv_m)\Big] \quad \text{for all} \quad T \in \mathscr{L}(\mathscr{V},\mathscr{W}) \tag{3.6}$$

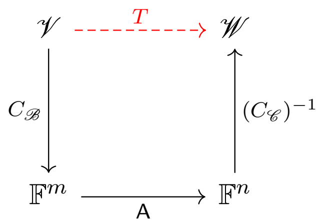
is an isomorphism.

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

Since the definition of $M_{\mathscr{C}}^{\mathscr{B}}(T)$ coincides with (3.5), equality (3.3) yields

$$C_{\mathscr{C}}(Tv) = \big(M_{\mathscr{C}}^{\mathscr{B}}(T)\big)C_{\mathscr{B}}(v). \tag{3.7}$$

The most direct way to prove that $M_{\mathscr{C}}^{\mathscr{B}}$ is an isomorphism is to construct its inverse. The inverse is suggested by the diagram (3.8) below. In the following diagram, T is the only unknown:



(3.8)

Define

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

by

$$\big(N_{\mathscr{C}}^{\mathscr{B}}(\mathsf{A})\big)(v) = (C_{\mathscr{C}})^{-1}\big(\mathsf{A}C_{\mathscr{B}}(v)\big), \quad \text{for all} \quad \mathsf{A} \in \mathbb{F}^{n\times m}. \tag{3.9}$$

Next we prove that

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

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

$$\begin{aligned}
\Big(\big(N_{\mathscr{C}}^{\mathscr{B}} \circ M_{\mathscr{C}}^{\mathscr{B}}\big)(T)\Big)(v) &= (C_{\mathscr{C}})^{-1}\Big(\big(M_{\mathscr{C}}^{\mathscr{B}}(T)\big)(C_{\mathscr{B}}(v))\Big) && \text{by (3.9)} \\
&= (C_{\mathscr{C}})^{-1}\big(C_{\mathscr{C}}(Tv)\big) && \text{by (3.7)} \\
&= Tv.
\end{aligned}$$

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

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

$$\big(M_{\mathscr{C}}^{\mathscr{B}} \circ N_{\mathscr{C}}^{\mathscr{B}}\big)(\mathsf{A}) = M_{\mathscr{C}}^{\mathscr{B}}\big(N_{\mathscr{C}}^{\mathscr{B}}(\mathsf{A})\big)$$

$$
\begin{aligned}
\boxed{\text{by (3.6)}} &= \Big[C_{\mathscr{C}}\big(\big(N_{\mathscr{C}}^{\mathscr{B}}(\mathsf{A})\big)(v_1)\big) \ \cdots \ C_{\mathscr{C}}\big(\big(N_{\mathscr{C}}^{\mathscr{B}}(\mathsf{A})\big)(v_m)\big) \Big] \\
\boxed{\text{by (3.9)}} &= \Big[\mathsf{A} C_{\mathscr{B}}(v_1) \ \cdots \ \mathsf{A} C_{\mathscr{B}}(v_m) \Big] \\
\boxed{\text{matrix mult.}} &= \mathsf{A}\Big[C_{\mathscr{B}}(v_1) \ \cdots \ C_{\mathscr{B}}(v_m) \Big] \\
\boxed{\text{def. of } C_{\mathscr{B}}} &= \mathsf{A}\,\mathsf{I}_m \\
\boxed{\text{matrix mult.}} &= \mathsf{A}.
\end{aligned}
$$

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

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

The following corollary is a special case of (3.7).

Corollary 3.2. *Let $m \in \mathbb{N}$ and let $\mathscr{V}$ be finite-dimensional vector spaces over $\mathbb{F}$ with $m = \dim \mathscr{V}$. Let $\mathscr{B}$ and $\mathscr{C}$ be bases for $\mathscr{V}$. Then*

$$
\forall v \in \mathscr{V} \quad C_{\mathscr{C}}(v) = \big(M_{\mathscr{C}}^{\mathscr{B}}(I_{\mathscr{V}})\big) C_{\mathscr{B}}(v)
$$

Definition 3.3. In the setting of Corollary 3.2, the $m \times m$ matrix $M_{\mathscr{C}}^{\mathscr{B}}(I_{\mathscr{V}})$ is called *change of coordinates matrix*, or *change of bases matrix* ◊

Theorem 3.4. *Let $k, m, n \in \mathbb{N}$, let $\mathscr{U}$, $\mathscr{V}$ and $\mathscr{W}$ be finite-dimensional vector spaces over $\mathbb{F}$, $k = \dim \mathscr{U}$, $m = \dim \mathscr{V}$, $n = \dim \mathscr{W}$, let $\mathscr{A}$ be a basis for $\mathscr{U}$, let $\mathscr{B}$ be a basis for $\mathscr{V}$, and let $\mathscr{C}$ be a basis for $\mathscr{W}$. Let $S \in \mathscr{L}(\mathscr{U}, \mathscr{V})$ and $T \in \mathscr{L}(\mathscr{V}, \mathscr{W})$. Let $M_{\mathscr{B}}^{\mathscr{A}}(S) \in \mathbb{F}^{m\times k}$, $M_{\mathscr{C}}^{\mathscr{B}}(T) \in \mathbb{F}^{n\times m}$ and $M_{\mathscr{C}}^{\mathscr{A}}(TS) \in \mathbb{F}^{n\times k}$ be as defined in Theorem 3.1. Then*

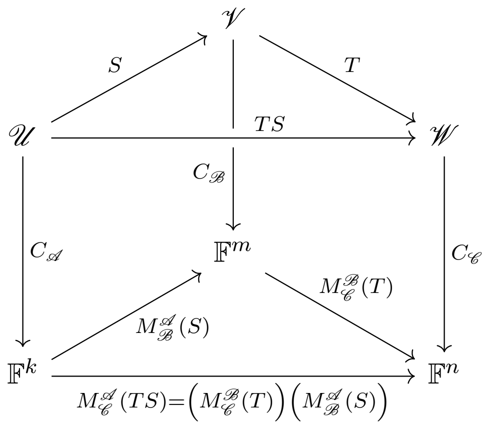
$$
M_{\mathscr{C}}^{\mathscr{A}}(TS) = \big(M_{\mathscr{C}}^{\mathscr{B}}(T)\big)\big(M_{\mathscr{B}}^{\mathscr{A}}(S)\big).
$$

Proof. Let $\mathscr{A} = (u, \ldots, u_k)$ and calculate, first using the definition in (3.6), then the boxed tools

$$
\begin{aligned}
M_{\mathscr{C}}^{\mathscr{A}}(TS) &= \Big[C_{\mathscr{C}}(TSu_1) \ \cdots \ C_{\mathscr{C}}(TSu_k) \Big] \\
\boxed{\text{by (3.7)}} &= \Big[\big(M_{\mathscr{C}}^{\mathscr{B}}(T)\big)\big(C_{\mathscr{B}}(Su_1)\big) \ \cdots \ \big(M_{\mathscr{C}}^{\mathscr{B}}(T)\big)\big(C_{\mathscr{B}}(Su_k)\big) \Big] \\
\boxed{\text{matrix mult.}} &= \big(M_{\mathscr{C}}^{\mathscr{B}}(T)\big)\Big[C_{\mathscr{B}}(Su_1) \ \cdots \ C_{\mathscr{B}}(Su_k) \Big] \\
\boxed{\text{by (3.6)}} &= \big(M_{\mathscr{C}}^{\mathscr{B}}(T)\big)\big(M_{\mathscr{B}}^{\mathscr{A}}(S)\big).
\end{aligned}
$$

□

The following diagram illustrates the content of Theorem 3.4.



Definition 3.5. Let $\mathscr{V}$ and $\mathscr{W}$ be vector spaces over $\mathbb{F}$. Operators $S \in \mathscr{L}(\mathscr{V})$ and $T \in \mathscr{L}(\mathscr{W})$ are said to be *similar* if there exist an isomorphisms $\Phi : \mathscr{V} \to \mathscr{W}$ such that

$$T = \Phi S \Phi^{-1}.$$

$\lozenge$

Definition 3.6. Let $n \in \mathbb{N}$ and let A and B be $n \times n$ matrices. Matrices A and B are said to be *similar* if there exists an invertible $n \times n$ matrix P such that $\mathsf{B} = P\mathsf{A}P^{-1}$. $\lozenge$

Theorem 3.7. *Let $n \in \mathbb{N}$ and let $\mathscr{V}$ and $\mathscr{W}$ be finite-dimensional vector spaces over $\mathbb{F}$ such that $n = \dim \mathscr{V} = \dim \mathscr{W}$. Let $S \in \mathscr{L}(\mathscr{V})$ and $T \in \mathscr{L}(\mathscr{W})$. The following statements are equivalent.*

(i) *The operators S and T are similar.*

(ii) *For every basis $\mathscr{A}$ of $\mathscr{V}$ (respectively, every basis $\mathscr{B}$ of $\mathscr{W}$), there exists a basis $\mathscr{B}$ of $\mathscr{W}$ (respectively, a basis $\mathscr{A}$ of $\mathscr{V}$) such that*

$$M_{\mathscr{A}}^{\mathscr{A}}(S) = M_{\mathscr{B}}^{\mathscr{B}}(T). \tag{3.10}$$

(iii) *For every basis $\mathscr{A}$ of $\mathscr{V}$ and every basis $\mathscr{B}$ of $\mathscr{W}$ the $n \times n$ matrices $M_{\mathscr{A}}^{\mathscr{A}}(S)$ and $M_{\mathscr{B}}^{\mathscr{B}}(T)$ are similar.*

4. Problems

Problem 4.1. Let $\mathscr{V}$ and $\mathscr{W}$ be vector spaces over a scalar field $\mathbb{F}$. Let $\mathscr{S}$ be a subspace of the direct product vector space $\mathscr{V} \times \mathscr{W}$, let $\mathscr{G}$ be a subspace of $\mathscr{V}$ and let $\mathscr{H}$ be a subspace of $\mathscr{W}$. Then

$$\mathscr{S}(\mathscr{G}) = \big\{ w \in \mathscr{W} \ : \exists v \in \mathscr{G} \text{ such that } \ (v,w) \in \mathscr{S} \big\}$$

is a subspace of $\mathscr{W}$ and

$$\mathscr{S}^{-1}(\mathscr{H}) = \big\{ v \in \mathscr{V} \ : \exists w \in \mathscr{H} \text{ such that } \ (v,w) \in \mathscr{S} \big\}$$

is a subspace of $\mathscr{V}$. $\lozenge$

Problem 4.2. Let $\mathscr{V}$ and $\mathscr{W}$ be finite-dimensional vector spaces over a scalar field $\mathbb{F}$. Let $\mathscr{S}$ be a subspace of the direct product vector space $\mathscr{V}\times\mathscr{W}$. The following four sets are subspaces

$$\begin{aligned}\operatorname{dom}\mathscr{S} &= \bigl\{v\in\mathscr{V} : \exists\, w\in\mathscr{W} \text{ such that } (v,w)\in\mathscr{S}\bigr\},\\ \operatorname{ran}\mathscr{S} &= \bigl\{w\in\mathscr{W} : \exists\, v\in\mathscr{V} \text{ such that } (v,w)\in\mathscr{S}\bigr\},\\ \operatorname{nul}\mathscr{S} &= \bigl\{v\in\mathscr{V} : \ (v,0_{\mathscr{W}})\in\mathscr{S}\bigr\},\\ \operatorname{mul}\mathscr{S} &= \bigl\{w\in\mathscr{W} : \ (0_{\mathscr{V}},w)\in\mathscr{S}\bigr\}.\end{aligned}$$

and the following equality holds:

$$\dim\operatorname{dom}\mathscr{S} + \dim\operatorname{mul}\mathscr{S} = \dim\operatorname{ran}\mathscr{S} + \dim\operatorname{nul}\mathscr{S}.$$

Hint: The following equivalence holds. For all $v\in\mathscr{V}$ and all $w\in\mathscr{W}$ we have:

$$(v,w)\in\mathscr{S} \quad\Leftrightarrow\quad (v+x,w+y)\in\mathscr{S} \quad \forall x\in\operatorname{nul}\mathscr{S} \text{ and } \forall y\in\operatorname{mul}\mathscr{S}.$$

◊

Problem 4.3. Let $\mathscr{V}$ and $\mathscr{W}$ be finite-dimensional vector spaces over a scalar field $\mathbb{F}$ and recall that $\mathscr{V}\times\mathscr{W}$ and $\mathscr{W}\times\mathscr{V}$ are the direct product vector spaces. Prove that the function

$$R:\mathscr{V}\times\mathscr{W}\to\mathscr{W}\times\mathscr{V}$$

defined by

$$R(v,w)=(w,v) \quad \text{for all} \quad (v,w)\in\mathscr{V}\times\mathscr{W}$$

is an isomorphism. ◊

Problem 4.4. Let $\mathscr{V}$ and $\mathscr{W}$ be finite-dimensional vector spaces over a scalar field $\mathbb{F}$ and recall that $\mathscr{V}\times\mathscr{W}$ and $\mathscr{W}\times\mathscr{V}$ are the direct product vector spaces. Let $\mathscr{T}$ be a subset of $\mathscr{V}\times\mathscr{W}$. Then $\mathscr{T}$ is an isomorphism between $\mathscr{V}$ and $\mathscr{W}$ if and only if the set

$$\Bigl\{(w,v)\in\mathscr{W}\times\mathscr{V} : (v,w)\in\mathscr{T}\Bigr\} = R\mathscr{T}$$

is an isomorphism between $\mathscr{W}$ and $\mathscr{V}$. (Use Problem 4.3 and Propositions 2.3 and 2.4 to prove this equivalence.) ◊

Problem 4.5. This problem explores the vector space $\mathscr{P}_3=\mathbb{R}[x]_{<4}$ of polynomials od degree at most 3 with real coefficients. Recall that the standard basis for this vector space is the basis which consists of monomials:

$$\mathscr{M}=\bigl(1,x,x^2,x^3\bigr).$$

(I) Consider the linear operator $T\in\mathscr{L}\bigl(\mathscr{P}_3,\mathscr{P}_3\bigr)=\mathscr{L}\bigl(\mathscr{P}_3\bigr)$ defined by

$$\forall p\in\mathscr{P}_3 \quad (Tp)(x)=p(-1)+p(-3)x+p(1)x^2+p(3)x^3.$$

(a) Find the 4×4 matrix $\mathsf{A}=M_{\mathscr{M}}^{\mathscr{M}}(T)$, the matrix of T relative to the basis of monomials $\mathscr{M}$.

(b) Find the inverse of the matrix A considering it as the matrix of the linear operator T^{-1}. That is, find $M_{\mathscr{M}}^{\mathscr{M}}(T^{-1})$ by using logic of polynomials, not using matrix calculations.

(c) The operator T has four eigenvalues. Determine these eigenvalues and the corresponding eigen-polynomials. That is, find the real numbers λ_1, λ_2, λ_3, λ_4 and the corresponding polynomials $p_1(x)$, $p_2(x)$, $p_3(x)$, $p_4(x)$ in $\mathscr{P}_3$ such that

$$\forall j \in \{1,2,3,4\} \quad (Tp_j)(x) = \lambda_j p_j(x).$$

(II) Instead of the special numbers $-1, -3, 1, 3 \in \mathbb{R}$ used in (I), consider an arbitrary

$$\mathbf{c} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{bmatrix} \in \mathbb{R}^4,$$

and define $L_{\mathbf{c}} \in \mathscr{L}\big(\mathscr{P}_3\big)$ by

$$\forall p \in \mathscr{P}_3 \quad \big(L_{\mathbf{c}} p\big)(x) = p(\gamma_1) + p(\gamma_2)x + p(\gamma_3)x^2 + p(\gamma_4)x^3.$$

(a) Consider the function $\Psi : \mathbb{R}^4 \to \mathscr{L}\big(\mathscr{P}_3\big)$ defined by $\Psi(\mathbf{c}) = L_{\mathbf{c}}$. Is this function linear? Is this function a surjection? Is this function an injection?

(b) Characterize $\mathbf{c} \in \mathbb{R}^4$ for which $L_{\mathbf{c}} \in \mathscr{L}\big(\mathscr{P}_3\big)$ is an isomorphism.

(c) Does there exist $\mathbf{c}, \mathbf{d} \in \mathbb{R}^4$ such that $L_{\mathbf{c}} L_{\mathbf{d}} = I_{\mathscr{P}_3}$?

$\lozenge$

Problem 4.6. Let $n \in \mathbb{N}$ and denote by $\mathscr{P}_n = \mathbb{R}[x]_{\leq n}$ the vector space of all polynomials of degree at most n with coefficients in $\mathbb{R}$.

Denote by $D = \frac{\mathrm{d}}{\mathrm{d}x}$ the differentiation operator defined on $\mathscr{P}_n$.

Let $t \in \mathbb{R}$. Denote by S_t the shift operator defined on $\mathscr{P}_n$ by

$$\forall f \in \mathscr{P}_n \qquad (S_t f)(x) = f(x+t).$$

We have proved that $D, S_t \in \mathscr{L}\big(\mathscr{P}_n, \mathscr{P}_n\big) = \mathscr{L}\big(\mathscr{P}_n\big)$. All statements below are in the vector space $\mathscr{L}\big(\mathscr{P}_n\big)$. The operators D, S_t depend on n as everything else in this problem. However, for conciseness, we choose not to emphasise that by indexing them by n.

(a) Prove that for all $t \in \mathbb{R}$ we have

$$S_t = \sum_{k=0}^{n} \frac{1}{k!} t^k D^k. \tag{4.1}$$

While this is not exclusively a linear algebra task, it is an important mathematical fact with significant linear algebraic consequences.

(b) The preceding result implies that

$$\forall t \in \mathbb{R} \qquad S_t \in \operatorname{span}\Big\{ D^k : k \in \mathbb{N} \cup \{0\} \Big\},$$

where $D^0 \stackrel{\text{def}}{=} I_{\mathscr{P}_n}$. Consequently,

$$\operatorname{span}\Big\{S_t : t \in \mathbb{R}\Big\} \subseteq \operatorname{span}\Big\{D^k : k \in \mathbb{N} \cup \{0\}\Big\}.$$

Does the converse inclusion hold?

(c) Set

$$\mathscr{D} = \operatorname{span}\Big\{D^k : k \in \mathbb{N} \cup \{0\}\Big\}, \quad \mathscr{S} = \operatorname{span}\Big\{S_t : t \in \mathbb{R}\Big\}.$$

Find a basis for each of these spaces consisting of the operators that are being spanned. Be specific and as thorough as possible, in particular when discussing $\mathscr{S}$.

(d) In formula (4.1), a shift operator is expressed as a specific linear combination of the powers of the derivative operator. Do analogous converse formulas exist? That is, can the powers of the differentiation operator be expressed as specific linear combinations of shift operators, or as linear combinations of the powers of a fixed shift operator?

$\Diamond$

Problem 4.7. In an earlier problem we considered special subspaces

$$\mathscr{S}_\omega \stackrel{\text{def}}{=} \Big\{f \in \mathbb{R}^{\mathbb{R}} : \exists\, a, b \in \mathbb{R} \ \text{ such that } \ f(x) = a\sin(\omega x + b), \ t \in \mathbb{R}\Big\}.$$

of the vector space $\mathbb{R}^{\mathbb{R}}$ of all real valued functions defined on $\mathbb{R}$. Here ω is an arbitrary fixed nonzero real number. We proved that

$$\mathscr{T}_\omega = \big(\cos(\omega x), \sin(\omega x)\big)$$

is a basis of $\mathscr{S}_\omega$.

(a) Denote by $D = \frac{d}{dx}$ the differentiation operator defined on $\mathscr{S}_\omega$. Prove that $D \in \mathscr{L}\big(\mathscr{S}_\omega, \mathscr{S}_\omega\big)$ and calculate $M_{\mathscr{T}_\omega}^{\mathscr{T}_\omega}(D)$.

(b) Let $t \in \mathbb{R}$. Denote by S_t the shift operator defined on $\mathscr{S}_\omega$ by

$$\forall\, f \in \mathscr{S}_\omega \qquad (S_t f)(x) = f(x+t).$$

Prove that $S_t \in \mathscr{L}\big(\mathscr{S}_\omega, \mathscr{S}_\omega\big)$ and calculate $M_{\mathscr{T}_\omega}^{\mathscr{T}_\omega}(S_t)$.

(c) Prove that

$$\frac{d}{dt}\Big(M_{\mathscr{T}_\omega}^{\mathscr{T}_\omega}(S_t)\Big) = M_{\mathscr{T}_\omega}^{\mathscr{T}_\omega}(D) M_{\mathscr{T}_\omega}^{\mathscr{T}_\omega}(S_t).$$

(d) Set

$$\mathsf{D} = M_{\mathscr{T}_\omega}^{\mathscr{T}_\omega}(D).$$

Prove that

$$M_{\mathscr{T}_\omega}^{\mathscr{T}_\omega}(S_t) = \sum_{k=0}^{\infty} \frac{1}{k!} t^k \mathsf{D}^k.$$

$\Diamond$

Problem 4.8. Let $n \in \mathbb{N}$ and denote by $\mathscr{P}_n = \mathbb{R}[x]_{\leq n}$ the vector space of all polynomials of degree less or equal to n. Denote by

$$\mathscr{M}_n = \big(1, x, \ldots, x^n\big)$$

the basis of $\mathscr{P}_n$ consisting of the monomials in $\mathscr{P}_n$.

(a) Denote by $D = \frac{d}{dx}$ the differentiation operator defined on $\mathscr{P}_n$. Prove that $D \in \mathscr{L}\big(\mathscr{P}_n, \mathscr{P}_n\big)$ and calculate $M^{\mathscr{M}_n}_{\mathscr{M}_n}(D)$.

(b) Let $t \in \mathbb{R}$. Denote by S_t the shift operator defined on $\mathscr{P}_n$ by

$$\forall f \in \mathscr{P}_n \qquad (S_t f)(x) = f(x+t).$$

Prove that $S_t \in \mathscr{L}\big(\mathscr{P}_n, \mathscr{P}_n\big)$ and calculate $M^{\mathscr{M}_n}_{\mathscr{M}_n}(S_t)$.

(c) Prove that

$$\frac{d}{dt}\Big(M^{\mathscr{M}_n}_{\mathscr{M}_n}(S_t)\Big) = M^{\mathscr{M}_n}_{\mathscr{M}_n}(D) M^{\mathscr{M}_n}_{\mathscr{M}_n}(S_t).$$

(d) Set

$$\mathsf{D} = M^{\mathscr{M}_n}_{\mathscr{M}_n}(D).$$

Prove that

$$M^{\mathscr{P}_n}_{\mathscr{P}_n}(S_t) = \sum_{k=0}^{n} \frac{1}{k!} t^k \mathsf{D}^k.$$

◊

Remark 4.9. The content of items (c) and (d) in Problems 4.7 and 4.8 indicates that the formula

$$e^{tD} = S_t$$

makes sense on the space of all differentiable functions.

◊