This chapter briefly presents the algebraic structure of vector spaces starting with the rudimentary concepts of groups and fields. This chapter should be read along with Chapter 4 Part A of Naylor & Sell. Both solved examples and exercises in Naylor & Sell would be very useful.

1 Monoids and Groups

Let \( S \) be a nonempty set. With \( m \in \mathbb{N} \), an \( m \)-ary operation \( \ast \) is a function \( \ast : S^m \to S \), i.e., mapping \( S \times S \times \cdots S \) \( m \) times into \( S \). Most commonly encountered operations are binary, i.e., \( m = 2 \), where \( \ast : S \times S \to S \).

This section focuses on binary algebras that are algebraic systems with a single binary operation. We first move from a primitive algebra called semigroup to monoid, and then to a highly structured algebraic system called group. Concepts of groups are extensively used in Physics and Engineering.

Definition 1.1. (Binary Algebra) Let \( S \) be a nonempty set. The operation \( \ast : S \times S \to S \) is called a binary algebra and is referred to as \((S; \ast)\).

Definition 1.2. (Semigroup) A binary algebra \((S; \ast)\) is called a semigroup if it satisfies the associativity property, i.e., \((\alpha \ast \beta) \ast \gamma = \alpha \ast (\beta \ast \gamma) \quad \forall \alpha, \beta, \gamma \in S\). A semigroup is called commutative if \(\alpha \ast \beta = \beta \ast \alpha \quad \forall \alpha, \beta \in S\).

Definition 1.3. (Identity Element) An element \( 1 \ell \in S \) is said to be a left identity of the binary algebra \((S; \ast)\) if \( 1 \ell \ast \alpha = \alpha \quad \forall \alpha \in S \). Similarly, an element \( 1 r \in S \) is said to be a right identity of the binary algebra \((S; \ast)\) if \( \alpha \ast 1 r = \alpha \quad \forall \alpha \in S \). Finally, an element \( 1 \in S \) is said to be an identity of the binary algebra \((S; \ast)\) if \( 1 \ast \alpha = \alpha = \alpha \ast 1 \quad \forall \alpha \in S \).

Definition 1.4. (Zero Element) An element \( 0 \ell \in S \) is said to be a left zero of the binary algebra \((S; \ast)\) if \( 0 \ell \ast \alpha = 0 \ell \quad \forall \alpha \in S \). Similarly, an element \( 0 r \in S \) is said to be a right zero of the binary algebra \((S; \ast)\) if \( \alpha \ast 0 r = 0 r \quad \forall \alpha \in S \). Finally, an element \( 0 \in S \) is said to be a zero of the binary algebra \((S; \ast)\) if \( 0 \ast \alpha = \alpha = \alpha \ast 0 \quad \forall \alpha \in S \).

Definition 1.5. (Monoid) A semigroup with an identity element is called a monoid. That is, a monoid is a binary algebra that satisfies the associativity property and has an identity element. A monoid is called commutative if \(\alpha \ast \beta = \beta \ast \alpha \quad \forall \alpha, \beta \in S\).
Definition 1.6. (Invertible Element) Let \((S; \oplus)\) be a monoid with an identity element \(1\). Then, an element \(a \in S\) is called left invertible if there exists \(a_l^{-1} \in S\) such that \(a_l^{-1} \oplus a = 1\). Similarly, an element \(a \in S\) is called right invertible if there exists \(a_r^{-1} \in S\) such that \(a \oplus a_r^{-1} = 1\). An element \(a \in S\) is called invertible if there exists \(a^{-1} \in S\) such that \(a^{-1} \oplus a = a \oplus a^{-1} = 1\).

Definition 1.7. (Group) A monoid whose every element is invertible is called a group. That is, a binary algebra \((S; \oplus)\) is called a group if the following properties are satisfied:

- **Closure Property:** \(a \oplus b \in S\) \(\forall a, b \in S\)
- **Associativity Property:** \((a \oplus b) \oplus c = a \oplus (b \oplus c)\) \(\forall a, b, c \in S\).
- **Existence of an Identity:** There exists \(1 \in S\) such that \(1 \oplus a = a \oplus 1 = a\) \(\forall a \in S\).
- **Existence of an Inverse for each Element:** There exists \(a^{-1} \in S\) such that \(a^{-1} \oplus a = a \oplus a^{-1} = 1\) \(\forall a \in S\).

Definition 1.8. (Abelian Group) A group \((S; \oplus)\) is called commutative (also called Abelian) if the following additional property: \(a \oplus b = b \oplus a\) \(\forall a, b \in S\) is satisfied. It is customary to denote the binary operator \(\oplus\) as + for commutative groups. That is, \((S, +)\) indicates an Abelian group. The inverse of an element \(a\) is denoted as \(-a\) and the identity as \(0\).

Example 1.1. Let \(S = \mathbb{R} \triangleq (-\infty, \infty)\) and let + be the usual addition operation. Then, \((S, +)\) is an Abelian group whose identity element is 0.

Example 1.2. Let \(S = \mathbb{R} \triangleq (-\infty, \infty) \setminus \{0\}\) and let \(\cdot\) be the usual multiplication operation. Then, \((S, \cdot)\) is an Abelian group whose identity element is 1.

Example 1.3. Let \(S\) be the set of all \(n \times n\) invertible matrices with real entries. Let \(\oplus\) be the operation of matrix multiplication. Then, \((S; \oplus)\) is a group whose identity element is the \(n \times n\) identity matrix, but it is NOT Abelian. Note that if \(M\) is the set of all \(n \times n\) matrices with real entries, then \((M; \oplus)\) is not a group because some of the matrices in \(M\) are not invertible; however, \((M; \circ)\) is a monoid.

Example 1.4. Let \(S\) be the set of all \(m \times n\) matrices (\(m\) is not necessarily equal to \(n\)) with real entries. Let \(\oplus\) be the operation of matrix addition. Then, \((S; \oplus)\) is an Abelian group whose identity element is the \((m \times n)\) zero matrix.

Definition 1.9. (Subgroup) Let \((S; \oplus)\) be a group and let \(\bar{S} \subseteq S\) be closed under \(\oplus\), i.e., \(\oplus: \bar{S} \times \bar{S} \rightarrow \bar{S}\). If \((\bar{S}, \oplus)\) satisfies the properties of a group as delineated in Definition 1.7, then \((\bar{S}, \oplus)\) is called a subgroup of the group \((S; \oplus)\). In this case, if \(\bar{S} \subset S\), i.e., \(\bar{S}\) is a proper subset of \(S\), then \((\bar{S}, \oplus)\) is called a proper subgroup of the group \((S; \oplus)\). If \(\bar{S} = S\) or if \(\bar{S} = \{1\}\), where \(1\) is the identity element of \(S\), then \((\bar{S}, \oplus)\) is called a trivial subgroup of \((S; \oplus)\). The largest proper subgroup of \((S; \oplus)\) is called the maximal subgroup.

Example 1.5. \((\mathbb{R}, +)\) is a proper subgroup of \((\mathbb{C}, +)\) and \((\mathbb{Q}, +)\) is a proper subgroup of \((\mathbb{R}, +)\).
Definition 1.10. (Cosets and Normal Subgroup) Let \((G; \oplus)\) be a group and \((S; \oplus)\) be a subgroup of \((G; \oplus)\). If \(a \in G\), then the set \(S \oplus a \triangleq \{ x = b \oplus a : b \in S \}\) is called a right coset of \(S\) in \((G; \oplus)\) and the set \(a \oplus S \triangleq \{ x = a \oplus b : b \in S \}\) is called a left coset of \(S\) in \((G; \oplus)\). If, for all \(a \in G\), we have \(S \oplus a = a \oplus S\), then \((S; \oplus)\) is called a normal subgroup of \((G; \oplus)\), and the right and left cosets are simply called cosets.

Remark 1.1. Every subgroup of an Abelian group is a normal subgroup.

2 Rings and Fields

This section focuses on rings and fields that deal with two binary operations. Additional algebraic structures are imposed on rings to construct fields.

Definition 2.1. (Ring) Let \((S; +, \cdot)\) be an algebraic system, i.e., \(+ : S \times S \to S\) and \(\cdot : S \times S \to S\). Then, \((S; +, \cdot)\) is called a ring if \((S; +)\) is an Abelian group (with \(0 \in S\) as the additive identity element) and if, for every \(\alpha, \beta, \gamma \in S\), the following conditions are satisfied:

- **Associative property**: \((\alpha \cdot \beta) \cdot \gamma = \alpha \cdot (\beta \cdot \gamma)\).
- **Distributive property**: \(\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma\) and \((\beta + \gamma) \cdot \alpha = \beta \cdot \alpha + \gamma \cdot \alpha\).

A ring is called commutative if, in addition, it satisfies the commutative property, i.e., \(\alpha \cdot \beta = \beta \cdot \alpha\) \(\forall \alpha, \beta \in S\).

Remark 2.1. A ring \((S; +, \cdot)\) may or may not satisfy the following properties:

- **Existence of multiplicative Identity**: There exists \(1 \in S\) such that \(1 \cdot \alpha = \alpha \cdot 1 = \alpha\) \(\forall \alpha \in S\).
- **Cancellation**: If \(\alpha \neq 0\), then \(\forall \beta, \gamma \in S\)

\[
\begin{align*}
(\alpha \cdot \beta = \alpha \cdot \gamma) & \Rightarrow (\beta = \gamma) \\
(\beta \cdot \alpha = \gamma \cdot \alpha) & \Rightarrow (\beta = \gamma)
\end{align*}
\]

Definition 2.2. (Ideals) Let \((S; +, \cdot)\) be a ring and let \(D\) be a nonempty subset of \(S\) that is closed under the operations of \(+\) and \(\cdot\) in \(S\). If \((D; +, \cdot)\) is itself a ring, then it is called a subring of \((S; +, \cdot)\). A subring \((D; +, \cdot)\) of a ring \((S; +, \cdot)\) is called a left ideal if

\[
\left( s \in S \text{ and } x \in D \right) \Rightarrow \left( s \cdot x \in D \right)
\]

and is called a right ideal if

\[
\left( s \in S \text{ and } x \in D \right) \Rightarrow \left( x \cdot s \in D \right)
\]

The subring \((D; +, \cdot)\) is called an ideal if it is both a left ideal and a right ideal.

Remark 2.2. Just as normal groups play an important role in the theory of groups (see Definition 1.10), so do ideals in the theory of rings.
Definition 2.3. (Zero-divisor) Let \((S; +, \bullet)\) be a ring. Then, an element \(\alpha \in S \setminus \{0\}\) is called a left [resp. right] zero-divisor if there exists \(\beta \in S \setminus \{0\}\) such that \(\alpha \bullet \beta = 0\) [resp. \(\beta \bullet \alpha = 0\)]. An element of \(S\), which is both a left and a right zero-divisor, is called a zero-divisor.

Definition 2.4. (Invertible Element) Let \((S; +, \bullet)\) be a ring with \(1\) as the multiplicative identity (i.e., the identity with respect to the operation \(\bullet\)). Then, an element \(\alpha \in S\) is called a left [resp. right] invertible if there exists \(\beta \in S\) [resp. \(\gamma \in S\)] such that \(\beta \bullet \alpha = 1\) [resp. \(\alpha \bullet \gamma = 1\)]. The element \(\beta\) [resp. \(\gamma\)] is called a left [resp. right] inverse of \(\alpha\). If an element \(\alpha \in S\) is both left and right invertible, then \(\alpha\) is called an invertible element.

Definition 2.5. (Integral Domain) Let \((S; +, \bullet)\) be a commutative ring with the multiplicative identity \(1 \neq 0\) and no zero divisors. Then, \((S; +, \bullet)\) is called an integral domain.

Definition 2.6. (Division Ring) Let \((S; +, \bullet)\) be a ring with the multiplicative identity \(1 \neq 0\) and let \((S \setminus \{0\}, \bullet)\) be a group with \(1\) as the identity. Then, \((S; +, \bullet)\) is called a division ring.

Definition 2.7. (Field) A commutative division ring is called a field.

Remark 2.3. A field \((F; +, \bullet)\) has the following properties.

- \((F; +, \bullet)\) is a commutative ring.
- There exists \(1 \in F\) such that \(1 \cdot \alpha = \alpha \cdot 1 = \alpha\) \(\forall \alpha \in F\).
- \((F \setminus \{0\}; \bullet)\) is an Abelian group with \(1 \in F \setminus \{0\}\) as the identity element.
- If \(\alpha \in (F \setminus \{0\})\), then there exists \(\alpha^{-1} \in (F \setminus \{0\})\) such that \(\alpha^{-1} \cdot \alpha = \alpha \cdot \alpha^{-1} = 1\).

Example 2.1. \((\mathbb{R} : +; \bullet)\) and \((\mathbb{C} : +; \bullet)\) are fields that are very commonly used in engineering analysis. Note that \((\mathbb{Q} : +; \bullet)\) is also a field but it is seldom used because, as we have seen, \(\mathbb{Q}\) is not a complete set whereas \(\mathbb{R}\) and \(\mathbb{C}\) are. Note that \((\mathbb{Z} : +; \bullet)\) is a commutative ring but it is not a field because no element of \(\mathbb{Z} \setminus \{1\}\) has a multiplicative inverse.

Example 2.2. Fields can be both infinite and finite. Examples of infinite fields are: \((\mathbb{R} : +; \bullet)\) and \((\mathbb{C} : +; \bullet)\). An example of a finite field (also called Galois field) is \(GF(2) \triangleq (\{0,1\}; +_2; \cdot_2)\), where \(+_2\) is the addition operation under modulo 2 and \(\cdot_2\) is the multiplication operation under modulo 2. This is the smallest field.

Example 2.3. A polynomial over a field \((F; +, \bullet)\) is defined as an expression of the form: \(a_0 + a_1 x + a_1 x^2 + \cdots + a_n x^n\) where \(a_i \in F\) and \(n\) is a non-negative integer; and \(x\) is called the indeterminate. Examples of the indeterminate include real numbers, complex numbers, and square matrices.

Definition 2.8. (Homomorphism) Let \((U; \oplus, \circ)\) and \((V; \star, \bullet)\) be two algebraic systems and let \(h : U \to V\). Then, \(h\) is called a homomorphism from \(U\) to \(V\) (carrying \(\oplus\) to \(\star\) and \(\circ\) to \(\bullet\)) if, for all \((u_1, u_2) \in U \times U\), the following conditions hold: \(h(u_1 \oplus u_2) = h(u_1) \star h(u_2)\) and \(h(u_1 \circ u_2) = h(u_1) \bullet h(u_2)\).
If the function \( h \) is injective, then \( h \) is called a monomorphism; if \( h \) is surjective, then it is called a epimorphism; and if \( h \) is bijective, then it is called an isomorphism. In case of isomorphism, the inverse function \( h^{-1} \) exists and it is legitimate to say that \((U; \circlearrowright, \circlearrowleft)\) and \((V; \star, \bullet)\) are isomorphic to each other.

3 Modules and Vector Spaces

This section focuses on modules and vector spaces. Modules over a ring are a generalization of Abelian groups that are necessary for further study of algebra. Important examples of modules are vector spaces that are of prime focus in this course. In particular, a vector space is an Abelian group defined over a field whose elements are called scalars.

Definition 3.1. (Module) Let \((S; +, \bullet)\) be a ring. Then, a (left) \(S\)-module is an additive Abelian (i.e., commutative) group \((M, \oplus)\) together with a function \(S \times M \to M\), where the image of \((s, m) \in S \times M\) is denoted as \(sm \in M\) such that \(\forall r, s \in S\) and \(\forall u, v \in M\) the following conditions hold:

- \(r(u \oplus v) = ru \oplus rv\)
- \((r + s)u = ru \oplus su\)
- \(r(su) = (rs)u\)

If the ring \((S; +, \bullet)\) has a (multiplicative) identity element \(1\), and if \(1u = u\ \forall u \in M\), then \(M\) is called a unitary (left) \(S\)-module. If \((S; +, \bullet)\) is a division ring, then a unitary (left) \(S\)-module is called a left vector space. Similar definitions hold for right \(S\)-module, unitary right \(S\)-module, and right vector space.

Remark 3.1. Modules over a ring are generalization of Abelian groups that, in turn, are modules over \((\mathbb{Z}; +, \bullet)\).

Definition 3.2. (Vector Space) Let \((V; \oplus)\) be an Abelian group and \((F; +, \bullet)\) be a field whose multiplicative identity is 1 (i.e., \(\alpha \bullet 1 = 1 \bullet \alpha = \alpha \ \forall \alpha \in V\)) and the additive identity (i.e., zero-element) is 0 (i.e., \(\alpha \oplus 0 = 0 \oplus \alpha = \alpha \ \forall \alpha \in V\)). Then, \((V; \oplus)\) is called a vector space over the field \((F; +, \bullet)\) if the following properties are satisfied:

- Closure: There exists a function \(\otimes : F \times V \to V\), called scalar multiplication, such that
  \[
  \begin{cases}
  \alpha \otimes x \in V \ \forall \alpha \in F \ \forall x \in V \\
  1 \otimes x = x \ \forall x \in V
  \end{cases}
  \]
- Associativity: \(\alpha \otimes (\beta \otimes x) = (\alpha \bullet \beta) \otimes x \ \forall \alpha, \beta \in F \ \forall x \in V\)
- Distributivity: \(\forall \alpha, \beta \in F \ \forall x, y \in V\)
  \[
  \begin{cases}
  \alpha \otimes (x \oplus y) = (\alpha \otimes x) \oplus (\alpha \otimes y) \\
  (\alpha + \beta) \otimes x = (\alpha \otimes x) \oplus (\beta \otimes y)
  \end{cases}
  \]

The elements of \(V\) are called vectors and the elements of \(F\) are called scalars. Often the multiplicative operators \(\bullet\) and \(\otimes\) are simply omitted, i.e., we write \(\alpha \bullet \beta\) as \(\alpha \beta\) and \(\alpha \otimes x\) as \(\alpha x\). However, it is important to distinguish between the operators of
the scalar addition $+$ and the vector addition $\oplus$. We denote the zero (i.e., additive identity) of the field $(F; +, \bullet)$ as $0$ and the zero vector (i.e., the identity of the Abelian group $(V; \oplus)$) as $0$.

**Remark 3.2.** It follows from the definition of a vector space that $\forall x \in V \ \forall \alpha \in F$

\begin{itemize}
  \item $0x = 0$ because $0x = (1 + (-1))x = 1x \oplus (-1)x = x \oplus (-x) = 0$
  \item $(-1)x = -x$ because $x \oplus (-1)x = 1x \oplus (-1)x = (1 + (-1))x = 0x = 0$
  \item $\alpha 0 = 0$ because $\alpha 0 = \alpha (x \oplus (-x)) = \alpha x \oplus \alpha (-x) = \alpha x \oplus \alpha (-1)x = \alpha x \oplus (-\alpha)x = (\alpha + (-\alpha))x = 0x = 0$
\end{itemize}

**Remark 3.3.** Every vector space must contain the zero vector $0$. That is, the set $V$ in a vector space $(V; \oplus)$ can never be empty. A geometric interpretation of this fact is that every coordinate frame of a vector space must have the origin.

**Definition 3.3.** (Subspace) Let $(V; \oplus)$ be a vector space over a field $(F; +, \bullet)$. Let $U \subseteq V$ such that $(U; \oplus)$ is a vector space over a field $(F; +, \bullet)$. Then, $(U; \oplus)$ is a subspace of $(V; \oplus)$. In addition, if $U$ is a proper subset of $V$, then $(U; \oplus)$ is a proper subspace of $(V; \oplus)$.

**Remark 3.4.** If $(U; \oplus)$ is a subspace of $(V; \oplus)$, then it follows that $(U; \oplus)$ is a subgroup of the group $(V; \oplus)$.

**Proposition 3.1.** $(U; \oplus)$ is a subspace of a vector space $(V; \oplus)$ over a field $(F; +, \bullet)$ if and only if the following condition holds: $(\alpha x \oplus y) \in U \ \forall x, y \in U \ \forall \alpha \in F$.

**Proof.** The proof follows from the definition of a subspace. \qed

**Definition 3.4.** (Linear Combination) Let $(V; \oplus)$ be a vector space over a field $(F; +, \bullet)$ and let $S$ be a nonempty (finite or countably infinite or uncountable) set of vectors. Then, $x \in V$ is said to be a linear combination of vectors in $S$ if there exists a finite set of vectors $\{u^1, \ldots, u^n\}$ and a finite set of scalars $\{\alpha_1, \ldots, \alpha_n\}$, where $n \in \mathbb{N}$, such that $x = \bigoplus_{j=1}^{n} \alpha_j u^j$.

**Definition 3.5.** (Linear Dependence) Let $(V; \oplus)$ be a vector space over a field $(F; +, \bullet)$ and let $S$ be a nonempty (finite or countably infinite or uncountable) set of vectors. Then, the vectors in $S$ are said to linearly independent if, for each $x \in S$, $x$ is not a linear combination of the vectors in $S \setminus \{x\}$ (i.e., the set $S$ with $x$ removed). The set $S$ in the vector space $(V; \oplus)$ is linearly dependent if it is not linearly independent, i.e., there exists a vector $x \in S$ such that $x$ is a linear combination of vectors in $S \setminus \{x\}$.

**Theorem 3.1.** Let $(V; \oplus)$ be a vector space over a field $(F; +, \bullet)$ and let $S$ be a nonempty (finite or countably infinite or uncountable) set of vectors. Then, $S$ is linearly independent if and only if, for each nonempty finite subset of $S$, say, $\{u^1, \ldots, u^n\}$, the only $n$-tuple of scalars satisfying the equation: $\bigoplus_{j=1}^{n} \alpha_j u^j = 0$, is the trivial solution: $\alpha_1 = \cdots = \alpha_n = 0$.

**Proof.** See Naylor & Sell pp. 177-178. \qed
Remark 3.5. A (nonempty) set $S$ in a vector space $(V; \oplus)$ is linearly dependent if and only if there exists a nonempty finite subset of $S$, say, $\{u^1, \cdots, u^n\}$ and scalars $\alpha_1, \cdots, \alpha_n$, where not all $\alpha_i$'s are zero, such that $\bigoplus_{j=1}^n \alpha_j u^j = 0$.

Definition 3.6. (Spanning) Let $(V; \oplus)$ be a vector space over a field $(F; +, \cdot)$ and let $S$ be a nonempty (finite or countably infinite or uncountable) set of vectors. Then, the set of all (finite) linear combinations of vectors in $S$ is the space spanned by $S$ and is denoted as $\text{span}(S)$.

Remark 3.6. It follows from Definition 3.6 that $\text{span}(S)$ is the smallest subspace of $(V; \oplus)$ containing the set $S$, i.e., $\text{span}(S)$ is the intersection of all subspaces of $(V; \oplus)$ that contain the set $S$.

3.1 Hamel Basis and Dimension of Vector Spaces

Hamel basis is a purely algebraic concept and is not the only concept of basis that arises in the analysis of vector spaces. There are other concepts of basis (e.g., Schauder basis and orthonormal basis) that involves both algebraic and topological concepts. These issues will be dealt with in later chapters.

Definition 3.7. (Hamel Basis) Let $(V; \oplus)$ be a vector space over a field $(F; +, \cdot)$ and let $B$ be a nonempty set of vectors. Then, $B$ is said to be a Hamel basis for $(V; \oplus)$ if the following conditions hold: (i) $B$ is a linearly independent set, and (ii) $\text{span}(B) = V$.

Remark 3.7. Let $(V; \oplus)$ be a vector space and let $S$ be a linearly independent set of vectors in $V$. Then, there exists a Hamel basis $B$ such that $S \subseteq B$, and it follows that every vector space has a Hamel basis.

Remark 3.8. Let $B_1$ and $B_2$ be two Hamel bases of a vector space $(V; \oplus)$. Then, $B_1$ and $B_2$ have the same cardinal number. (See Naylor & Sell, p. 184).

Definition 3.8. (Dimension) Let $(V; \oplus)$ be a vector space over a field $(F; +, \cdot)$. The dimension of $(V; \oplus)$ is defined to be the cardinal number of its Hamel basis and is denoted by $\dim(V)$.

Remark 3.9. It follows from the above definition that dimension is a property of the vector space and is independent of any particular Hamel basis.

Definition 3.9. (Vector Space Isomorphism) Let $(V; \oplus)$ and $(W; \otimes)$ be two vector space over the same field $(F; +, \cdot)$. Then, $(V; \oplus)$ and $(W; \otimes)$ are said to be isomorphic if there exists a linear bijective mapping $h : V \rightarrow W$. That is, $h : V \rightarrow W$ satisfies the following conditions.

- **Linearity:** $h(\alpha x \oplus y) = \alpha h(x) \oplus h(y) \quad \forall x, y \in V \quad \forall \alpha \in F$
- **Injectivity:** $(h(x) = h(y)) \Rightarrow (x = y) \quad \forall x, y \in V$
- **Surjectivity:** $\{h(x) : x \in V\} = W$

Remark 3.10. The role of the field in Definition 3.9 is very critical in the sense that a given Abelian group could generate two different vector spaces when defined over two different fields. For example, the Abelian group of complex numbers $\mathbb{C}$
forms a two-dimensional vector space over the real field \( \mathbb{R} \), which is isomorphic to the two-dimensional real vector space \( \mathbb{R}^2 \). The same Abelian group of complex numbers \( \mathbb{C} \) forms a one-dimensional vector space over the complex field \( \mathbb{C} \). The implications of this fact are illustrated below by a physical example.

Let an unforced single-degree-of-freedom underdamped linear time-invariant mass-spring-damper (or inductance-capacitance-resistance) system, with non-zero initial conditions, be governed by the following equation.

\[
\frac{d^2 y(t)}{dt^2} + 2 \xi \omega_n \frac{dy(t)}{dt} + \omega_n^2 y(t) = 0
\]

where \( y(t) \) is the time-dependent displacement (or charge); \( \omega_n \) is the natural frequency; and \( \xi \) is the damping coefficient \((0 \leq \xi < 1)\). A state-space representation of the above equation in the two-dimensional vector space \( \mathbb{R}^2 \) over the real field \( \mathbb{R} \) is given below.

\[
\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix}
-\xi \omega_n & \sqrt{1-\xi^2} \omega_n \\
-\sqrt{1-\xi^2} \omega_n & -\xi \omega_n
\end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}
\]

where a choice for the states is: \( x_1 = \xi \omega_n y + \frac{dy}{dt} \) and \( x_2 = -\sqrt{1-\xi^2} \omega_n y \) and the state transition matrix \( \begin{bmatrix}
-\xi \omega_n & \sqrt{1-\xi^2} \omega_n \\
-\sqrt{1-\xi^2} \omega_n & -\xi \omega_n
\end{bmatrix} \in \mathbb{R}^{2 \times 2} \) maps \( \mathbb{R}^2 \) into \( \mathbb{R}^2 \).

Let us construct a vector space over the complex field \( \mathbb{C} \) instead of the real field \( \mathbb{R} \). To this end, let us define a complex-valued state \( z(t) \triangleq x_1(t) + ix_2(t) \), where \( i = \sqrt{-1} \); and \( x_1 \) and \( x_2 \) are as defined above. Then it follows that

\[
\frac{dx_1}{dt} + i\frac{dx_2}{dt} = \left(-\xi \omega_n x_1 + \sqrt{1-\xi^2} \omega_n x_2\right) + i \left(-\sqrt{1-\xi^2} \omega_n x_1 - \xi \omega_n x_2\right),
\]

which reduces to

\[
\frac{d}{dt}[x_1 + ix_2] = \left(-\xi + i\sqrt{1-\xi^2}\omega_n\right)[x_1 + ix_2].
\]

Consequently, the state-space representation of the above equation in the one-dimensional vector space \( \mathbb{C}^1 \) over the complex field \( \mathbb{C} \) is as given below.

\[
\frac{d}{dt}[z] = \left(-\xi + i\sqrt{1-\xi^2}\omega_n\right)[z]
\]

where the state transition matrix \( \left(-\xi + i\sqrt{1-\xi^2}\omega_n\right) \in \mathbb{C}^{1 \times 1} \) maps \( \mathbb{C}^1 \) into \( \mathbb{C}^1 \). Note that, in the disciplines of Physics and Engineering, such a 2nd order underdamped system is often referred to as a single degree-of-freedom system. Here, we have shown that an underdamped system is represented on a vector space of dimension 2 over the real field \( \mathbb{R} \), or on a vector space of dimension 1 over the complex field \( \mathbb{C} \).

Next we consider an overdamped system, i.e., the damping coefficient \( \xi > 1 \). A state-space representation of the overdamped system in the two-dimensional vector space \( \mathbb{R}^2 \) over the real field \( \mathbb{R} \) is given below.

\[
\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix}
-\xi \omega_n & \sqrt{\xi^2 - 1} \omega_n \\
\sqrt{\xi^2 - 1} \omega_n & -\xi \omega_n
\end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}
\]
where a choice for the states is: $x_1 = \xi \omega_n y + \frac{dy}{dt}$ and $x_2 = \sqrt{\xi^2 - 1} \omega_n y$ and the state transition matrix

$$\begin{bmatrix}
-\xi \omega_n & \sqrt{\xi^2 - 1} \omega_n \\
\sqrt{\xi^2 - 1} \omega_n & -\xi \omega_n
\end{bmatrix} \in \mathbb{R}^{2 \times 2} \text{ maps } \mathbb{R}^2 \text{ into } \mathbb{R}^2.
$$

Similar to what was done for the underdamped system, the vector space is constructed over the complex field $\mathbb{C}$ instead of the real field $\mathbb{R}$. The complex-valued state is defined as $z(t) \triangleq x_1(t) + ix_2(t)$, where $i = \sqrt{-1}$; and $x_1$ and $x_2$ are as defined above. Then it follows that

$$\frac{dx_1}{dt} + i \frac{dx_2}{dt} = \left( -\xi \omega_n x_1 + \sqrt{\xi^2 - 1} \omega_n x_2 \right) + i \left( \sqrt{\xi^2 - 1} \omega_n x_1 - \xi \omega_n x_2 \right),$$

which reduces to

$$\frac{d}{dt} [x_1 + ix_2] = \omega_n (-\xi x_1 + \sqrt{\xi^2 - 1} x_2) + i(\sqrt{\xi^2 - 1} x_1 - \xi x_2))\]$$

Consequently, the state-space representation of the above equation in the one-dimensional vector space $\mathbb{C}^1$ over the complex field $\mathbb{C}$ as given below.

$$\frac{d}{dt}[z] = \left[ - (\xi + i \sqrt{\xi^2 - 1}) \omega_n \right] [z] - \left[ \sqrt{\xi^2 - 1} \omega_n \right] (z - \bar{z})$$

where $\bar{z}$ is the complex conjugation of $z \in \mathbb{C}^1$. Note that the operation of complex conjugation may not be linear because, given any $z, \bar{z}, \gamma \in \mathbb{C}$, it follows that $\bar{z + \gamma} = \bar{z} + \gamma \bar{\bar{z}}$ instead of $\bar{z + \gamma} \bar{\bar{z}} = \bar{\bar{z}} + \bar{\gamma} \bar{\bar{z}}$.

For the critically damped system (i.e., $\xi = 1$), The system dynamics is given as

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -\omega_n & \omega_n \\ 0 & -\omega_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

where a choice for the states is: $x_1 = \omega_n y$ and $x_2 = \omega_n y + \frac{dy}{dt}$ and the state transition matrix

$$\begin{bmatrix}
-\omega_n & \omega_n \\
0 & -\omega_n
\end{bmatrix} \in \mathbb{R}^{2 \times 2} \text{ maps } \mathbb{R}^2 \text{ into } \mathbb{R}^2.$$

Letting $z(t) \triangleq x_1(t) + ix_2(t)$, it follows that

$$\frac{d}{dt}[z] = [-\omega_n][z] - \left[ \frac{i}{2} \right] (z - \bar{z}).$$

**Theorem 3.2.** Let $(V; \oplus)$ and $(W; \otimes)$ be two vector spaces over the same field $\mathbb{F}$. Then, $(V; \oplus)$ and $(W; \otimes)$ are isomorphic if and only if $\dim(V) = \dim(W)$.

**Proof.** Let $(V; \oplus)$ and $(W; \otimes)$ be isomorphic and let $B_V$ be a basis for $V$. Let $h : V \to W$ be an isomorphism (i.e., linear bijection) of $V$ onto $W$. Then, due to linearity of $h$, the image $h(B_V)$ of $B_V$ under $h$ generates linear combinations that would span $W$. Furthermore, since $h$ is surjective, it follows that span($h(B_V)$) = $W$. Therefore, $h(B_V)$ is a Hamel basis of $W$. Finally, since $h$ is injective, it follows that $\left( \text{card}(B_V) = \text{card}(h(B_V)) \right) \Rightarrow \left( \dim(V) = \dim(W) \right)$.

To prove the other way, let $\dim(V) = \dim(W)$ be given. Therefore, there exists a one-to-one correspondence $\hat{h}$ between the spaces $V$ and $W$. Using this correspondence $\hat{h}$, let us define a linear mapping $h : V \to W$.

Let $B_V$ and $B_W$ be Hamel bases of the spaces $V$ and $W$, respectively. Let $x \in V$ that is expressed uniquely as a (finite) linear combination of vectors in $B_V$
in the form: \( x = \bigoplus_{j=1}^{n} \alpha_j v^j \), where \( v^j \in B_V \) and the scalar \( \alpha_j \in F \) and at least one \( \alpha_j \neq 0 \). Now we construct \( h \) such that

\[
\begin{align*}
  h(x) &= \bigotimes_{j=1}^{n} \alpha_j \hat{h}(v^j) \\
  h(0) &= 0
\end{align*}
\]

Since \( B_W \) is a basis, \( h \) is a bijective mapping from \( v \) onto \( V \). Furthermore, \( h \) is linear and hence \( h \) is an isomorphism. \( \square \)

**Theorem 3.3.** (Vector Space Isomorphism) \((V; \otimes)\) be a finite-dimensional vector space over the field \( F \), where \( \dim(V) = n \in \mathbb{N} \). Let \((F^n, \oplus)\) denote the vector space made up of ordered \( n \)-tuples of scalars in \( F \). Then, the vector space \((V; \otimes)\) is isomorphic to \((F^n, \oplus)\).

**Proof.** Given that \((V; \otimes)\) is an \( n \)-dimensional vector space with scalars in \( F \), let \( \{v^1, \ldots, v^n\} \) be a basis of \( V \), i.e., the vectors \( v^j \), \( j = 1, \ldots, n \) are linearly independent and span the space \((V; \otimes)\).

To establish an isomorphism between the spaces \((F^n, \oplus)\) and \((V; \otimes)\), it is necessary to find a function \( h : F^n \rightarrow V \) that is linear and bijective. Let \( h(a) \triangleq \bigotimes_{j=1}^{n} a_j v^j \), where \( a = [a_1 \cdots a_n] \) with \( a_j \in F \).

**Linearity of \( h \):** Let \( \alpha \in F \) and \( a, b \in F^n \). Then, \( \alpha a \oplus b = F^n \) and \( h(\alpha a \oplus b) = \bigotimes_{j=1}^{n} (\alpha a_j + b_j) v^j = (\alpha h(a)) \otimes h(b) \Rightarrow \) linearity is established.

**Injectivity of \( h \):** Let \( h(a) = h(b) \) for \( a, b \in F^n \). By linearity of \( h \) and using \(-1 \in F\), it follows that \( 0 = h(a \oplus (-1)b) \Rightarrow \bigotimes_{j=1}^{n} (a_j + (-1)b_j) v^j \). Because of linear independence of \( v^j \)'s, it follows that \( a_j + (-1)b_j = 0 \Rightarrow a_j = b_j \ \forall j \). Therefore, \( a = b \Rightarrow \) injectivity is established.

**Surjectivity of \( h \):** Since \( v^j \)'s form a basis of \((V; \otimes)\), there exist unique scalars \( c_j \)'s such that \( x = \bigotimes_{j=1}^{n} c_j v^j = h(c) \ \forall x \in V \), i.e., there exists \( c \in F^n \) \( \forall x \in V \Rightarrow \) surjectivity is established. \( \square \)

**Remark 3.11.** Since a finite-dimensional (say \( n \in \mathbb{N} \)) vector space \( V \) over a field \( F \) (e.g., the real field \( \mathbb{R} \) or the complex field \( \mathbb{C} \)) is isomorphic to \( F^n \), it is natural to study the properties of \( F^n \) in lieu of those of \( V \). The vectors in \( F^n \) can be interpreted as: (i) \( n \)-tuples of scalars in \( F \), or (ii) linear mappings from \( F \) to \( F^n \). Let us present a generalization of the second concept. Let \( N \triangleq \{1, 2, \ldots, n\} \) be the set of all integers from 1 to \( n \). Then, a function \( x : N \rightarrow F \) implies that a rule assigns a scalar \( x(j) \) to each \( j \in N \) and let the set of all such functions be denoted as \( F(N) \), i.e., specifying a function \( x \in F(N) \) amounts to presenting \( n \) consecutive scalars \( x(1), x(2), \ldots, x(n) \), which could be arranged as a column vector or as a row vector in \( F^n \). This provides a bijectivity between \( F(N) \) and \( F^n \). For all \( c \in F \) and \( x, y \in F(N) \), it is meaningful to define the operations of scalar multiplication and vector addition in \( F(N) \) respectively by

\[
(x \cdot c)(j) = cx(j) \quad \text{and} \quad (x \oplus y)(j) = x(j) + y(j)
\]

These operations make \( F(N) \) a vector space over the field \( F \), which reveals that there is no essential difference between \( F(N) \) and \( F^n \). Now we generalize this concept to infinite-dimensional vector spaces.

Let \( S \) be an arbitrary (i.e., finite, countable, or uncountable) nonempty set that serves as an index set and let \( F^S \) denote the set of all functions \( f : S \rightarrow F \). If \( S = \mathbb{N} \),
then $\mathbb{F}^S$ is the vector space of all possible sequences $s : \mathbb{N} \to \mathbb{F}$ which includes all (i.e., real or complex) scalar-valued sequences. Similarly, if $S = \mathbb{R}$, then $\mathbb{F}^S$ is the vector space of all possible (not necessarily linear) functions $f : \mathbb{R} \to \mathbb{F}$ that includes all (i.e., real or complex) scalar-valued time-dependent signals; for example, $f(t)$ could be the output voltage of an operational amplifier. This concept allows extension of the tools of linear algebra to functions spaces. In general, if $S$ is a countably infinite set, i.e., $S \sim \mathbb{N}$, the the vector $x \in \mathbb{F}^S$ is called a sequence; if $S$ is an uncountably infinite set, i.e., $S \sim \mathbb{R}$, the vector $x \in \mathbb{F}^S$ is called a function.

### 3.2 Linear Transformation of Vector Spaces

This subsection introduces the rudimentary concepts of linear transformation of vector spaces. More advanced materials on transformation of vector spaces will be presented in later chapters.

**Definition 3.10.** (Linear Transformation) Let $(V; \oplus)$ and $(W; \otimes)$ be two vector spaces over the same field $\mathbb{F}$. Then, $L : V \to W$ is called a linear transformation if the following condition holds for all $x^k \in V$ and all $\alpha_k \in \mathbb{F}$:

$$L \left( \bigoplus_{k=1}^{n} (\alpha_k x^k) \right) = \bigotimes_{k=1}^{n} \left( \alpha_k L(x^k) \right) \text{ for any arbitrary } n \in \mathbb{N}$$

If $W = V$, then the transformation $L : V \to V$ is called an operator. If $W = \mathbb{F}$, then the transformation $L : V \to \mathbb{F}$ is called a functional. The space of all linear bounded functionals of a vector space $V$ is called the dual space $V^\star$.

**Definition 3.11.** (Null Space) The null space $\mathcal{N}(L)$ of a linear transformation $L : V \to W$ is the subspace of $V$ that satisfies the following condition:

$$\mathcal{N}(L) \triangleq \{ x \in V : Lx = 0_W \}$$

**Definition 3.12.** (Range Space) The range space $\mathcal{R}(L)$ of a linear transformation $L : V \to W$ is the subspace of $W$ that satisfies the following condition:

$$\mathcal{R}(L) \triangleq \{ Y \in W : y = Lx \text{ and } x \in V \}$$

**Definition 3.13.** (Sum, Direct Sum, and Algebraic Complement) Let $V$ be a vector space, where the operation of vector addition is $\oplus$, and let $U$ and $W$ be two subspaces of $V$. Then, the sum of $U$ and $W$ is a subspace of $V$, denoted as $Y = U + W$, which is spanned by all vectors in $U$ and $W$, i.e.,

$$U + W \triangleq \{ x = u \oplus w : u \in U \text{ and } w \in W \}$$

The direct sum of of $U$ and $W$ is a subspace of $V$, denoted as $Y = U \oplus W$, if $\forall y \in Y \exists u \in U$ and $w \in W$ such that there is a unique representation $y = u \oplus w$.

In this setting, $U$ is called the algebraic complement of $W$ (alternatively, $W$ is the algebraic complement of $U$) in $Y$.

**Definition 3.14.** (Projection) Let $V$ be a vector space. Then, a linear transformation $P : V \to V$ is called a projection on $V$ if $P^2 = P$, i.e., $P(Px) = Px \forall x \in V$. 

11
Therefore, \( z \) which implies \( x \) Therefore, \( R \)

**Proof.** First part: Let \( x \in R(P) \cap N(P) \). Since \( x \in R(P) \), there exists \( y \in V \) such that \( Py = x \Rightarrow Px = x \). Since \( x \in N(P) \), it follows that \( Px = 0_V \Rightarrow x = 0_V \). Therefore, \( R(P) \cap N(P) = \{0_V\} \).

Second part: Let \( x \in V \) and define \( y = Px \) and \( y \in R(P) \). Having \( z = x - y \) which implies \( x = y + z \), it follows that

\[
Pz = P(x - y) = Px - Py = Px - P^2x = Px - Px = 0_V.
\]

Therefore, \( z \in N(P) \). Since the choice of \( x \in V \) is arbitrary, it follows that \( R(P) \oplus N(P) = V \).

**Remark 3.12.** Linearity is a necessity for projection on a vector space. A nonlinear mapping \( f : V \to V \), which satisfies the condition \( f^2 = f \) is not necessarily a projection. For example, let \( f : \mathbb{R} \to \mathbb{R} \) be defined as:

\[
f(x) = \begin{cases} 
1 & \text{if } x \geq 1 \\
0 & \text{if } -1 < x < 1 \\
-1 & \text{if } x \leq -1 
\end{cases}
\]

In this case, \( f \circ f = f \) but \( f \) is not a projection because it is nonlinear.

**Lemma 3.1.** (Decomposition of projection) Let \( P \) be a projection defined on a vector space \( V \). Then, the null space \( N(P) \) and the range space \( R(P) \) of the projection \( P \) are algebraic complements of each other in \( V \), and the following conditions are satisfied.

1. \( R(P) \cap N(P) = \{0_V\} \).
2. \( R(P) \oplus N(P) = V \).

**Proof.** First part: Let \( x \in R(P) \cap N(P) \). Since \( x \in R(P) \), there exists \( y \in V \) such that \( Py = x \Rightarrow Px = x \). Since \( x \in N(P) \), it follows that \( Px = 0_V \Rightarrow x = 0_V \). Therefore, \( R(P) \cap N(P) = \{0_V\} \).

Second part: Let \( x \in V \) and define \( y = Px \) and \( y \in R(P) \). Having \( z = x - y \) which implies \( x = y + z \), it follows that

\[
Pz = P(x - y) = Px - Py = Px - P^2x = Px - Px = 0_V.
\]

Therefore, \( z \in N(P) \). Since the choice of \( x \in V \) is arbitrary, it follows that \( R(P) \oplus N(P) = V \).

**Remark 3.13.** Let \( U \) and \( W \) be two subspaces of a vector space \( V \) such that \( U \cap W = \{0_V\} \) and \( U \oplus W = V \). Then, there exists a projection \( P \) on \( V \) such that \( R(P) = U \) and \( N(P) = W \). This issue will be further dealt with in Chapter Four (see also Naylor & Sell, p. 200).

**Remark 3.14.** Some of the important observations on sum, direct sum, and algebraic complement of subspaces are listed below: (see Naylor & Sell, pp.198-199.)

1. The sum \( U + W \) is the direct sum \( U \oplus W \) if and only if \( U \cap W = 0_V \).
2. The natural mapping \( \varphi \), denoted as \( \varphi[(x, y)] = x \oplus y \), of the direct sum \( U \oplus W \) is an isomorphism if and only if \( U \cap W = 0 \).
3. If \( U \) is a subspace of a vector space \( V \), then \( U \) must have an algebraic complement in \( V \).

### 3.3 Matrix Representation of Finite-dimensional Spaces

Let \( V \) and \( W \) be two finite-dimensional vector spaces over the same field \( \mathbb{F} \), where \( \dim(V) = n \) and \( \dim(W) = m \); let \( \mathcal{A} \in L(V,W) \) be a linear transformation from \( V \) to \( W \). If \( B \triangleq \{b_1, b_2, \ldots, b_n\} \) is a basis for \( V \), then each vector \( x \in V \) can be uniquely represented as \( x = \sum_{k=1}^{n} b_k \beta_k \), where \( \beta_k \in \mathbb{F} \). Let the vector of coordinates \( \beta_k \) be denoted as \( [x]^B \triangleq [\beta_1, \beta_2, \ldots, \beta_n]^T \), which implies that \( [x]^B \in \mathbb{F}^n \). Similarly, if \( C \triangleq \{c_1, c_2, \ldots, c_m\} \) is a basis for \( W \), then the image of \( x \in V \) under \( \mathcal{A} \in L(V,W) \),
Next, to show that the spaces $L$ and $F^{m \times n}$ are equivalent, it follows that

Let $\exists$ a unique matrix $T$ to be shown are that $P_k$ similar to that of Theorem 3.3.

Example 3.1. (Integration of a Polynomial) Let the vector spaces $V$ and $W$ over the real field $\mathbb{R}$ be defined as: $V = P^{n-1}((0, \infty))$ and $W = P^n((0, \infty))$, where $P^n((0, \infty))$ is the space of polynomials of degree $n$ or less and the indeterminate $x \in V$.

$A \xrightarrow{\mathcal{A}} Ax \in W$

$[x]^B \in F^n$

$\mathcal{A}^{B,C}$

$[Ax]^C \in F^m$

Figure 1: Isomorphism between $L(V, W)$ and $F^{m \times n}$

i.e., the vector $Ax \in W$ can be uniquely represented as $Ax = \sum_{k=1}^{m} c^k \gamma_k$, where $\gamma_k \in F$; thus, the vector of coordinates $\gamma_k$ is denoted as $[Ax]^C \equiv [\gamma_1 \gamma_2 \cdots \gamma_m]^T$, which implies that $[Ax]^C \in F^m$.

**Definition 3.15.** (Matrix representation) The linear transformations from the coordinates, $[x]^B \in F^n$, of a vector $x \in V$ to the coordinates, $[Ax]^C \in F^m$, of the vector $Ax \in W$ is an $(m \times n)$ matrix. This matrix, denoted by $A \in F^{m \times n}$, serves as the coordinates of the linear transformation $\mathcal{A} \in L(V, W)$ relative to the bases $(B, C)$.

**Theorem 3.4.** (Matrix Representation Theorem) Let $B \triangleq \{b^1, b^2, \cdots, b^n\}$ and $C \triangleq \{c^1, c^2, \cdots, c^m\}$ be bases for the vector spaces $V$ and $W$, respectively. Let the $k^{th}$ column of a matrix $A \in F^{m \times n}$ be defined as $a^k \triangleq [Ab^k]^C$ for $1 \leq k \leq n$ and $k \in \mathbb{N}$. Then, $L(V, W)$ is isomorphic to $F^{m \times n}$ and $[Ax]^C = A[x]^B \ \forall x \in V$.

**Proof.** Let $x \in V$ and $x = \sum_{k=1}^{n} b^k \beta_k$. Then, $[x]^B \equiv [\beta_1 \beta_2 \cdots \beta_n]^T$. Since $\mathcal{A}$ is a linear transformation, it follows that

$$[Ax]^C = \left[ A \sum_{k=1}^{n} b^k \beta_k \right]^C = \sum_{k=1}^{n} \left[ Ab^k \right]^C \beta_k = \sum_{k=1}^{n} a^k \beta_k = A[x]^B$$

Next, to show that the spaces $L(V, W)$ and $F^{m \times n}$ are isomorphic, let us define a linear transformation $T : L(V, W) \to F^{m \times n}$ such that, for each $A \in L(V, W)$, there exists a unique matrix $A \in F^{m \times n}$ whose $k^{th}$ column is $a^k = [Ab^k]^C$. What remain to be shown are that $T$ is linear, injective, surjective. This part of the proof is similar to that of Theorem 3.3.

In view of Theorem 3.4, every finite-dimensional linear transformation can be represented by a matrix of appropriate dimension as seen in Figure 1 that pictorially illustrates the isomorphic relationship between the spaces $L(V, W)$ and $F^{m \times n}$. The matrix $A \in F^{m \times n}$ is the matrix representation of the linear transformation $\mathcal{A}$ induced by the bases $B$ and $C$.

**Example 3.1.** (Integration of a Polynomial) Let the vector spaces $V$ and $W$ over the real field $\mathbb{R}$ be defined as: $V = P^{n-1}((0, \infty))$ and $W = P^n((0, \infty))$, where $P^n((0, \infty))$ is the space of polynomials of degree $n$ or less and the indeterminate
is a real scalar in the range of $(0, \infty)$. Let us consider the integral transform $A \in L(V,W)$, i.e., for each $x \in V$ and each $t \in (0, \infty)$,

$$(Ax)(t) \triangleq \int_0^t d\tau \; x(\tau)$$

Let $B$ be the basis of elementary polynomials, i.e., $B = \{1, t, t^2, \cdots\}$ for $t \in (0, \infty)$. Then, the matrix representation of $A$ induced by the bases of elementary polynomials is computed as:

$$a_k = [At^{k-1}]^B = \left[ \int_0^t d\tau \; \tau^{k-1} \right]^B = \left[ \frac{t^k}{k} \right]^B = \frac{1}{k} \bf{1}^{k+1}$$

where $\bf{1}^k$ denotes the $k^{th}$ column of the identity matrix. Thus, for $n = 3$, it follows that the matrix representation $A$ of the integral transform $A$ mapping the coordinates of a polynomial into the coordinates of its integral is given as:

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$$

For example, if $x(t) = 4 - 3t + t^2$, then $[x]^B = [4 - 3 \ 1]^T$ and the coordinates of $Ax$ are obtained as

$$[Ax]^B = A[x]^B = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 4 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ -\frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

Thus, the integral of $(4 - 3t + t^2)$ is $(4t - \frac{3}{2}t^2 + \frac{1}{3}t^3)$.

**Example 3.2. (Differentiation of a Polynomial)** Let the vector spaces $V$ and $W$ over the real field $\mathbb{R}$ be defined as: $V = P^n((0, \infty))$ and $W = P^{n-1}((0, \infty))$, where $P^n((0, \infty))$ is the space of polynomials of degree $n$ or less and the indeterminate is a real scalar in the range of $(0, \infty)$. Let us consider the derivative transform $A \in L(V,W)$, i.e., for each $x \in V$ and each $t \in (0, \infty)$,

$$(Ax)(t) \triangleq \frac{d}{dt} x(t)$$

Let $B$ be the basis of elementary polynomials, i.e., $B = \{1, t, t^2, \cdots\}$, $t \in (0, \infty)$. Then, the matrix representation of $A$ induced by the bases of elementary polynomials is computed as:

$$a_k = [At^k]^B = \left[ \frac{d}{dt} t^k \right]^B = \left[ kt^{k-1} \right]^B = k \bf{1}^k$$

where $\bf{1}^k$ denotes the $k^{th}$ column of the identity matrix. Thus, for $n = 3$, it follows that the matrix representation $A$ of the derivative transform $A$ mapping the coordinates of the polynomial into the coordinates of its derivative is given as:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$
Let $\mathbf{B}$ be two finite-dimensional vector spaces over the same field $\mathbb{F}$.

### Theorem 3.5

**Transformation Theorem.** Let $\mathbf{A} \in L(V,W)$, where $V \sim \mathbb{F}^n$ and $W \sim \mathbb{F}^m$. Then, $\mathbf{A}^B$ represents the same linear transformation in a different coordinate system. Here each vector in $L^V$ and $L^W$ has $n \times m$ coordinates.

For example, if $x(t) = -4 + 5t - 2t^2 + t^3$, then $[x]^B = [-4 \ 5 \ -2 \ 1]^T$ and the coordinates of $Ax$ are obtained as

$$
[Ax]^B = A[x]^B = \begin{bmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 2 & 0 \\
0 & 0 & 0 & 3 \\
\end{bmatrix} \begin{bmatrix}
-4 \\
5 \\
-2 \\
1 \\
\end{bmatrix} = \begin{bmatrix}
5 \\
-4 \\
3 \\
\end{bmatrix}
$$

Thus, the derivative of $(-4 + 5t - 2t^2 + t^3)$ is $(5 - 4t + 3t^2)$.

### 3.4 Coordinate Transformation in Finite-dimensional Spaces

This subsection introduces the basic concepts of coordinate transformation in finite-dimensional vector spaces through the usage of linear transformation. The concept of matrix representation in Theorem 3.4 leads to the fact that, by choosing bases for two finite-dimensional vector spaces $V \sim \mathbb{F}^n$ and $W \sim \mathbb{F}^m$ over the same field $\mathbb{F}$, each vector in $L(V,W)$ can be represented by its coordinates as a vector in $\mathbb{F}^{m \times n}$.

To develop a strategy for selecting bases of $V$ and $W$, we encounter the problem of transforming from one matrix representation to another. That is, given the matrix representation of $\mathbf{A} \in L(V,W)$ relative to a given pair of bases, the problem is how to find its matrix representation relative to another pair of bases. Here each matrix represents the same linear transformation in a different coordinate system. Consequently, moving from one matrix representation to another is a coordinate transformation problem.

**Theorem 3.5. (Coordinate Transformation Theorem)** Let $V \sim \mathbb{F}^n$ and $W \sim \mathbb{F}^m$ be two finite-dimensional vector spaces over the same field $\mathbb{F}$ and let $\mathbf{A} \in L(V,W)$. Let $B = \{b^1, b^2, \ldots, b^n\}$ and $C = \{c^1, c^2, \ldots, c^n\}$ be two bases for $V$, and let $D = \{d^1, d^2, \ldots, d^m\}$ and $E = \{e^1, e^2, \ldots, e^n\}$ be two bases for $W$. Let the matrix $P \in \mathbb{F}^{n \times m}$ whose $k^{th}$ column is $[c^k]^B$, and let the matrix $Q \in \mathbb{F}^{m \times m}$ whose $k^{th}$ column is $[d^k]^E$. Then, it follows that

$$
[A]^C,E = Q [A]^{B,D} P
$$

**Proof.** First we show that the matrix $P$ maps $C$-coordinates into $B$-coordinates.

For each $x \in V$, we have

$$
[x]^B = \left[\sum_{k=1}^{n} c^k [x]^C \right]^B = \sum_{k=1}^{n} [c^k]^B [x]^C = \sum_{k=1}^{n} p^k [x]^C = P [x]^C
$$
By using a similar argument we show that the matrix $Q$ maps $D$-coordinates into $E$-coordinates. That is, for each $y \in W$, we have $[y]^E = Q[y]^D$. Therefore, for each $x \in V$, it follows that


Since $x \in V$ is arbitrarily chosen, the proof follows by application of Theorem 3.4.

Pre-multiplication and post-multiplication by the respective transformation matrices is needed to obtain one matrix representation from another as seen pictorially in Figure 2 that summarizes the relationship between a linear transformation and two of its matrix representations. A matrix itself is a linear transformation. As such, it is possible to find a matrix representation of a matrix.

**Example 3.3.** (Coordinate Transformation) This example deals with a linear transformation, namely, differentiation on the space of polynomials, and two of its matrix representations. Let $V = P^n((0, \infty)$ and let $A$ be the derivative transformation on $V$. That is, for each $x \in V$ and each $t \in (0, \infty)$,

$$(Ax)(t) \triangleq \frac{d}{dt} x(t)$$

Then, $A \in L(V,W)$, where $W = P^{n-1}((0, \infty))$. Let $B$ be the standard basis of elementary polynomials, i.e., $B = \{1, t, t^2, \cdots \}$. If $A \in \mathcal{F}^{n \times (n+1)}$ is the matrix representation of $A$ induced by the bases of elementary polynomials, then it follows from Example 3.2 for the case $n = 3$ that

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

Next, let us consider the following alternative basis for the space of polynomials: $D = \{1, 2t, 3t^2, \cdots \}$. If $D \in \mathcal{F}^{n \times (n+1)}$ is the matrix representation of $A$ induced by the the bases associated with $D$, then it follows from Theorem 3.5 that $D = QAP$, where the coordinate transformation matrix $P$ is computed as $p^k = [kt^{k-1}]^B = k1^k$ for $1 \leq k \leq (n+1)$, and the coordinate transformation matrix $Q$ is computed as $q^k = [kt^{k-1}]^D = \frac{1}{k}1^k$ for $1 \leq k \leq n$). For $n = 3$, it follows that

$$P = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 2 & 0 & 0 \\
0 & 0 & 3 & 0 \\
0 & 0 & 0 & 4
\end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix}
1 & 0 & 0 \\
0 & \frac{1}{2} & 0 \\
0 & 0 & \frac{1}{3}
\end{bmatrix}$$

Therefore,

$$D = QAP = \begin{bmatrix}
0 & 2 & 0 & 0 \\
0 & 0 & 3 & 0 \\
0 & 0 & 0 & 4
\end{bmatrix}$$

As a check, let us use Theorem 3.4 to compute the matrix $D$ directly. For $1 \leq k \leq (n+1)$, we have

$$d^k = [A(kt^{k-1})]^D = \left[ \frac{d}{dt}(kt^{k-1}) \right]^D = [k(kt^{k-1})]^{k-1} = k \ 1^{k-1} \text{ and } 1^0 = 0 \quad \quad \quad \quad$$
4 Tensor Product of Vector Spaces

The notion of tensor product is rather a language than a substantial mathematical theory. Nevertheless this language is very useful in the disciplines of Physics and Engineering and is often indispensable. In particular tensors provide a uniform description of different aspects of linear algebra within a single algebraic structure.

**Definition 4.1.** (Multilinearity) Let $V_1, \cdots, V_p$, where $p \in \mathbb{N}$, and $U$ be vector spaces over a given field $\mathbb{F}$. Then, the map

$$\varphi : V_1 \times \cdots \times V_p \to U$$

is called multilinear (specifically, $p$-linear) if it is linear in each of the $p$ arguments when the other arguments are held constant. Such a map is denoted as $\text{Hom}(V_1, \cdots, V_p, U)$ and forms a vector space, which is a subspace of the space of all maps from $V_1, \cdots, V_p$ to $U$.

If the spaces $V_1, \cdots, V_p$ and $U$ are finite-dimensional (but not necessarily of the same dimension), then the space $\text{Hom}(V_1, \cdots, V_p, U)$ is also finite-dimensional and $\dim(\text{Hom}(V_1, \cdots, V_p, U)) = \dim(V_1) \cdots \dim(V_p) \dim(U)$. This is so because a multilinear map $\varphi : V_1 \times \cdots \times V_p \to U$ is determined by the images of the basis vectors of the vector spaces $V_1, \cdots, V_p$, which, in turn, are determined by their coordinates in a basis of $U$. Furthermore, if $U = \mathbb{F}$, then $\text{Hom}(V_1, \cdots, V_p, \mathbb{F})$ is a space of multilinear functions on $V_1, \cdots, V_p$. In particular, if $p = 1$, then $\text{Hom}(V, \mathbb{F})$ is called the dual space $V^*$ of the vector space $V$ provided that $\text{Hom}(V, \mathbb{F})$ is bounded.

The tensor products of two vector spaces $V$ and $W$ (over the same field) arises naturally for bilinear maps $\varphi : V \times W \to U$, and the corresponding space $U$ is called the tensor product of $V$ and $W$.

**Theorem 4.1.** (Bilinear Properties) Let $V$ and $W$ be two vector spaces (over the same field) with bases $\{v_i, \ i \in I\}$ and $\{w_j, \ j \in J\}$. The following properties of a bilinear map $\varphi : V \times W \to U$ are equivalent.

- (i) The set of vectors $\{\varphi(v_i, w_j), \ i \in I, \ j \in J\}$ forms a basis of $U$.
- (ii) Every vector $u \in U$ decomposes uniquely as $u = \sum_i \varphi(v_i, y_i), \ y_i \in W$.
- (iii) Every vector $u \in U$ decomposes uniquely as $u = \sum_j \varphi(x_j, w_j), \ x_j \in V$.

If the spaces $V$ and $W$ are infinite-dimensional, the respective sums are required to be finite.

**Proof.** Let $u = \sum_i \sum_j u_{ij} \varphi(v_i, w_j)$. Then, by setting $y_i = \sum_j u_{ij} w_j$, it follows that $u = \sum_i \varphi(v_i, y_i), \ y_i \in W$ and vice versa. Therefore, equivalence of the properties (i) and (ii) is established. Similarly, by setting $x_j = \sum_i u_{ij} v_i$, it follows that $u = \sum_j \varphi(x_j, w_j), \ x_j \in V$. Thus, equivalence of (i) and (iii) is established.

**Remark 4.1.** If the property (i) of Theorem 4.1 holds for some bases of $V$ and $W$, then it holds for any bases of $V$ and $W$. 

17
Definition 4.2. (Tensor Product) A tensor product of two vector spaces \( V \) and \( W \), over the same field \( F \), is a vector space \( T \) over \( F \) with a bilinear map
\[
\otimes : V \times W \rightarrow T, \quad (x, y) \mapsto (x \otimes y)
\]
that satisfies the following condition: If \( \{v_i : i \in I\} \) and \( \{w_j : j \in J\} \) are bases of \( V \) and \( W \), respectively, then \( \{v_i \otimes w_j : i \in I, j \in J\} \) is a basis of \( T \). This condition is independent of the choice of the bases in \( V \) and \( W \).

Remark 4.2. The tensor product of vector spaces \( V \) and \( W \) in Definition 4.2 is denoted as \( V \otimes W \); if it is necessary to specify the field \( F \) over which the vector spaces are defined, then the tensor product is denoted as \( V \otimes_F W \). Furthermore, if both \( V \) and \( W \) are finite-dimensional, then it follows that
\[
\dim(V \otimes W) = \dim V \times \dim W.
\]

A tensor product exists for any two vector spaces over the same field. For example, let a vector space \( T \) have a basis \( \{t_{ij} : i \in I, j \in J\} \) and let a bilinear map be defined as: \( \otimes : V \times W \rightarrow T \) so that \( v_i \otimes w_j = t_{ij} \) for the basis vectors \( v_i \) and \( w_j \).

A tensor product is unique in the following sense: If \( (T_1, \otimes_1) \) and \( (T_2, \otimes_2) \) are two tensor products of vector spaces \( V \) and \( W \), then there exists a unique isomorphism \( \psi : T_1 \rightarrow T_2 \) such that \( \psi(x \otimes_1 y) = x \otimes_2 y \ \forall x \in V \) and \( \forall y \in W \). Indeed, for basis vectors, the isomorphism can be constructed as: \( \psi(v_i \otimes_1 w_j) = v_i \otimes_2 w_j \).

Example 4.1. Let a bilinear map \( \otimes : F[x] \times F[y] \rightarrow F[x, y] \) be defined as:
\[
(v \otimes w)(x, y) = v(x)w(y)
\]
(Note that \( F[\bullet] \) denotes a polynomial algebra over the field \( F \).) The products \( x^i \otimes y^j = x^i y^j \), where \( i, j = 0, 1, 2, \ldots \), form a basis for \( F[x, y] \). Hence,
\[
F[x, y] = F[x] \otimes F[y].
\]

Similarly, a multi-linear tensor will follow the rule:
\[
F[x_1, \ldots, x_m; y_1, \ldots, y_n] = F[x_1, \ldots, x_m] \otimes F[y_1, \ldots, y_n]
\]

Theorem 4.2. Let \( V \) and \( W \) be vector spaces over the same field \( F \). Then, for any bilinear map \( \varphi : V \times W \rightarrow U \), there exists a unique linear map \( \psi : V \otimes W \rightarrow U \) such that \( \varphi(x, y) = \psi(x \otimes y) \ \forall x \in V, \forall y \in W \).

Proof. On basis vectors of the vector space \( V \otimes W \), the linear map \( \psi : V \otimes W \rightarrow U \) is determined as:
\[
\psi(v_i \otimes w_j) = \varphi(v_i, w_j).
\]
Every element \( u \in V \otimes W \) decomposes uniquely as:
\[
u = \sum_i \sum_j u_{ij} v_i \otimes w_j \text{ where } u_{ij} \in F
\]
where the scalars \( u_{ij} \) are called the coordinates of \( u \) with respect to the given bases of \( V \) and \( W \). Specifically, in finite-dimensional vector spaces, \( u \) is described by a \((m \times n)\) matrix \((u_{ij})\), where \( m = \dim(V) \) and \( n = \dim(W) \).
Definition 4.3. (Decomposition) A vector \( u \in V \otimes W \) is called decomposable if it decomposes as:

\[
u = x \otimes y \text{ for some } x \in V \text{ and } y \in W.
\]

Remark 4.3. Let \( x = \sum_i x_i v_i \) and \( y = \sum_j y_j w_j \); then, for \( u = x \otimes y \), it follows that \( u_{ij} = x_i y_j \). This implies that, in the finite-dimensional case, \( \text{rank}(u_{ij}) \leq 1 \). Thus, decomposable elements comprise a small part of the space \( V \otimes W \) unless either \( V \) or \( W \) is one-dimensional; nevertheless they span the whole space.

Example 4.2. (Tensor Product) Let \( V \) and \( W \) be two finite-dimensional vector spaces over the same field \( F \) with \( \dim(V) = n \) and \( \dim(W) = n \). Let \( A = (a_{ij}) \) be the matrix of a linear operator \( A \) in a basis \( \{v_1, \ldots, v_n\} \) of \( V \), and let \( B = (b_{kl}) \) be the matrix of a linear operator \( B \) in a basis \( \{w_1, \ldots, w_m\} \) of \( W \). Then, the matrix of linear operator \( A \otimes B \) in the basis \( \{v_1 \otimes w_1, \ldots, v_1 \otimes w_m, v_2 \otimes w_1, \ldots, v_2 \otimes w_m, \ldots, v_n \otimes w_1, \ldots, v_n \otimes w_m\} \) of the space \( V \otimes W \) is

\[
\begin{pmatrix}
    a_{11}B & a_{12}B & \cdots & a_{1n}B \\
    a_{21}B & a_{22}B & \cdots & a_{2n}B \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{n1}B & a_{n2}B & \cdots & a_{nn}B
\end{pmatrix}
\]

which is called the tensor product of matrices \( A \) and \( B \) and is denoted as \( A \otimes B \). It follows that

- \( \text{Trace}(A \otimes B) = \text{Trace}(A) \times \text{Trace}(B) \).
- \( \text{Det}(A \otimes B) = \text{Det}(A)^m \times \text{Det}(B)^n \).