

# MATH251 - Algebra 2

Vector spaces, linear (in)dependence, span, bases; linear transformations, kernel, image, isomorphisms.

Based on lectures from Winter, 2024 by Prof. Anush Tserunyan  
Notes by Louis Meunier

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Vector Spaces . . . . .	2
1.2	Creating Spaces from Other Spaces . . . . .	4
1.3	Linear Combinations and Span . . . . .	8
1.4	Linear Dependence and Span . . . . .	11
<b>2</b>	<b>Linear Transformations</b>	<b>18</b>
2.1	Definitions . . . . .	18
2.2	Isomorphisms, Kernel, Image . . . . .	20
2.3	The Space $\text{Hom}(V, W)$ . . . . .	25
2.4	Matrix Representation of Linear Transformations, Finite Fields . . . . .	26
2.5	Matrix Representation of Linear Transformations, General Spaces . . . . .	29
2.6	Composition of Linear Transformations, Matrix Multiplication . . . . .	31
2.7	Inverses of Transformations and Matrices . . . . .	32
2.7.1	Invariant Subspaces and Nilpotent Transformations . . . . .	34

# 1 Introduction

**Remark 1.1.** *This course is about vector spaces and linear transformations between them; a vector space involves multiplication by scalars, where the scalars come from some field. We recall first examples of fields, then vector spaces, as a motivation, before presenting a formal definition.*

## 1.1 Vector Spaces

**Remark 1.2.** *Much of this is recall from [Algebra 1](#).*

### ⊛ Example 1.1: Examples of Fields

1.  $\mathbb{Q}$ ; the field of rational numbers.
2.  $\mathbb{R}$ ; the field of real numbers;  $\mathbb{Q} \subseteq \mathbb{R}$ .
3.  $\mathbb{C}$ ; the field of complex numbers;  $\mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ .
4.  $\mathbb{F}_p \equiv \mathbb{Z}/p\mathbb{Z} \equiv \{0, 1, \dots, p-1\}$ ; the (unique) field of  $p$  elements, where  $p$  prime.<sup>a</sup>
  - (a)  $p = 2$ ;  $\mathbb{F}_2 \equiv \{0, 1\}$ .
  - (b)  $p = 3$ ;  $\mathbb{F}_3 \equiv \{0, 1, 2\}$ .
  - (c)  $\dots$

---

<sup>a</sup>where  $a +_p b :=$  remainder of  $\frac{a+b}{p}$ ,  $a \cdot_p b :=$  remainder of  $\frac{a \cdot b}{p}$ .

**Remark 1.3.** *Throughout the course, we will denote an abstract field as  $\mathbb{F}$ .*

⊛ **Example 1.2: Examples of Vector Spaces**

1.  $\mathbb{R}^3 := \{(x, y, z) : x, y, z \in \mathbb{R}\}$ . We can add elements in  $\mathbb{R}^3$ , and multiply them by real scalars.
2.  $\mathbb{F}^n := \underbrace{\mathbb{F} \times \mathbb{F} \times \cdots \times \mathbb{F}}_{n \text{ times}} := \{(a_1, a_2, \dots, a_n) : a_i \in \mathbb{F}\}$ , where  $n \in \mathbb{N}^1$ ; this is a generalization of the previous example, where we took  $n = 3$ ,  $\mathbb{F} = \mathbb{R}$ . Operations follow identically; addition:

$$(a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) := (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)$$

and, taking a scalar  $\lambda \in \mathbb{F}$ , multiplication:

$$\lambda \cdot (a_1, a_2, \dots, a_n) := (\lambda \cdot a_1, \lambda \cdot a_2, \dots, \lambda \cdot a_n).$$

We refer to these elements  $(a_1, \dots, a_n)$  as *vectors* in  $\mathbb{F}^n$ ; the vector for which  $a_i = 0 \forall i$  is the *0 vector*, and is the additive identity, making  $\mathbb{F}^n$  an abelian group under addition, that admits multiplication by scalars from  $\mathbb{F}$ .

3.  $C(\mathbb{R}) := \{f : \mathbb{R} \rightarrow \mathbb{R} : f \text{ continuous}\}$ . Here, we have the constant zero function as our additive identity ( $x \mapsto 0 \forall x$ ), and addition/scalar multiplication of two continuous real functions are continuous.
4.  $\mathbb{F}[t] := \{a_0 + a_1t + a_2t^2 + \cdots + a_nt^n : a_i \in \mathbb{F} \forall i, n \in \mathbb{N}\}$ , ie, the set of all polynomials in  $t$  with coefficients from  $\mathbb{F}$ . Here, we can add two polynomials;

$$(a_0 + a_1t + \cdots + a_nt^n) + (b_0 + b_1t + \cdots + b_mt^m) := \sum_{i=0}^{\max\{n,m\}} (a_i + b_i)t^i,$$

(where we “take” undefined  $a_i/b_i$ ’s as 0; that is, if  $m > n$ , then  $a_{m-n}, a_{m-n+1}, \dots, a_m$  are taken to be 0). Scalar multiplication is defined

$$\lambda \cdot (a_0 + a_1t + a_2t^2 + \cdots + a_nt^n) := \lambda a_0 + \lambda a_1t + \lambda a_2t^2 + \cdots + \lambda a_nt^n.$$

Here, the zero polynomial is simply 0 (that is,  $a_i = 0 \forall i$ ).

<sup>1</sup>Where we take  $0 \in \mathbb{N}$ , for sake of consistency. Moreover, by convention, we define  $\mathbb{F}^0$  (that is, when  $n = 0$ ) to be  $\{0\}$ ; the trivial vector space.

### ↪ **Definition 1.1: Vector Space**

A *vector space*  $V$  over a field  $\mathbb{F}$  is an *abelian group* with an operation denoted  $+$  (or  $+_V$ ) and identity element<sup>2</sup> denoted  $0_V$ , equipped with *scalar multiplication* for each scalar  $\lambda \in \mathbb{F}$  satisfying the following axioms:

1.  $1 \cdot v = v$  for  $1 \in \mathbb{F}, \forall v \in V$ .
2.  $\alpha \cdot (\beta \cdot v) = (\alpha \cdot \beta)v, \forall \alpha, \beta \in \mathbb{F}, v \in V$ .
3.  $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v, \forall \alpha, \beta \in \mathbb{F}, v \in V$ .
4.  $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v, \forall \alpha \in \mathbb{F}, u, v \in V$ .

We refer to elements  $v \in V$  as *vectors*.

### ↪ **Proposition 1.1**

For a vector space  $V$  over a field  $\mathbb{F}$ , the following holds:

1.  $0 \cdot v = 0_V, \forall v \in V$  (where  $0 := 0_{\mathbb{F}}$ )
2.  $-1 \cdot v = -v, \forall v \in V$  (where  $1 := 1_{\mathbb{F}}$ )<sup>3</sup>
3.  $\alpha \cdot 0_V = 0_V, \forall \alpha \in \mathbb{F}$

Proof. 1.  $0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v \implies 0 \cdot v = 0_V$  (by “cancelling” one of the  $0 \cdot v$  terms on each side).

2.  $v + (-1 \cdot v) = (1 \cdot v + (-1) \cdot v) = (1 - 1) \cdot v = 0 \cdot v = 0_V \implies (-1 \cdot v) = -v$ .

3.  $\alpha \cdot 0_V = \alpha \cdot (0_V + 0_V) = \alpha \cdot 0_V + \alpha \cdot 0_V \implies \alpha \cdot 0_V = 0_V$  (by, again, cancelling a term on each side).

■

↪ Lecture 01; Last Updated: Wed Jan 10 14:16:29 EST 2024

## 1.2 Creating Spaces from Other Spaces

---

<sup>2</sup>The “zero vector”.

<sup>3</sup>NB: “additive inverse”

↪ **Definition 1.2: Product/Direct Sum of Vector Spaces**

For vector spaces  $U, V$  over the same field  $\mathbb{F}$ , we define their *product* (or *direct sum*) as the set

$$U \times V = \{(u, v) : u \in U, v \in V\},$$

with the operations:

$$(u_1, v_1) + (u_2, v_2) := (u_1 + u_2, v_1 + v_2)$$

$$\lambda \cdot (u, v) := (\lambda \cdot u, \lambda \cdot v)$$

⊗ **Example 1.3:  $\mathbb{F}$**

$\mathbb{F}^2 = \mathbb{F} \times \mathbb{F}$ , where  $\mathbb{F}$  is considered as the vector space over  $\mathbb{F}$  (itself).

↪ **Definition 1.3: Subspace**

For a vector space  $V$  over a field  $\mathbb{F}$ , a *subspace* of  $V$  is a subset  $W \subseteq V$  s.t.

1.  $0_V \in W$ <sup>4</sup>
2.  $u + v \in W \forall u, v \in W$  (closed under addition)
3.  $\alpha \cdot u \in W \forall u \in W, \alpha \in \mathbb{F}$ <sup>5</sup>

Then,  $W$  is a vector space in its own right.

---

<sup>4</sup>This is equivalent to requiring that  $W \neq \emptyset$ ; stated this way, axiom 3. would necessitate that  $0 \cdot w = 0_V \in W$ .

<sup>5</sup>Note that these axioms are equivalent to saying that  $W$  is a subgroup of  $V$  with respect to vector addition; 2. ensures closed under addition, and 3. ensures the existence of additive inverses (as per  $-1 \cdot v = -v$ ).

⊛ **Example 1.4: Examples of Subspaces**

1. Let  $V := \mathbb{F}^n$ .

- $W := \{(x_1, x_2, \dots, x_n) \in \mathbb{F}^n : x_1 = 0\} = \{(0, x_2, x_3, \dots, x_n) : x_i \in \mathbb{F}\}$ .
- $W := \{(x_1, x_2, \dots, x_n) \in \mathbb{F}^n : x_1 + 2 \cdot x_2 = 0\}$

*Proof.* Let  $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in W$ . Then,  $x + y = (x_1 + y_1, \dots, x_n + y_n)$ , and  $x_1 + y_1 + 2 \cdot (x_2 + y_2) = x_1 + 2 \cdot x_2 + y_1 + 2 \cdot y_2 = 0 + 0 = 0 \implies x + y \in W$ . Similar logic follows for axioms 2., 3. ■

- (More generally)

$$W := \{(x_1, \dots, x_n) \in \mathbb{F}^n : \begin{array}{l} a_{11}x_1 + \dots + a_{1n}x_n = 0 \\ a_{21}x_1 + \dots + a_{2n}x_n = 0 \\ \vdots \\ a_{k1}x_1 + \dots + a_{kn}x_n = 0 \end{array}\},$$

that is, a linear combination of homogenous “conditions” on each term.

- $W^* := \{(x_1, \dots, x_n) : x_1 + x_2 = 1\}$  is *not* a subspace; it is not closed under addition, nor under scalar multiplication.

2. Let  $\mathbb{F}[t]_n := \{a_0 + a_1t + \dots + a_nt^n : a_i \in \mathbb{F}\}$ . Then,  $\mathbb{F}[t]_n$  is a subspace of  $\mathbb{F}[t]$ , the more general polynomial space. *However*, the set of all polynomials of degree *exactly*  $n$  (all axioms fail, in fact) is not a subspace of  $\mathbb{F}[t]_n$ .

- $W := \{p(t) \in \mathbb{F}[t]_n : p(1) = 0\}$ .
- $W := \{p(t) \in \mathbb{F}[t]_n : p''(t) + p'(t) + 2p(t) = 0\}$ .

3. Let  $V := C(\mathbb{R})$  be the space of continuous function  $\mathbb{R} \rightarrow \mathbb{R}$ .

- $W := \{f \in C(\mathbb{R}) : f(\pi) + 7f(\sqrt{2}) = 0\}$ .
- $W := C^1(\mathbb{R}) :=$  everywhere differentiable functions.
- $W := \{f \in C(\mathbb{R}) : \int_0^1 f \, dx = 0\}$ .

### ↪ **Proposition 1.2**

Let  $W_1, W_2$  be subspaces of a vector space  $V$  over  $\mathbb{F}$ . Then, define the following:

1.  $W_1 + W_2 := \{w_1 + w_2 : w_1 \in W_1, w_2 \in W_2\}$
2.  $W_1 \cap W_2 := \{w \in V : w \in W_1 \wedge w \in W_2\}$

These are both subspaces of  $V$ .

Proof.

1. (a)  $0_V \in W_1$  and  $0_V \in W_2 \implies 0_V = 0_V + 0_V \in W_1 + W_2$ .  
(b)  $(u_1 + u_2) + (v_1 + v_2) = (u_1 + v_1) + (u_2 + v_2) \in W_1 + W_2$ .  
(c)  $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v \in W_1 + W_2$
2. (a)  $0_V \in W_1$  and  $0_V \in W_2 \implies 0_V = 0_V + 0_V \in W_1 \cap W_2$ .  
(b)  $u, v \in W_1 \cap W_2 \implies u + v \in W_1 \wedge u + v \in W_2 \implies u + v \in W_1 \cap W_2$ .  
(c)  $\alpha \cdot u \in W_1 \wedge \alpha \cdot u \in W_2 \implies \alpha \cdot u \in W_1 \cap W_2$ .

■

## 1.3 Linear Combinations and Span

### ↪ **Definition 1.4: Linear Combination**

Let  $V$  be a vector space over a field  $\mathbb{F}$ . For finitely many vectors  $v_1, v_2, \dots, v_n$ , their *linear combination* is a sum of the form

$$\sum_{i=1}^n a_i v_i = a_1 \cdot v_1 + \dots + a_n \cdot v_n,$$

where  $a_i \in \mathbb{F} \forall i$ .

A linear combination is called *trivial* if  $a_i = 0 \forall i$ , that is, all coefficients are 0.

If  $n = 0$  (ie, we are “summing up” 0 vectors), we define the sum as the zero vector;  $\sum_{i=1}^0 a_i v_i := 0_V$ .

↪ Lecture 02; Last Updated: Tue Jan 23 15:04:02 EST 2024

### ↪ **Definition 1.5: A More General Definition of Linear Combination**

For a (possibly infinite) set  $S$  of vectors from  $V$ , a *linear combination* of vectors in  $S$  is a linear combination of  $a_1 v_1 + \dots + a_n v_n$  for some finite subset  $\{v_1, \dots, v_n\} \subseteq S$ .<sup>6</sup>

<sup>6</sup>That is, we do not allow infinite sums.



↪ **Definition 1.6: Span**

For a subset  $S \subseteq V$ , we define its *span* as

$$\text{Span}(S) := \text{set of all linear combinations of } S := \{a_1v_1 + \cdots + a_nv_n : a_i \in \mathbb{F}, v_i \in S\}.$$

By convention, we set  $\text{Span}(\emptyset) = \{0_V\}$ .

⊛ **Example 1.5**

Let  $S := \{(1, 0, -1), (0, 1, -1), (1, 1, -2)\} \subseteq \mathbb{R}^3$ . Then,

$$0_{\mathbb{R}^3} = (0, 0, 0) = 1 \cdot (1, 0, -1) + 1 \cdot (0, 1, -1) + -1 \cdot (1, 1, -2).$$

We claim, moreover, that  $\text{Span}(S) = U := \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$  (a plane through the origin).

Proof. Note that  $S \subseteq U$ , hence  $S \subseteq \text{Span } S \subseteq U$ . OTOH, if  $(x, y, z) \in U$ , we have  $z = -x - y$ , and so

$$(x, y, z) = (x, y, -x - y) = x \cdot (1, 0, -1) + y \cdot (0, 1, -1) \in \text{Span}(S)$$

hence  $U \subseteq \text{Span}(S)$  and thus  $\text{Span}(S) = U$ . ■

**Remark 1.4.** We implicitly used the following claim in the proof above; we prove it more generally.

↪ **Proposition 1.3**

Let  $V$  be a vector space over  $\mathbb{F}$  and let  $S \subseteq V$ . Then,  $\text{Span}(S)$  is always a subspace. Moreover, it is the smallest (minimal) subspace containing  $S$  (that is, for any subspace  $U \supseteq S$ , we have that  $U \supseteq \text{Span}(S)$ ).

Proof. Because adding/scalar multiplying linear combinations of elements of  $S$  again results in a linear combination of elements of  $S$ , and  $0_V \in \text{Span}(S)$  by definition, we have that  $\text{Span}(S)$  is indeed a subspace.

If  $U \supset S$  is a subspace of  $V$  containing  $S$ , then by definition  $U$  is closed under addition, that is, taking linear combinations of its elements (in particular, of elements of  $S$ ); hence,  $U \supset \text{Span}(S)$ . ■

↪ **Lemma 1.1**

For  $S \subseteq V$  and  $v \in V$ ,  $v \in \text{Span}(S) \iff \text{Span}(S \cup \{v\}) = \text{Span}(S)$ .

Proof. ( $\implies$ ) Let  $v \in \text{Span}(S) \implies v = a_1v_1 + \cdots + a_nv_n, a_i \in \mathbb{F}, v_i \in S$ . Then, for any linear combination

$$b_1u_1 + \cdots + b_mu_m + b \cdot v = b_1u_1 + \cdots + b_mu_m + b(a_1v_1 + \cdots + a_nv_n)$$

is a linear combination of vectors in  $S \cup \{v\}$  (first equality) or equivalently, a combination of vectors in  $S$  (second equality) and thus  $\text{Span}(S \cup \{v\}) \subseteq \text{Span } S$ . The reverse inclusion follows trivially.

$$(\Leftarrow) \text{Span}(S \cup \{v\}) = \text{Span } S \implies v \in \text{Span}(S).$$

### ⊛ Example 1.6

(From the above example) We have

$$\text{Span}(\{(1, 0, -1), (0, 1, -1)\} \cup \{(1, 1, -2)\}) = \text{Span}(\{(1, 0, -1), (0, 1, -1)\}),$$

since  $(1, 1, -2) \in \text{Span}(\{(1, 0, -1), (0, 1, -1)\})$  (it was redundant, as it could be generated by the other two vectors).

### ↪ Definition 1.7: Spanning Set

Let  $V$  be a vector space over a field  $\mathbb{F}$ . We call  $S \subseteq V$  a *spanning set* for  $V$  if  $\text{Span}(S) = V$ . We call such a spanning set *minimal* if no proper subset of  $S$  is a spanning set ( $\nexists v \in S$  s.t.  $S \setminus \{v\}$  spanning).

**Remark 1.5.** Note that any  $S \subseteq V$  is a spanning for  $\text{Span}(S)$ . But,  $S$  may not be minimal; indeed, consider the previous example. We were able to remove a vector from  $S$  while having the same span.

### ⊛ Example 1.7

For  $\mathbb{F}^n$  as a vector space over  $\mathbb{F}$ , the *standard spanning set*

$$\text{St}_n := \{\underbrace{(1, \dots, 0)}_{:=e_1}, \underbrace{(0, 1, 0, \dots, 0)}_{:=e_2}, \dots, \underbrace{(0, \dots, 1)}_{e_n}\}.$$

Given any  $x := (x_1, \dots, x_n) \in \mathbb{F}^n$ , we can write

$$x = x_1 \cdot e_1 + \dots x_n \cdot e_n.$$

This is clearly minimal; removing any  $e_i$  would then result in a 0 in the  $i$ th “coordinate” of a vector, hence  $\text{St} \setminus \{e_i\}$  would span only vectors whose  $i$ th coordinate is 0.

### ↪ Definition 1.8: Linear Dependence

Let  $V$  be a vector space over a field  $\mathbb{F}$ . A set  $S \subseteq V$  is said to be *linearly dependent* if there is a nontrivial linear combination of vectors in  $S$  that is equal to  $0_V$ .

Conversely,  $S$  is called *linearly independent* if there is no nontrivial linear combination of vectors in  $S$  that is equal to  $0_V$ ; all linear combinations of vectors in  $S$  that equal  $0_V$  are trivial.

⊛ **Example 1.8**

1. The empty set  $\emptyset$  is linearly independent; there are no non-trivial linear combinations that equal  $0_V$  (there are no linear combinations at all).
2. For  $v \in V$ , the set  $\{v\}$  is linearly dependent iff  $v = 0_V$ .
3.  $S := \{(1, 0, -1), (0, 1, -1), (1, 1, -2)\} := \{v_1, v_2, v_3\}$ ;  $S$  is linearly dependent ( $v_1 + v_2 - v_3 = (0, 0, 0)$ ).
4.  $V := \mathbb{F}^3$ ;  $S := \{(1, 0, -1), (0, 1, -1), (0, 0, 1)\} = \{v_1, v_2, v_3\}$  is linearly independent.

Proof. Suppose

$$\begin{aligned} a_1 v_1 + a_2 v_2 + a_3 v_3 &= 0_V \\ \implies a_1 &= 0 \wedge a_2 = 0 \wedge -a_1 - a_2 + a_3 = 0 \implies a_3 = 0 \\ \implies a_1 &= a_2 = a_3 = 0 \end{aligned}$$

Hence only a trivial linear combination is possible. ■

5.  $\text{St}_n$  is linearly independent.

Proof.

$$\sum_{i=1}^n a_i e_i = 0_{\mathbb{F}^n} \implies a_i = 0 \forall i$$
■

↔ **Lemma 1.2**

Let  $V$  be a vector space over a field  $\mathbb{F}$ , and  $S \subseteq V$  (possibly infinite).

1.  $S$  is linearly dependent  $\iff$  there is a finite subset  $S_0 \subseteq S$  that is linearly dependent.
2.  $S$  is linearly independent  $\iff$  all finite subsets of  $S$  are linearly independent.

Proof. 2. follows from the negation of 1.

( $\Leftarrow$ ) Trivial.

( $\Rightarrow$ ) Suppose  $S$  linearly dependent. Then,  $0_V =$  some nontrivial linear combination of vectors  $v_1, \dots, v_n$  in  $S$ . Let  $S_0 = \{v_1, \dots, v_n\}$ , then,  $S_0$  is linearly dependent itself. ■

## 1.4 Linear Dependence and Span

↪ **Proposition 1.4**

Let  $V$  be a vector space over a field  $\mathbb{F}$  and  $S \subseteq V$ .

1.  $S$  linearly dependent  $\iff \exists v \in \text{Span}(S \setminus \{v\})$ .
2.  $S$  linearly independent  $\iff$  there is no  $v \in \text{Span}(S \setminus \{v\})$ .

Proof. 2. follows from the negation of 1.

( $\implies$ ) Suppose  $S$  linearly dependent. Then,  $0_V = \sum_{i=1}^n a_i v_i$  for some nontrivial linear combination of distinct vectors  $S$ . At least one of  $a_i \neq 0$ ; we can assume wlog (reindexing)  $a_1 \neq 0$ . Then,

$$a_1 v_1 = - \sum_{i=2}^n a_i v_i \implies v_1 = (-a_1^{-1}) \sum_{i=2}^n a_i v_i = \sum_{i=2}^n (-a_1^{-1} a_i) v_i,$$

hence,  $v_1 \in \text{Span}(\{v_2, \dots, v_n\}) \subseteq \text{Span}(S \setminus \{v\})$

( $\impliedby$ ) Suppose  $v \in \text{Span}(S \setminus \{v\})$ , then  $v = a_1 v_1 + \dots + a_n v_n$ , with  $v_1, \dots, v_n \in S \setminus \{v\}$ , thus

$$0_V = a_1 v_1 + \dots + a_n v_n - v,$$

which is not a trivial combination ( $-1$  on the  $v$ ;  $v$  cannot “merge” with the other vectors), hence  $S$  is linearly dependent. ■

↪ **Corollary 1.1**

$S \subseteq V$  is linearly independent  $\iff S$  a minimal spanning set of  $\text{Span } S$ .

Proof. Follows from proposition 1.4, 2. ■

↪ **Definition 1.9: Maximally Independent**

Let  $V$  be a vector space over a field  $\mathbb{F}$ . A set  $S \subseteq V$  is called *maximally independent* if  $S$  is linearly independent and  $\nexists v \in V \setminus S$  s.t.  $S \cup \{v\}$  is still linearly independent.

In other words, there is no proper supset  $\tilde{S} \supsetneq S$  that is still independent.

↪ **Lemma 1.3**

If  $S \subseteq V$  maximally independent, then  $S$  is spanning for  $V$ .

Proof. Let  $S \subseteq V$  be maximally independent. Let  $v \in V$ ; supposing  $v \notin S$  (in the case that  $v \in S$ , then  $v \in \text{Span}(S)$  trivially). By maximality,  $S \cup \{v\}$  is linearly dependent, hence there exists a nontrivial linear combination that equals

$0_V$ . Since  $S$  independent, this combination must include  $v$ , with a nonzero coefficient. We can write

$$av + \sum_{i=1}^n a_i v_i = 0_V \quad a \neq 0, v_i \in S$$

$$\implies v = \sum_{i=1}^n (-a^{-1}a_i)v_i \in \text{Span } S.$$

■

### ↪ **Theorem 1.1**

Let  $V$  be a vector space over a field  $\mathbb{F}$  and let  $S \subseteq V$ . TFAE:

1.  $S$  is a minimal spanning set;
2.  $S$  is linearly independent and spanning;
3.  $S$  is a maximally linearly independent set;
4. Every vector in  $V$  is equal to *unique* linear combination of vectors in  $S$ .

↪ Lecture 04; Last Updated: Fri Jan 26 13:32:10 EST 2024

Proof. (1.  $\implies$  2.) Suppose  $S$  is spanning for  $V$  and is minimal. Then, by corollary 1.1, we have that  $S$  is linearly independent, and is thus both linearly independent and spanning.

(2.  $\implies$  3.) Suppose  $S$  is linearly independent and spanning. Let  $v \in V \setminus S$ ;  $S$  is spanning, hence  $v \in \text{Span } S$ , that is, there exists a linear combination of vectors in  $S$  that is equal to  $v$ :

$$v = a_1 v_1 + \cdots + a_n v_n, a_i \in \mathbb{F}, v_i \in S.$$

Thus,  $0_V = a_1 v_1 + \cdots + a_n v_n - v$ , thus  $S \cup \{v\}$  is linearly dependent, and so  $S$  is maximally linearly independent.

(3.  $\implies$  1.) Suppose  $S$  is maximally linearly independent. By lemma 1.3,  $S$  is spanning, and since  $S$  is linearly independent, by corollary 1.1,  $S$  is minimally spanning for  $\text{Span } S$ .

(2.  $\implies$  4.) Suppose  $S$  is linearly independent and spans  $V$ , and let  $v \in V$ . We have that  $v \in \text{Span } S$  and hence is equal to a linear combination of vectors in  $S$ . This gives existence; we now need to prove uniqueness.

Suppose there exist two linear combinations that equal  $v$ ,

$$v = a_1 v_1 + \cdots + a_n v_n = b_1 u_1 + \cdots + b_m u_m,$$

$a_i, b_j \in \mathbb{F}, v_i, u_j \in S$ . With appropriate reindexing/relabelling and allowing certain scalars to equal 0, we can assume that the combinations use the same vectors (with potentially different coefficients), that is,

$$v = a_1 w_1 + \cdots + a_k w_k = b_1 w_1 + \cdots + a_k w_k.$$

This implies, then,

$$(a_1 - b_1)w_1 + \cdots + (a_k - b_k)w_k = 0_V,$$

and by the assumed linear independent of  $S$ , each coefficient  $(a_i - b_i) = 0 \forall i \implies a_i = b_i \forall i$ , hence, these are indeed the same representations, and thus this representation is unique.

(4.  $\implies$  2.) Suppose every vector in  $V$  admits a unique linear combination of vectors in  $S$ . Clearly, then,  $S$  is spanning. It remains to show  $S$  is linearly independent. Suppose

$$0_V = a_1 v_1 + \cdots + a_n v_n$$

for  $v_i \in S$ . But we have that every vector has a unique representation, and we know that  $a_i = 0 \forall i$  is a (valid) linear combination that gives  $0_V$ ; hence, this must be the unique combination,  $a_i = 0 \forall i$ , and the linear combination above is trivial. Hence,  $S$  is linearly independent and spanning. ■

### ↪ **Definition 1.10: Basis**

If any (hence all) of the above statements hold, we call  $S$  a *basis* for  $V$ .

In the words of 4., we call the unique linear combination of vectors in  $S$  that is equal to  $v$  the *unique representation of  $v$  in  $S$* . Its coefficients are called the *Fourier coefficients of  $v$  in  $S$* .

### ⊗ **Example 1.9**

1.  $\text{St}_n = \{e_i : 1 \leq i \leq n\}$  is a basis for  $\mathbb{F}^n$ .

2. In  $\mathbb{F}^3$ , the set

$$\{(1, 0, -1), (0, 1, -1), (0, 0, 1)\}$$

is a basis; it is linearly independent and spanning.

3. For  $\mathbb{F}[t]_n$ , the standard basis is

$$\{1, t, t^2, \dots, t^n\}.$$

4. For  $\mathbb{F}[t]$ , the standard basis is

$$S := \{1, t, t^2, \dots\} = \{t^n : n \in \mathbb{N}\}.$$

5. Let  $\mathbb{F}[[t]]$  denote the space of all formal power series  $\sum_{n \in \mathbb{N}} a_n t^n$ ; polynomials are an example, but with only finite nonzero coefficients. Note that, then, the set  $S$  defined above is not a basis for this “extended” set. We *can* in fact find a basis for this set; we need more tools first.

### ↪ **Theorem 1.2**

Every vector space has a basis.

**Remark 1.6.** *This theorem relies on assuming the Axiom of Choice.*

*Proof (Attempt).* (Of theorem 1.2) We will try to “inductively” build a maximally independent set, as follows:

Begin with an empty set  $S_0 := \emptyset$ , and iteratively add more vectors to it. Let  $v_0 \in V$  be a non-zero vector, and let  $S_1 := \{v_0\}$ .

If  $S_1$  is maximal, then we are done. Otherwise, there exists a new vector  $v_1 \in V \setminus S_1$  s.t.  $S_2 := \{v_0, v_1\}$  is still independent.

If  $S_2$  is maximal, then we are done. Otherwise, there exists a new vector  $v_2 \in V \setminus S_2$  s.t.  $S_3 := \{v_0, v_1, v_2\}$  is still independent.

Continue in this manner; this would take arbitrarily many finite, or even infinite, steps; we would need some “choice function” that would “allow” us to choose any particular  $i$ th vector  $v_i$ .

We can make this construction precise via the Axiom of Choice and transfinite induction (on ordinals); alternatively, we will prove a statement equivalent to the Axiom of Choice, Zorn’s Lemma. ■

**Remark 1.7.** Before stating Zorn’s Lemma, we introduce the following terminology.

### ↪ **Axiom 1.1: Axiom of Choice**

Let  $X$  be a set of nonempty sets. Then, there exists a choice function  $f$  defined on  $X$  that maps each set of  $X$  to an element of that set.

### ↪ **Definition 1.11: Inclusion-Maximal Element**

A *inclusion-maximal* element of  $I$  is a set  $S \in I$  s.t. there is no strict super set  $S' \supsetneq S$  s.t.  $S' \in I$ .

### ↪ **Definition 1.12: Chain**

Let  $X$  a set. Call a collection  $\mathcal{C} \subseteq \mathcal{P}(X)$  a *chain* if any two  $A, B \in \mathcal{C}$  are comparable, ie,  $A \subseteq B$  or  $B \subseteq A$ .

### ↪ **Definition 1.13: Upper Bound**

An *upper bound* of a collection  $\tau \subseteq \mathcal{P}(X)$  is a set  $U \subseteq X$  s.t.  $U \supseteq J \forall J \in \tau$ ;  $U$  contains the union of all sets in  $J$ .

### ⊛ **Example 1.10: Of The Previous Definitions**

Let  $X := \mathbb{N}$ ,  $I := \{\emptyset, \{0\}, \{1, 2\}, \{1, 2, 3\}\} \subseteq \mathcal{P}(\mathbb{N})$ .

The maximal elements of  $I$  would be  $\{0\}$  and  $\{1, 2, 3\}$ .

Chains would include  $\mathcal{C}_0 := \{\emptyset, \{1, 2\}, \{1, 2, 3\}\}$ ,  $\mathcal{C}_1 := \{\emptyset, \{0\}\}$ ,  $\mathcal{C}_2 := \{\emptyset\}$  (or any set containing a single element).

The sets  $\{0, 1, 2, 3\}$  and  $\{0, 1, 2, 3, 4, 5\}$  are upper bounds for  $I$ , while neither is an element of  $I$ . The set  $\{1, 2, 3\}$  is an upper bound for  $\mathcal{C}_0$ . A chain  $\{\emptyset, \{0\}, \{0, 1\}, \{0, 1, 2\}, \dots\}$  has an upper bound of  $\mathbb{N}$ .

### ↪ **Lemma 1.4: Zorn's Lemma**

Let  $X$  be an ambient set and  $I \subseteq \mathcal{P}(X)$  be a nonempty collection of subsets of  $X$ . If every chain  $\mathcal{C} \subseteq I$  has an upper bound in  $I$ , then  $I$  has a maximal element.

*“Proof”.* This is equivalent to the Axiom of Choice; proving it is beyond the scope of this course :( ■

*Proof of theorem 1.2, cnt'd.* We obtain a maximal independent set using Zorn's Lemma.

Let  $I$  be the collection of all linearly independent subsets of  $V$ .  $I$  is nonempty;  $\emptyset \in I$ , as is  $\{v\} \in I$  for any nonzero  $v \in V$ . To apply Zorn's, we need to show that every chain  $\mathcal{C}$  of sets in  $I$  has an upper bound in  $I$ ; that is, every linearly independent set has an upper bound that itself is linearly independent.

Let  $\mathcal{C}$  be a chain in  $I$ . Let  $S := \bigcup \mathcal{C}$  be the union of all sets in  $\mathcal{C}$ . To show  $S$  is linearly independent, it suffices to show that every finite subset  $\{v_1, \dots, v_n\} \subseteq S$  is linearly independent. Let  $S_i \in \mathcal{C}$  be s.t.  $v_i \in S_i$  for each  $i$ . Because  $\mathcal{C}$  a chain, for each  $i, j$  we have either  $S_i \subseteq S_j$  or  $S_j \subseteq S_i$ , and so we can order  $S_1, \dots, S_n$  in increasing order w.r.t  $\subseteq$ . This implies, then, there is a maximal  $S_{i_0}$  s.t.  $S_{i_0} \supseteq S_i \forall i \in \{1, \dots, n\}$ . Moreover, we have that  $\{v_1, \dots, v_n\} \in S_{i_0}$ , and that  $S_{i_0}$  is linearly independent and thus  $\{v_1, v_2, \dots, v_n\}$  is also linearly independent.

Thus, as we can apply Zorn's Lemma, we conclude that  $I$  has a maximal element, ie, there is a maximal independent set, and thus a  $V$  indeed has a basis. ■

↪ Lecture 06; Last Updated: Fri Jan 19 13:36:58 EST 2024

### ↪ **Theorem 1.3**

For every vector space  $V$  over a field  $\mathbb{F}$ , any two bases  $\mathcal{B}_1, \mathcal{B}_2$  are equinumerous/of equal size/cardinality, ie, there is a bijection between  $\mathcal{B}_1$  and  $\mathcal{B}_2$ .

**Remark 1.8.** We will only prove this for vector spaces that admit a finite basis.

### ↪ **Lemma 1.5: Steinitz Substitution**

Let  $V$  be a vector space over a field  $\mathbb{F}$ . Let  $Y \subseteq V$  be a (possibly infinite) linearly independent set and let  $Z \subseteq V$  be a finite spanning set. Then:

1.  $k := |Y| \leq |Z| =: n$
2. There is  $Z' \subseteq Z$  of size  $n - k$  s.t.  $Y \cup Z'$  is still spanning.

*Proof.* We prove by induction on  $k$ .

$k = 0$  gives that  $Y = \emptyset$ , and so  $Z' = Z$  itself works ( $Z' \cup Y = Z$ ) as a spanning set.

Suppose the statement holds for some  $k \geq 0$ . Let  $Y$  be an independent set such that  $|Y| = k + 1$ , ie

$$Y := \{y_1, y_2, \dots, y_k, y_{k+1}\}, \quad y \in V.$$

By our inductive assumption, we can consider  $Y' := \{y_1, \dots, y_k\} \subseteq Y$  of size  $k$ , to obtain a set

$$Z' = \{z_1, z_2, \dots, z_{n-k}\} \subseteq Z, \text{ s.t. } Y' \cup Z' = \{y_1, \dots, y_k, z_1, \dots, z_{n-k}\}$$



is spanning. As this is spanning, we can write  $y_{k+1}$  as a linear combination of vectors in  $Y' \cup Z'$ , ie

$$y_{k+1} = a_1 y_1 + \cdots + a_k y_k + b_1 z_1 + \cdots + b_{n-k} z_{n-k}, \quad a_i, b_j \in \mathbb{F}.$$

It must be that at least one of  $b_j$ 's must be nonzero; if they were all zero, then  $y_{k+1}$  would simply be a linear combination of vector  $y_i$  giving that  $y_{k+1}$  linearly dependent, contradicting our construction of  $Y$  linearly independent.

Assume, wlog,  $b_{n-k} \neq 0$ . Then, we can write

$$z_{n-k} = b_{n-k}^{-1} y_{k+1} - b_{n-k}^{-1} a_1 y_1 - \cdots - b_{n-k}^{-1} a_k y_k - b_{n-k}^{-1} b_1 z_1 - \cdots - b_{n-k}^{-1} b_{n-k-1} z_{n-k-1},$$

and hence

$$z_{n-k} \in \text{Span}\{y_1, \dots, y_{k+1}, z_1, \dots, z_{n-k-1}\} = \text{Span} \left( \underbrace{\{y_1, \dots, y_{k+1}\}}_Y \cup \underbrace{\{z_1, \dots, z_{n-k-1}\}}_{:=Z''} \right).$$

We had that  $Y' \cup Z'$  was spanning, and  $(Y' \cup Z') \setminus (Y \cup Z'') = \{z_{n-k}\} \subseteq \text{Span}(Y \cup Z'')$ , and we thus have that  $Y \cup Z''$  is also spanning. ■

### ↔ **Corollary 1.2: Finite Basis Case for theorem 1.3**

Let  $V$  be a vector space that admits a finite basis. Then, any two bases of  $V$  are equinumerous.

*Proof.* Let  $Y, Z$  be two finite bases for  $V$ . Then,  $Y$  is independent and  $Z$  is spanning, so by Steinitz Substitution,  $|Y| \leq |Z|$ . OTOH,  $Z$  is independent, and  $Y$  is spanning, so by Steinitz Substitution,  $|Z| \leq |Y|$ , and we conclude that  $|Y| = |Z|$ . Let  $n := |Y|$ .

It remains to show that there exist no infinite bases for  $V$ ; it suffices to show that there is no independent set of size  $n + 1$ . To this end, let  $I \subseteq V$  such that  $|I| = n + 1$  be an independent set.  $Y$  is still spanning, hence, by the substitution lemma,  $n + 1 \leq n$ , a contradiction. Hence,  $I$  as defined cannot exist and so any basis of  $V$  must be of size  $n$ . ■

### ↔ **Definition 1.14: Dimension**

Let  $V$  be a vector space over a field  $\mathbb{F}$ . The *dimension* of  $V$ , denote

$$\dim(V)$$

as the cardinality/size of any basis for  $V$ . We call  $V$  *finite dimensional* if  $\dim(V)$  is a natural number, i.e.  $V$  admits a finite basis. Otherwise, we say  $V$  is infinite dimensional.

### ↪ **Corollary 1.3: of Steinitz Substitution**

Let  $V$  be a finite dimensional vector space over  $\mathbb{F}$  and denote  $n := \dim(V)$ . Then:

1. Every linearly independent subset  $I \subseteq V$  has size  $\leq n$ ;
2. Every spanning set  $S \subseteq V$  for  $V$  has size  $\geq n$ ;
3. Every independent set  $I$  can be completed to a basis to  $V$ , ie, there exists a basis  $B$  for  $V$  s.t.  $I \subseteq B$ .

Proof. Fix a basis  $B$  for  $V$ ,  $|B| = n$ .

1. If  $I$  is a independent set, then because  $B$  spanning, Steinitz Substitution gives  $|I| \leq |B|$ .
2. If  $S$  spanning for  $V$ , then because  $B$  is linearly independent, Steinitz Substitution gives  $|B| \leq |S|$ .
3. Let  $I$  be an independent set. Then, because  $B$  is spanning, Steinitz Substitution gives  $B' \subseteq B$  of size  $n - |I|$  s.t.  $I \cup B'$  is spanning. Moreover,  $|I \cup B'| \leq n$ , and by 2. it must have size  $\geq n$ , and thus has size precisely  $n$  and is thus a minimally spanning set and thus a basis.

■

### ↪ **Corollary 1.4: Monotonicity of Dimension**

Let  $V$  be a vector space over a field  $\mathbb{F}$ . For any subspace  $W \subseteq V$ ,  $\dim W \leq \dim V$ , and

$$\dim W = \dim V \iff W = V.$$

Proof. Let  $B \subseteq W$  be a basis for  $W$ . Because  $B$  is independent,  $|B| \leq \dim(V)$  by 1. of corollary 1.3, so  $\dim(W) = |B| \leq \dim(V)$ .

If  $|B| = \dim(V)$ , then  $B$  is a basis for  $V$  again by 1. of corollary 1.3, so  $W = \text{Span}(B) = V$ .

■

↪ Lecture 07; Last Updated: Mon Jan 22 13:43:44 EST 2024

## 2 Linear Transformations

### 2.1 Definitions

### ↪ Definition 2.1: Linear Transformation

Let  $V, W$  be vector spaces over a field  $\mathbb{F}$ . A function  $T : V \rightarrow W$  is called a *linear transformation* if it preserves the vector space structures, that is,

1.  $T(v_0 + v_1) = T(v_0) + T(v_1), \forall v_0, v_1 \in V$ ;
2.  $T(\alpha \cdot v) = \alpha \cdot T(v), \forall \alpha \in \mathbb{F}, v \in V$ ;
3.  $T(0_V) = 0_W$ .

**Remark 2.1.** Note that 3. is redundant, implied by 2., but included for emphasis:

$$T(0_V) = T(0_{\mathbb{F}} \cdot 0_V) = 0_{\mathbb{F}} \cdot T(0_V) = 0_W.$$

### ⊛ Example 2.1: Linear Transformations

1.  $T : \mathbb{F}^2 \rightarrow \mathbb{F}^2, T(a_1, a_2) := (a_1 + 2a_2, a_1)$ .
2. Let  $\theta \in \mathbb{R}$ , and let  $T_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  be the rotation by  $\theta$ . The linearity of this is perhaps most obvious in polar coordinates, ie  $v \in \mathbb{R}^2, v = r(\cos \alpha, \sin \alpha)$  for appropriate  $r, \alpha$ , and  $T_\theta(v) = r(\cos(\alpha + \theta), \sin(\alpha + \theta))$ .
3.  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , a reflection about the  $x$ -axis, ie,  $T(x, y) = (x, -y)$ .
4. Projections,  $T : \mathbb{F}^n \rightarrow \mathbb{F}^n$ .
5. The transpose on  $M_n(\mathbb{F})$ , ie,  $T : M_n(\mathbb{F}) \rightarrow M_n(\mathbb{F})$ , where  $A \mapsto A^t$ .
6. The derivative on space of polynomials of degree leq  $n$ ,  $D : \mathbb{F}[t]_{n+1} \rightarrow \mathbb{F}[t]_n, p(t) \mapsto p'(t)$ .

### ↪ Theorem 2.1

Linear transformations are completely determined by their values on a basis.

That is, let  $\mathcal{B} := \{v_1, \dots, v_n\}$  be a basis for a vector space  $V$  over  $\mathbb{F}$ . Let  $W$  also be a vector space over  $\mathbb{F}$  and let  $w_1, \dots, w_n \in W$  be arbitrary vectors. Then, there is a unique linear transformation  $T : V \rightarrow W$  s.t.  $T(v_i) = w_i \forall i = 1, \dots, n$ .

Proof. We aim to define  $T(v)$  for arbitrary  $v \in V$ . We can write

$$v = a_1 v_1 + \dots + a_n v_n$$

as the unique representation of  $v$  in terms of the basis  $\mathcal{B}$ . Then, we simply define

$$T(v) := a_1 w_1 + \dots + a_n w_n,$$

for our given  $w_i$ 's. Then,  $T(v_i) = 1 \cdot w_i = w_i$ , as desired, and  $T$  is linear;

1. Let  $u, v \in V$ ;  $u := \sum_n a_i v_i, v := \sum_n b_i v_i$ . Then,

$$T(u + v) = T\left(\sum_n a_i v_i + \sum_n b_i v_i\right) = T\left(\sum_n (a_i + b_i) v_i\right) = \sum_n (a_i + b_i) w_i = \sum_n a_i w_i + \sum_n b_i w_i = T(u) + T(v).$$

2. Scalar multiplication follows similarly.

To show uniqueness, suppose  $T_0, T_1$  are two linear transformations satisfying  $T_0(v_i) = w_i = T_1(v_i)$ . Let  $v \in V$ , and write  $v = \sum_n a_i v_i$ . By linearity,

$$T_k(v) = T_k\left(\sum_n a_i v_i\right) = \sum_n a_i T_k(v_i) = \sum_n a_i w_i,$$

for  $k = 0, 1$ , hence,  $T_1(v) = T_0(v)$  for arbitrary  $v$ , hence the transformations are equivalent. ■

### ↪ **Definition 2.2: Some Important Transformations**

We denote  $T_0 : V \rightarrow W$  by  $T_0(v) := 0_W \forall v \in V$  the *zero transformation*. We denote  $I_V : V \rightarrow V, I_V(v) := v \forall v \in V$ , as the *identity transformation*.

↪ Lecture 08; Last Updated: Thu Jan 25 12:38:49 EST 2024

## 2.2 Isomorphisms, Kernel, Image

### ↪ **Definition 2.3: Isomorphism**

Let  $V, W$  be vector spaces over  $\mathbb{F}$ . An *isomorphism* from  $V$  to  $W$  is a linear transformation  $T : V \rightarrow W$  (a homomorphism for vector spaces) which admits an inverse  $T^{-1}$  that is also linear.

If such an isomorphism exists, we say  $V$  and  $W$  are *isomorphic*.

### ↪ **Proposition 2.1**

$T : V \rightarrow W$  is an isomorphism  $\iff T$  is linear and bijective.

Proof. The direction  $\implies$  is trivial.

Suppose  $T : V \rightarrow W$  is linear and bijective, ie  $T^{-1}$  exists. We need to show that  $T^{-1}$  is linear. Let  $w_1, w_2 \in W, a_1, a_2 \in \mathbb{F}$ . Then:

$$\begin{aligned} T^{-1}(a_1 w_1 + a_2 w_2) &= T^{-1}(a_1 T(T^{-1}(w_1)) + a_2 T(T^{-1}(w_2))) \\ \text{(by linearity of } T) &= T^{-1}(T(a_1 T^{-1}(w_1) + a_2 T^{-1}(w_2))) \\ &= a_1 T^{-1}(w_1) + a_2 T^{-1}(w_2). \end{aligned}$$

**Remark 2.2.** This proposition holds for all structures that only have operations; it does not for those with relations, such as graphs, orders, etc..

↪ **Theorem 2.2**

For  $n \in \mathbb{N}$ , every  $n$ -dimensional vector space  $V$  over  $\mathbb{F}$  is isomorphic to  $\mathbb{F}^n$ . In particular, all  $n$ -dim vector spaces over  $\mathbb{F}$  are isomorphic.

Proof. Fix a basis  $\mathcal{B} := \{v_1, \dots, v_n\}$  for  $V$ , and let  $T : V \rightarrow \mathbb{F}^n$  be the unique linear transformation determined by  $\mathcal{B}$  with  $T(v_i) = e_i$ , where  $\{e_1, \dots, e_n\}$  is the standard basis for  $\mathbb{F}^n$ . We show that  $T$  is a bijection.

(Injective) Suppose  $T(x) = T(y)$ ,  $x, y \in V$ . Write  $x = a_1v_1 + \dots + a_nv_n$ ,  $y = b_1v_1 + \dots + b_nv_n$ , the unique representation of  $x, y$  in the basis  $\mathcal{B}$ . We have:

$$a_1e_1 + \dots + a_ne_n = a_1T(v_1) + \dots + a_nT(v_n) = T(a_1v_1 + \dots + a_nv_n) = T(x) = T(y) = \dots = b_1e_1 + \dots + b_ne_n,$$

but by the uniqueness of representation in a basis, it follows that each  $a_i = b_i$ , hence,  $x = y$ .

(Surjective) Let  $w \in \mathbb{F}^n$ . Then,  $w = a_1e_1 + \dots + a_ne_n$  (uniquely). But then,

$$w = a_1T(v_1) + \dots + a_nT(v_n) = T(a_1v_1 + \dots + a_nv_n),$$

where  $a_1v_1 + \dots + a_nv_n \in V$ , hence  $T$  indeed surjective. ■

**Remark 2.3.** Replacing  $\mathbb{F}^n$  with an arbitrary  $n$ -dim vector space  $W$  over  $\mathbb{F}$  yields the following.

↪ **Theorem 2.3: Freeness of Vector Space**

Let  $W, V$  be vector spaces over  $\mathbb{F}$  and let  $\beta, \gamma$  be bases for  $V, W$  respectively. Every bijection  $T : \beta \rightarrow \gamma$  can be extended to an isomorphism  $\hat{T} : V \rightarrow W$ .

In particular, all vector spaces over  $\mathbb{F}$  with equinumerous bases are isomorphic.

**Remark 2.4.** The proof follows very similarly to the previous theorem, but extended to arbitrary, possibly infinite, spaces. ■

Proof.

↪ **Definition 2.4: Image/Kernel**

For a linear transformation  $T : V \rightarrow W$ , where  $V, W$  are vector spaces over  $\mathbb{F}$ , we define the *image*

$$\text{Im}(T) := T(V),$$

and its *kernel*

$$\text{Ker}(T) = T^{-1}(\{0_W\}).$$

↪ **Proposition 2.2**

$\text{Ker}(T)$  and  $\text{Im } T$  are subspaces of  $V, W$  resp.

Proof. ( $\text{Ker}(T)$ ) Let  $v_0, v_1 \in \text{Ker } T$  and  $a_0, a_1 \in \mathbb{F}$ , then

$$T(a_0v_0 + a_1v_1) = a_0T(v_0) + a_1T(v_1) = 0_W \implies a_0v_0 + a_1v_1 \in \text{Ker } T.$$

( $\text{Im}(T)$ ) Let  $w_0, w_1 \in \text{Im } T$ ,  $a_0, a_1 \in \mathbb{F}$ . Then  $w_i = T(v_i)$ ,  $v_i \in V$ , and so

$$a_0w_0 + a_1w_1 = a_0T(v_0) + a_1T(v_1) = T(a_0v_0 + a_1v_1) \implies a_0w_0 + a_1w_1 \in \text{Im } T.$$

■

↪ **Proposition 2.3**

Let  $T : V \rightarrow W$  be a linear transformation, where  $V, W$  vector spaces over  $\mathbb{F}$ . Let  $\beta$  be a (possibly infinite) basis for  $V$ . Then,  $T(\beta)$  spans  $\text{Im}(T)$ .

In particular,  $T$  is surjective iff  $T(\beta)$  spans  $W$ .

Proof. Let  $w \in \text{Im}(T)$ , so  $w = T(v)$  for some  $v \in V$ , where we have  $v := a_1v_1 + \cdots + a_nv_n$ ,  $v_i \in \beta$ . Then,

$$w = T(v) = a_1T(v_1) + \cdots + a_nT(v_n) \in \text{Span}(\{T(v_1), \dots, T(v_n)\}) \subseteq \text{Span}(T(\beta)).$$

■

↪ Lecture 09; Last Updated: Fri Jan 26 13:39:26 EST 2024

↪ **Proposition 2.4**

Let  $T : V \rightarrow W$  be a linear transformation, where  $V, W$  vector spaces over  $\mathbb{F}$ . TFAE:

1.  $T$  is injective.
2.  $\text{Ker}(T)$  is the trivial subspace  $\{0_V\}$ .
3.  $T(\beta)$  is independent for each basis  $\beta$  for  $V$ .
- 3'.  $T(\beta)$  is independent for some basis  $\beta$  for  $V$ .

Proof. (1.  $\implies$  2.) Trivial; only  $0_V$  can be mapped to  $0_W$ .

(2.  $\implies$  1.) Suppose  $\text{Ker}(T) = \{0_V\}$  and let  $T(x) = T(y)$ ,  $x, y \in V$ . By linearity,

$$T(x - y) = T(x) - T(y) = 0_W \implies x - y \in \text{Ker}(T) \implies x - y = 0_V \implies x = y.$$

(2.  $\implies$  3.) Fix a basis  $\beta$  for  $V$ . To show that  $T(\beta)$  linearly independent, take an arbitrary linear combination  $a_1w_1 + \cdots + a_nw_n \in T(\beta)$ . Suppose  $\sum_i a_iw_i = 0_W$ . Since  $w_i \in T(\beta)$ ,  $w_i = T(v_i)$ ,  $v_i \in \beta$ , hence

$$\begin{aligned} 0_W &= a_1w_1 + \cdots + a_nw_n = a_1T(v_1) + \cdots + a_nT(v_n) = T(a_1v_1 + \cdots + a_nv_n) \\ &\implies a_1v_1 + \cdots + a_nv_n \in \text{Ker}(T) \\ &\implies a_1v_1 + \cdots + a_nv_n = 0_V, \end{aligned}$$

but each  $v_i$  is linearly independent, hence this must be a trivial linear combination, and thus  $a_i = 0 \forall i$ .

(3)  $\implies$  (3') Trivial; stronger statement implies weaker statement.

(3')  $\implies$  (2) Suppose  $T(\beta)$  linearly independent for some basis  $\beta$  for  $V$ . Suppose  $T(v) = 0_W$ ,  $v \in V$ . We write

$$v = a_1v_1 + \cdots + a_nv_n, v_i \in \beta.$$

Then,

$$0_W = T(v) = T(a_1v_1 + \cdots + a_nv_n) = a_1T(v_1) + \cdots + a_nT(v_n),$$

but  $\{T(v_i)\} \subseteq T(\beta)$  is linearly independent, hence, this combination must be trivial and each  $a_i = 0$ , and thus  $v = 0_V$  and so  $\text{Ker}(T) = \{0_V\}$  is trivial. ■

#### $\hookrightarrow$ **Definition 2.5: Rank, nullity**

Let  $V, W$  be vector spaces over  $\mathbb{F}$  and  $T : V \rightarrow W$  be linear. Define *rank* of  $T$  as

$$\text{rank}(T) := \dim(\text{Im}(T)),$$

and *nullity* of  $T$  as

$$\text{nullity}(T) := \dim(\text{Ker}(T)).$$

#### $\hookrightarrow$ **Theorem 2.4: Rank-Nullity Theorem**

Let  $V, W$  be vector spaces over  $\mathbb{F}$ ,  $\dim(V) < \infty$ . Let  $T : V \rightarrow W$  be a linear transformation. Then,

$$\text{nullity}(T) + \text{rank}(T) = \dim(V).$$

**Remark 2.5.** Intuitively: the nullity is the number of vectors we “collapse”; the rank is what is left. Together, we have the entire space.

**Remark 2.6.** This follows directly from the first isomorphism theorem for vector spaces, and the fact that  $\dim(V / \text{Ker}(T)) = \dim(V) - \dim(\text{Ker}(T))$ ; however, we will prove it without this result below.

*Proof.* Let  $\{v_1, \dots, v_k\}$  be a basis for  $\text{Ker}(T)$ , and complete it to a basis  $\beta := \{v_1, \dots, v_k, u_1, \dots, u_{n-k}\}$  for  $V$ , where  $n := \dim(V)$ . We need to show that  $\dim(\text{Im}(T)) = n - k$ .

Recall that  $\{T(v_1), \dots, T(v_k), T(u_1), \dots, T(u_{n-k})\}$  spans  $\text{Im}(T)$ . But  $v_1, \dots, v_k \in \text{Ker}(T)$ , so  $T(v_i) = 0_W \forall i = 1, \dots, k$ . Hence, letting  $\gamma := \{T(u_1), \dots, T(u_{n-k})\}$  spans  $\text{Im}(T)$ . It remains to show that  $\gamma$  is independent.

Let  $a_1T(u_1) + \cdots + a_{n-k}T(u_{n-k}) = 0_W$ ; by linearity,

$$\begin{aligned} T(a_1u_1 + \cdots + a_{n-k}u_{n-k}) &= 0_W \\ \implies a_1u_1 + \cdots + a_{n-k}u_{n-k} &\in \text{Ker}(T) \\ \implies a_1u_1 + \cdots + a_{n-k}u_{n-k} &= b_1v_1 + \cdots + b_kv_k, \end{aligned}$$

but each of these  $u_i, v_j \in \beta$ , hence, each coefficient must be identically zero as  $\beta$  linearly independent, and thus  $\dim(\text{Im}(T)) = n - k$ . This completes the proof. ■

### ↪ **Corollary 2.1: Pigeonhole Principle for Dimension**

Let  $T : V \rightarrow W$  be a linear transformation. If  $T$  injective, then  $\dim(W) \geq \dim(V)$ .

*Proof.* If  $\dim(V) < \infty$ , then  $\dim(\text{Im}(T)) = \dim(V)$ , and we have that  $\dim(\text{Im}(T)) \leq \dim(W)$  and conclude  $\dim(V) \leq \dim(W)$ .

If  $\dim(V) = \infty$ , then  $\dim(\text{Im}(T)) = \infty$  and  $\dim(W) \geq \dim(\text{Im}(T)) = \infty$ . ■

### ↪ **Corollary 2.2**

Let  $n \in \mathbb{N}$  and  $V, W$  be  $n$ -dimensional vector spaces over  $\mathbb{F}$ . For a linear transformation  $T : V \rightarrow W$ , TFAE:

1.  $T$  injective;
2.  $T$  surjective;
3.  $\text{rank}(T) = n$ .

*Proof.* (2.  $\iff$  3.) Follows from  $\text{rank}(T) = \dim(\text{Im}(T)) = n \iff \text{Im}(T) = W$ .

(1.  $\implies$  3.) We have  $\text{nullity}(T) = 0$  so  $\text{rank}(T) = \dim(V) = n$ .

(3.  $\implies$  1.) If  $\text{rank}(T) = n$ , then  $\text{nullity}(T) = 0$ . ■

↪ Lecture 10; Last Updated: Mon Feb 5 14:03:23 EST 2024

### ↪ **Theorem 2.5: First Isomorphism Theorem for Vector Spaces**

Let  $V, W$  be vector spaces over  $\mathbb{F}$ . Let  $T : V \rightarrow W$  be a linear transformation. Then,

$$V / \text{Ker}(T) \cong \text{Im}(T),$$

by the isomorphism given by  $v + \text{Ker}(T) \mapsto T(v)$ .

*Proof.* From group theory, we know that  $\hat{T} : V / \text{Ker}(T) \rightarrow \text{Im}(T)$ , where  $\hat{T}(v + \text{Ker}(T)) := T(v)$  is well-defined, and is an isomorphism of abelian groups. We need only to check that  $\hat{T}$  is linear, namely, that it respects scalar multiplication.



We have

$$\begin{aligned}\hat{T}(a \cdot (v + \text{Ker}(T))) &= \hat{T}((a \cdot v) + \text{Ker}(T)) \\ &= T(av) = a \cdot T(v) \\ &= a\hat{T}(v + \text{Ker}(T)),\end{aligned}$$

as desired. ■

## 2.3 The Space $\text{Hom}(V, W)$

### ↪ **Definition 2.6: Homomorphism Space**

For vector spaces  $V, W$  over  $\mathbb{F}$ , let  $\text{Hom}(V, W)$  (also denoted  $\ell(V, W)$ ) denote the set of all linear transformations from  $V$  to  $W$ . We can turn this into a vector space over  $\mathbb{F}$  as follows:

1. *Addition of linear transformations:* for  $T_0, T_1 \in \text{Hom}(V, W)$ , define

$$(T_0 + T_1) : V \rightarrow W, \quad v \mapsto T_0(v) + T_1(v).$$

$(T_0 + T_1)$  is clearly a linear transformation, as the linear combination of linear transformations  $T_0, T_1$ .

2. *Scalar multiplication of linear transformations:* for  $T \in \text{Hom}(V, W)$ ,  $a \in \mathbb{F}$ , define

$$(a \cdot T) : V \rightarrow W, \quad v \mapsto a \cdot T(v),$$

which is again clearly linear in its own right.

### ↪ **Proposition 2.5**

Endowed with the operations described above,  $\text{Hom}(V, W)$  is a vector space over  $\mathbb{F}$ .

Proof. Follows easily from the definitions. ■

↪ **Theorem 2.6: Basis for  $\text{Hom}(V, W)$**

For vector spaces  $V, W$  over  $\mathbb{F}$  and bases  $\beta, \gamma$  for  $V, W$  resp., the following set

$$\{T_{v,w} = v \in \beta, w \in \gamma\},$$

is a basis for  $\text{Hom}(V, W)$ , where for each  $v \in \beta$  and  $w \in \gamma$ ,  $T_{v,w} \in \text{Hom}(V, W)$  defined as the unique linear transformation such that

$$T_{v,w}(v') = \begin{cases} w & v' = v \\ 0_W & v' \neq v \end{cases} \iff v' \in \beta \setminus \{v\}.$$

Proof. Left as a (homework) exercise. ■

↪ **Corollary 2.3**

If  $V, W$  finite dimensional, then  $\dim(\text{Hom}(V, W)) = \dim(V) \cdot \dim(W)$ .

↪ **Proposition 2.6**

Let  $\beta = \{v_1, \dots, v_n\}, \gamma = \{w_1, \dots, w_m\}$  be bases for  $V, W$  resp. Then, by theorem 2.6,

$$\{T_{v_i, w_j} : i \in \{1, \dots, n\}, j \in \{1, \dots, m\}\}$$

is a basis for  $\text{Hom}(V, W)$ , and it has  $n \cdot m$  vectors by construction.

## 2.4 Matrix Representation of Linear Transformations, Finite Fields

Consider a linear transformation  $T : \mathbb{F}^n \rightarrow \mathbb{F}^m$  between finite fields. We know that  $T$  is uniquely determined by its value of basis vectors, so fix the standard bases

$$\beta = \{e_1^{(n)}, \dots, e_n^{(n)}\} = \{v_1, \dots, v_n\},$$

and note that  $T$  is determined by  $\{T(v_1), \dots, T(v_n)\} \subseteq \mathbb{F}^m$ .

**Remark 2.7.** We denote vectors in  $\mathbb{F}^n$  as column vectors, ie  $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{F}^n$ .

Each  $T(v_i)$  is a column vector in  $\mathbb{F}^m$ , and we can put these into a  $m \times n$  matrix, namely:<sup>7</sup>

<sup>7</sup>Where  $[T]$  denotes a matrix named “ $T$ ”.

$$[T] := \left( \begin{array}{c|ccc|c} & & & & \\ & & & & \\ T(v_1) & & \cdots & & T(v_n) \\ & & & & \\ & & & & \end{array} \right) = \underbrace{\left( \begin{array}{ccc} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{array} \right)}_n \Bigg\}_m$$

We call this the *matrix representation* of  $T$  in the standard bases. The operation of multiplying an  $m \times n$  matrix and a  $n \times 1$  vector is precisely defined so that

↪ **Proposition 2.7**

$T(v) = [T] \cdot v$  for all  $v \in \mathbb{F}^n$ .

Proof. Let  $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ , where  $v = x_1 v_1 + \cdots + x_n v_n$ . Then

$$T(v) = x_1 T(v_1) + \cdots + x_n T(v_n)$$

$$T(v_i) = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix}$$

so

$$T(v) = \begin{pmatrix} a_{11} \cdot x_1 + \cdots + a_{1n} \cdot x_n \\ \vdots \\ a_{m1} \cdot x_1 + \cdots + a_{mn} \cdot x_n \end{pmatrix} = [T] \cdot v$$

■

↪ **Definition 2.7**

For a given  $m \times n$  matrix  $A$  over  $\mathbb{F}$ , define  $L_A : \mathbb{F}^n \rightarrow \mathbb{F}^m$  by  $L_A(v) := A \cdot v$ , where  $v$  is viewed as an  $n \times 1$  column. It follows from definition that the  $L_A$  is linear.

In other words, every  $T \in \text{Hom}(\mathbb{F}^n, \mathbb{F}^m)$  is equal to  $L_A$  for some  $A$ .

↪ **Proposition 2.8**

The map

$$\begin{aligned}\text{Hom}(\mathbb{F}^n, \mathbb{F}^m) &\rightarrow M_{m \times n}(\mathbb{F}) \\ T &\mapsto [T]\end{aligned}$$

is an isomorphism of vector spaces, with inverse

$$\begin{aligned}M_{m \times n}(\mathbb{F}) &\rightarrow \text{Hom}(\mathbb{F}^n, \mathbb{F}^m) \\ A &\mapsto L_A.\end{aligned}$$

Proof. Linearity: Let  $\beta = \{v_1, \dots, v_n\}$  be the standard basis for  $\mathbb{F}^n$ . Fix  $T_1, T_2 \in \text{Hom}(\mathbb{F}^n, \mathbb{F}^m)$  and  $\alpha \in \mathbb{F}$ .

1.

$$\begin{aligned}[T_1 + T_2] &= \begin{pmatrix} \cdots & \begin{array}{c} | \\ (T_1 + T_2)(v_i) \\ | \end{array} & \cdots \end{pmatrix} = \begin{pmatrix} \cdots & \begin{array}{c} | \\ T_1(v_i) + T_2(v_i) \\ | \end{array} & \cdots \end{pmatrix} \\ &= \begin{pmatrix} \cdots & \begin{array}{c} | \\ T_1(v_i) \\ | \end{array} & \cdots \end{pmatrix} + \begin{pmatrix} \cdots & \begin{array}{c} | \\ T_2(v_i) \\ | \end{array} & \cdots \end{pmatrix} \\ &= [T_1] + [T_2]\end{aligned}$$

2. It remains to show that  $\alpha \cdot [T] = [\alpha \cdot T]$ ; the proof follows similarly to 1.

Inverse: We need to show that 1.  $A \mapsto L_A \mapsto [L_A]$  is the identity on  $M_{m \times n}(\mathbb{F})$ , and conversely, that 2.  $T \mapsto [T] \mapsto L_{[T]}$  is the identity on  $\text{Hom}(\mathbb{F}^n, \mathbb{F}^m)$ .

1. We need to show that  $[L_A] = A$ . The  $j$ th column of  $[L_A]$  is  $L_A(v_j) = A \cdot v_j = j$ th column of  $A =: A^{(j)}$ . Hence, the  $j$ th column of  $[L_A]$  is equal to the  $j$ th column of  $A$ , and thus they are equal.
2. We showed this in proposition 2.7.

■

↪ **Corollary 2.4**

$$\dim(\text{Hom}(\mathbb{F}^n, \mathbb{F}^m)) = \dim(M_{m \times n}(\mathbb{F})) = m \cdot n.$$

**Remark 2.8.** This was stated previously in proposition 2.6 by constructing an explicit basis. Indeed, this basis is precisely the image of the standard basis for  $M_{m \times n}(\mathbb{F})$  under the map  $A \mapsto L_A$ .

## 2.5 Matrix Representation of Linear Transformations, General Spaces

**Remark 2.9.** The previous section was concerned with representing transformations between finite fields  $\mathbb{F}^n, \mathbb{F}^m$ ; this section aims to make the same construction for any finite dimensional  $V, W$ .

### ⇨ Definition 2.8: Coordinate Vector

Let  $V$  be a finite dimensional space over  $\mathbb{F}$  and let  $\beta := \{v_1, \dots, v_n\}$  be a basis for  $V$ . Let  $v \in V$ , with (unique) representation  $v = a_1v_1 + \dots + a_nv_n$ . We denote

$$[v]_\beta := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{F}^n$$

the coordinate vector of  $v$  in base  $\beta$ .

**Remark 2.10.** Recall that  $V \cong \mathbb{F}^n$  where  $\dim(V) = n$ , by the unique linear transformation  $v_i \mapsto e_i$ , where  $\{e_1, \dots, e_n\}$  the standard basis for  $\mathbb{F}^n$ . We denote this transformation

$$I_\beta : V \rightarrow \mathbb{F}^n.$$

For an arbitrary  $v \in V$ ,  $I_\beta(v)$  maps  $v$  to its coordinate vector:

$$I_\beta(v) = I_\beta(a_1v_1 + \dots + a_nv_n) = a_1I_\beta(v_1) + \dots + a_nI_\beta(v_n) \quad (1)$$

$$= a_1e_1 + \dots + a_ne_n = [v]_\beta. \quad (2)$$

### ⇨ Proposition 2.9

The map

$$I_\beta : V \rightarrow \mathbb{F}^n, \quad v \mapsto [v]_\beta$$

is an isomorphism.

Suppose we are given a linear transformation  $T : V \rightarrow W$ , where  $V, W$  finite dimensional spaces over  $\mathbb{F}$ . Fix  $\beta := \{v_1, \dots, v_n\}$  and  $\gamma := \{w_1, \dots, w_m\}$  as bases for  $V, W$  resp. We can denote  $[T(v_i)]_\gamma$  as  $T(v_i)$  in base  $\gamma$  (in the field  $m$ ), and construct a matrix for  $T$ :<sup>8</sup>

$$[T]_\beta^\gamma := \begin{pmatrix} | & & | \\ [T(v_1)]_\gamma & \cdots & [T(v_n)]_\gamma \\ | & & | \end{pmatrix}$$

We call this the *matrix representation* of  $T$  from  $\beta$  to  $\gamma$ .

<sup>8</sup>Where we denote  $[T]_\beta^\gamma$  as the matrix representation of the transform  $T : V \rightarrow W$ , with basis  $\beta, \gamma$  for  $V, W$  respectively.

↪ **Theorem 2.7**

Let  $T : V \rightarrow W$ ,  $\beta, \gamma$  as above.

1. The following diagram commutes:

$$\begin{array}{ccc} \bullet V & \xrightarrow{T} & \bullet W \\ I_\beta \downarrow & & \downarrow I_\gamma \\ \bullet \mathbb{F}^n & \xrightarrow{L_{[T]_\beta^\gamma}} & \bullet \mathbb{F}^m \end{array}$$

Namely,  $I_\gamma \circ T = L_{[T]_\beta^\gamma} \circ I_\beta$ , or equivalently, given  $v \in V$ ,  $[T(v)]_\gamma = [T]_\beta^\gamma \cdot [v]_\beta$ .

2. The map  $\text{Hom}(V, W) \rightarrow M_{m \times n}(\mathbb{F})$ ,  $T \mapsto [T]_\beta^\gamma$  is a vector space isomorphism with inverse given the map  $M_{m \times n}(\mathbb{F}) \rightarrow \text{Hom}(V, W)$ ,  $A \mapsto I_\gamma^{-1} \circ L_A \circ I_\beta$

Proof. 2. is left as a (homework) exercise; it follows directly from 1.

Fix  $v \in V$ . We need to show that  $I_\gamma \circ T(v) = L_{[T]_\beta^\gamma} \circ I_\beta(v)$ . We have

$$I_\gamma \circ T(v) = [T(v)]_\gamma.$$

OTOH,

$$L_{[T]_\beta^\gamma} \circ I_\beta(v) = L_{[T]_\beta^\gamma}([v]_\beta) = [T]_\beta^\gamma \cdot [v]_\beta.$$

We need to show, then, that  $[T(v)]_\gamma = [T]_\beta^\gamma \cdot [v]_\beta$ . Let  $v = a_1 v_1 + \cdots + a_n v_n$ , so  $[v]_\beta = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ . Recall that  $[T]_\beta^\gamma =$

$$\begin{pmatrix} | & & | \\ [T(v_1)]_\gamma & \cdots & [T(v_n)]_\gamma \\ | & & | \end{pmatrix}. \text{ Thus, we have}$$

$$\begin{aligned} [T]_\beta^\gamma \cdot [v]_\beta &= a_1 [T(v_1)]_\gamma + \cdots + a_n [T(v_n)]_\gamma = [a_1 T(v_1) + \cdots + a_n T(v_n)]_\gamma \quad (\text{by linearity of } I_\gamma) \\ &= [T(a_1 v_1 + \cdots + a_n v_n)]_\gamma \quad (\text{by linearity of } T) \\ &= [T(v)]_\gamma, \end{aligned}$$

which is precisely what we wanted to show. ■

**Remark 2.11.** For  $A \in M_{m \times n}(\mathbb{F})$  and  $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{F}^n$ , we have

$$A \cdot x = x_1 \cdot A^{(1)} + x_2 \cdot A^{(2)} + \cdots + x_n \cdot A^{(n)},$$

where  $A^{(j)}$  is the  $j$ th column of  $A$ ; thus  $A \cdot x$  is a linear combination of  $A$ , with coefficients given by the vector  $x$ ; this

## 2.6 Composition of Linear Transformations, Matrix Multiplication

### ↪ **Proposition 2.10**

Composition is associative; given  $T : V \rightarrow W$ ,  $S : W \rightarrow U$ , and  $R : U \rightarrow X$ , then

$$(R \circ S) \circ T = R \circ (S \circ T).$$

Proof. Fix  $v \in V$ . Then

$$(R \circ S) \circ T(v) = (R \circ S)(T(v)) = R(S(T(v)))$$

OTOH:

$$R \circ (S \circ T)(v) = R((S \circ T)(v)) = R(S(T(v))).$$

■

Let  $A \in M_{m \times n}(\mathbb{F})$  and  $B \in M_{l \times m}(\mathbb{F})$ . Then,  $L_A : \mathbb{F}^n \rightarrow \mathbb{F}^m$  and  $L_B : \mathbb{F}^m \rightarrow \mathbb{F}^l$ , and have composition  $L_B \circ L_A : \mathbb{F}^n \rightarrow \mathbb{F}^l$ . We know that  $L_B \circ L_A$  is a linear transformation, and thus must be equal to  $L_C$  for some matrix  $C \in M_{l \times n}(\mathbb{F})$ . Indeed,  $C$  is the matrix representation of the transformation  $[L_B \circ L_A]$ , as proven previously.

Let  $\beta = \{e_1, \dots, e_n\}$  for  $\mathbb{F}^n$ , then

$$[L_B \circ L_A] = \begin{pmatrix} | & & | \\ L_B \circ L_A(e_1) & \cdots & L_B \circ L_A(e_n) \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ B \cdot (A \cdot e_1) & \cdots & B \cdot (A \cdot e_n) \\ | & & | \end{pmatrix}$$

### ↪ **Definition 2.9: Matrix Multiplication**

For matrices  $A \in M_{m \times n}(\mathbb{F})$  and  $B \in M_{l \times m}(\mathbb{F})$ , define their product  $B \cdot A$  to be the matrix

$$[L_B \circ L_A] = \begin{pmatrix} | & & | \\ B \cdot (A \cdot e_1) & \cdots & B \cdot (A \cdot e_n) \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ B \cdot A^{(1)} & \cdots & B \cdot A^{(n)} \\ | & & | \end{pmatrix} = (c_{ij})_{1 \leq i \leq l, 1 \leq j \leq n}$$

where  $A^{(j)}$  is the  $j$ th column of  $A$ ,  $c_{ij} := \begin{pmatrix} - & B_{(i)} & - \end{pmatrix} \cdot \begin{pmatrix} | \\ A^{(j)} \\ | \end{pmatrix}$ .

↪ **Proposition 2.11**

$[L_B \circ L_A] = B \cdot A$ , ie  $L_B \circ L_A = L_{B \cdot A}$ .

Proof. Follows from our definition. ■

↪ **Corollary 2.5**

Matrix multiplication is association;  $C \cdot (B \cdot A) = (C \cdot B) \cdot A$  for  $A \in M_{m \times n}(\mathbb{F})$ ,  $B \in M_{l \times m}(\mathbb{F})$ ,  $C \in M_{k \times l}(\mathbb{F})$ .

Proof.  $C \cdot (B \cdot A) = [L_C \circ (L_B \circ L_A)] = [(L_C \circ L_B) \circ L_A] = (C \cdot B) \cdot A$ . ■

**Remark 2.12.** This is proven by the linear transformation representation of matrices; try proving this directly from our definition.

↪ **Corollary 2.6**

Let  $V, W, U$  be finite-dimensional vector spaces over  $\mathbb{F}$ ,  $T : V \rightarrow W, S : W \rightarrow U$  be linear transformations and  $\alpha, \beta, \gamma$  be bases for  $V, W, U$  resp. Then,

$$[S \circ T]_{\alpha}^{\gamma} = [S]_{\beta}^{\gamma} \cdot [T]_{\alpha}^{\beta}.$$

Proof. Follows from the commutativity of the diagrams:

$$\begin{array}{ccccc} V & \xrightarrow{T} & W & \xrightarrow{S} & U \\ \wr \downarrow & & \wr \downarrow & & \wr \downarrow \\ \mathbb{F}^n & \xrightarrow{[T]_{\alpha}^{\beta}} & \mathbb{F}^m & \xrightarrow{[S]_{\beta}^{\gamma}} & \mathbb{F}^l \end{array} \iff \begin{array}{ccc} V & \xrightarrow{T \circ S} & U \\ \wr \downarrow & & \wr \downarrow \\ \mathbb{F}^n & \xrightarrow{[S \circ T]_{\alpha}^{\gamma}} & \mathbb{F}^l \end{array}$$

In “words”, for  $v \in V$ ,

$$[S \circ T]_{\alpha}^{\gamma} \cdot [v]_{\alpha} = [(S \circ T)(v)]_{\alpha}^{\gamma} = [S(T(v))]_{\alpha}^{\gamma} = [S]_{\beta}^{\gamma} \cdot [T(v)]_{\beta} = [S]_{\beta}^{\gamma} \cdot [T]_{\alpha}^{\beta} \cdot [v]_{\alpha},$$

ie we have shown that  $L_{[S \circ T]_{\alpha}^{\gamma}} = L_{[S]_{\beta}^{\gamma} \cdot [T]_{\alpha}^{\beta}}$ . Because  $A \mapsto L_A$  is an isomorphism, it follows that  $[S \circ T]_{\alpha}^{\gamma} = [S]_{\beta}^{\gamma} \cdot [T]_{\alpha}^{\beta}$ . ■

↪ Lecture 13; Last Updated: Sat Feb 3 22:20:36 EST 2024

## 2.7 Inverses of Transformations and Matrices

**Remark 2.13.** Recall that, given a function  $f : X \rightarrow Y$ , a function  $g : Y \rightarrow X$  is called

1. a left inverse of  $f$  if  $g \circ f = \text{Id}_X$ ;
2. a right inverse of  $f$  if  $f \circ g = \text{Id}_Y$ ;
3. a (two-sided) inverse of  $f$  if  $g$  both a left and right inverse of  $f$ .

If an inverse exists, it is unique; let  $g_0, g_1$  be inverse of  $f$ , then,  $g_0 = g_0 \circ (f \circ g_1) = (g_0 \circ f) \circ g_1 = g_1$ .



↪ **Proposition 2.12**

Let  $f : X \rightarrow Y$ . Then,

1.  $f$  has a left-inverse  $\iff f$  injective;
2.  $f$  has a right-inverse  $\iff f$  surjective;
3.  $f$  has an inverse  $\iff f$  bijective.

Proof. ((a),  $\implies$ ) Suppose  $g : Y \rightarrow X$  is a left-inverse of  $f$  and  $f(x_1) = f(x_2)$ . Then,  $g \circ f(x_1) = g \circ f(x_2) \implies x_1 = x_2$  and so  $f$  injective.

((b),  $\implies$ ) Suppose  $g : Y \rightarrow X$  is a right-inverse of  $f$  and let  $y \in Y$ . Then,  $f(g(y)) = y \implies y \in f(X)$ .

The remainder of the cases and directions are left as an exercise. ■

**Remark 2.14.** Proof of (b),  $\Leftarrow$  uses Axiom of Choice.

⊛ **Example 2.2**

1. The differentiation transform  $\delta : \mathbb{F}[t]_{n+1} \rightarrow \mathbb{F}[t]_n, p(t) \mapsto p'(t)$  has a right inverse, the integration transform,  $\iota : \mathbb{F}[t]_n \rightarrow \mathbb{F}[t]_{n+1}, p(t) \mapsto$  antiderivative of  $p(t)$ ; conversely,  $\iota$  has left inverse  $\delta$ ; they do not admit inverses.
2. Let  $f : \mathbb{F}[[t]] \rightarrow \mathbb{F}[[t]]$  be the left-shift map, where  $\sum_{n=0}^{\infty} a_n t^n \mapsto \sum_{n=1}^{\infty} a_n t^{n-1}$ . Then,  $g : \mathbb{F}[[t]] \rightarrow \mathbb{F}[[t]]$  with  $\sum_{n=0}^{\infty} a_n t^n \mapsto \sum_{n=0}^{\infty} a_n t^{n+1}$ , the right-shift map, is a right inverse of  $f$ , but  $f$  has no left inverse (it is not injective).

**Remark 2.15.** The existence of only one-sided inverses existing happens only when in infinite-dimensional vectors spaces, or when the dimension of the domain is not the same as the dimension of the codomain.

↪ **Corollary 2.7: Of Rank-Nullity Theorem**

Let  $T : V \rightarrow W$  s.t.  $\dim(V) = \dim(W) < \infty$ . TFAE:

1.  $T$  has a left-inverse;
2.  $T$  has a right-inverse;
3.  $T$  is invertible (has an inverse).

Proof. We have already that  $T$  injective  $\iff T$  surjective  $\iff T$  bijective. ■

↪ **Definition 2.10: Matrix Inverse**

We call a  $n \times n$  matrix  $B$  over  $\mathbb{F}$  the *inverse* of an  $n \times n$  matrix  $A$  over  $\mathbb{F}$  if  $A \cdot B = B \cdot A = I_n$ . We denote  $B = A^{-1}$ .

↪ **Proposition 2.13**

Let  $A \in M_n(\mathbb{F})$ . Then,

1.  $L_A$  is invertible  $\iff A$  is invertible, in which case  $L_A^{-1} = L_{A^{-1}}$ ;
2.  $A$  is invertible  $\iff$  it has a left-inverse, ie  $B \cdot A = I_n \iff$  it has a right-inverse, ie  $A \cdot B = I_n$ .

Proof. 1.  $L_A$  invertible  $\iff \exists T : \mathbb{F}^n \rightarrow \mathbb{F}^n$ -linear s.t.  $L_A \circ T = T \circ L_A = I_{\mathbb{F}^n} \iff \exists$  a matrix  $B \in M_n(\mathbb{F})$  such that  $L_A \circ L_B = L_B \circ L_A = I_{\mathbb{F}^n} \iff$  there is a matrix  $B \in M_n(\mathbb{F})$  s.t.  $L_{AB} = L_{BA} = I_{\mathbb{F}^n} \iff$  there is a  $B \in M_n(\mathbb{F})$  s.t.  $A \cdot B = B \cdot A = I_n$ .

2. Follows directly from corollary 2.7 and part 1.



### 2.7.1 Invariant Subspaces and Nilpotent Transformations

↪ **Definition 2.11:  $T$ -Invariant**

Let  $T : V \rightarrow V$  be a linear transformation.<sup>9</sup>We call a subspace  $W \subseteq V$   *$T$ -invariant* if  $T(W) \subseteq W$ .

⊛ **Example 2.3: Examples of Invariant Subspaces**

1. For any  $T : V \rightarrow V$ ,  $\text{Im}(T)$  is  $T$ -invariant.
2. For any  $T : V \rightarrow V$ ,  $\text{Ker}(T)$  is  $T$ -invariant, since  $T(v) = 0_V \in \text{Ker}(T) \forall v \in \text{Ker}(T)$ . Moreover, for any  $n \in \mathbb{N}$ , the space  $\text{Ker}(T^n)$  is  $T$ -invariant.<sup>10</sup>

↪ Lecture 14; Last Updated: Wed Feb 7 13:36:32 EST 2024

↪ **Proposition 2.14**

For a linear operator  $T : V \rightarrow V$ , the following hold:

1.  $V \supseteq \text{Im}(T) \supseteq \text{Im}(T^2) \supseteq \dots \supseteq \text{Im}(T^n) \supseteq \dots$ . Moreover,  $\text{Im}(T^n)$  is  $T$ -invariant for any  $n \in \mathbb{N}$ .
2.  $\{0_V\} \subseteq \text{Ker}(T) \subseteq \text{Ker}(T^2) \subseteq \dots \subseteq \text{Ker}(T^n) \subseteq \dots$ . Moreover,  $\text{Ker}(T^n)$  is  $T$ -invariant for any  $n \in \mathbb{N}$ .

Proof. 1. If  $x \in \text{Im}(T^{n+1})$ , then  $x = T^{n+1}(y) = T^n(T(y)) \in \text{Im}(T^n)$  for some  $y \in V$ , hence  $\text{Im}(T^{n+1}) \subseteq \text{Im}(T^n)$ .  
If  $x \in \text{Im}(T^n)$ , then  $x = T^n(y)$  so  $T(x) = T(T^n(y)) = T^n(T(y)) \in \text{Im}(T^n)$ , so  $T(\text{Im}(T^n)) \subseteq \text{Im}(T^n)$ .

<sup>9</sup>Because the domain and codomain are the same, we often call  $T$  a “linear operator”.

<sup>10</sup> $T^n := T \circ T \circ \dots \circ T$ ,  $n$  times;  $T^0 := I_V$ .

2. If  $x \in \text{Ker}(T^n)$ , then  $T^{n+1}(x) = T(T^n(x)) = T(0_V) = 0_V$  hence  $x \in \text{Ker}(T^{n+1})$  so  $\text{Ker}(T^n) \subseteq \text{Ker}(T^{n+1})$ .  
 Moreover,  $T(x) \in \text{Ker}(T^n)$  since  $T(x) \in \text{Ker}(T^{n-1}) \subseteq \text{Ker}(T^n)$ , since  $T^{n-1}(T(x)) = T^n(x) = 0_V$  so  $T(\text{Ker}(T^n)) \subseteq \text{Ker}(T^n)$ . ■

⊛ **Example 2.4: More Examples of Invariant Subspaces**

Let  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  by  $T(x, y, z) := (2x + y, 3x - y, 7z)$ . Then, the  $x - y$  plane is  $T$ -invariant, as is the  $z$  axis. Hence, we can decompose  $\mathbb{R}^3$  into two  $T$ -invariant subspaces, namely the  $x - y$  plane  $\oplus$   $z$ -axis.

↪ **Definition 2.12: Nilpotent**

In a ring  $R$ , an element  $r \in R$  is called *nilpotent* if  $r^n = 0$  for some  $n \in \mathbb{N}^+$ .

More specifically, a linear transformation  $T : V \rightarrow V$  is called nilpotent if  $T^n = 0$  for some  $n \in \mathbb{N}^+$ .

For a matrix  $A \in M_n(\mathbb{F})$ ,  $A$  is called nilpotent if  $A^n = 0_n$  for some  $n \in \mathbb{N}^+$ .

⊛ **Example 2.5: Examples of Nilpotent Transformations**

1. Let  $V$ ,  $n$ -dimensional vector space over  $\mathbb{F}$  with basis  $\beta := \{v_1, \dots, v_n\}$ . Let  $T : V \rightarrow V$  be the unique linear transformation that “shifts”  $\beta$ : ie,  $T(v_1) := 0_V$ ,  $T(v_2) := v_1, \dots, T(v_n) = v_{n-1}$ .
2. The differentiation operation,  $\delta : \mathbb{F}[t]_n \rightarrow \mathbb{F}[t]_n$  is nilpotent, since  $\delta^{n+1} = 0$  for any polynomial.
3. For any matrix  $A \in M_n(\mathbb{F})$ ,  $A$  is nilpotent iff  $L_A : \mathbb{F}^n \rightarrow \mathbb{F}^n$  is nilpotent.

Proof.  $L_{A^k} = L_A^k \implies A^k = 0 \iff L_{A^k} = 0 \iff L_A^k = 0$  ■

4.  $n \times n$  matrices that are strictly upper triangular are nilpotent. For instance, for  $3 \times 3$ , we need to show

$$\begin{pmatrix} 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix}^3 = 0 \iff \begin{pmatrix} 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix}^2 \cdot \begin{pmatrix} * \\ * \\ * \end{pmatrix} = 0$$

We have:

$$\begin{aligned} \begin{pmatrix} 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix}^2 \begin{pmatrix} 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} * \\ * \\ * \end{pmatrix} &= \begin{pmatrix} 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix}^2 \begin{pmatrix} * \\ * \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} * \\ * \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} * \\ 0 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \end{aligned}$$

↪ **Proposition 2.15**

If  $V$  is  $n$ -dimension  $T : V \rightarrow V$  is linear nilpotent, then  $T^n = 0$ .

Proof. Left as a (homework) exercise. ■

↪ **Definition 2.13: Domain Restriction**

For a function  $f : X \rightarrow Y$  and  $A \subseteq X$ , we define the *restriction* of  $f$  to  $A$  as the function  $f|_A : A \rightarrow Y$  given by  $a \mapsto f(a)$ .

↪ **Theorem 2.8: Fitting's Theorem**

For finite dimensional vector space  $V$  over  $\mathbb{F}$  and a linear transformation  $T : V \rightarrow V$ , there is a decomposition

$$V = U \oplus W$$

as a direct sum of  $T$ -invariant subspaces  $U, W$  such that  $T|_U : U \rightarrow U$  is nilpotent and  $T|_W : W \rightarrow W$  is an isomorphism.

Proof. Recall that  $\text{Im}(T) \supseteq \cdots \supseteq \text{Im}(T^n)$  and  $\text{Ker}(T) \subseteq \cdots \subseteq \text{Ker}(T^n)$ . Both of these become constant eventually, ie the inequalities become strict equalities, hence  $\exists N \in \mathbb{N}^+$  such that  $\forall k \in \mathbb{N}$ ,  $\text{Im}(T^{N+k}) = \text{Im}(T^N)$  and  $\text{Ker}(T^{N+k}) = \text{Ker}(T^N)$ .

Let  $U := \text{Ker}(T^N)$  and  $W := \text{Im}(T^N)$ . These are clearly  $T$ -invariant.

$T^N(\text{Ker}(T^N)) = \{0_V\}$ , and  $T(\text{Im}(T^N)) = \text{Im}(T^{N+1}) = \text{Im}(T^N) = W$  and thus  $T|_W : W \rightarrow W$  is surjective and hence  $T|_W$  must be injective and thus an isomorphism.

It remains to show that  $V = U \oplus W$ .

Claim:  $U \cap W = \{0_V\}$ ; If  $v \in U \cap W$ ,  $T^N(v) = 0_V$  but  $T|_W$  an isomorphism so  $T^N(v) = 0 \iff v = 0_V$ , hence  $U \cap W = \{0_V\}$ .

Thus, we have  $\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W) = \dim(U) + \dim(W) = \dim(V)$ ; moreover, it must be that  $U + W = V$ . ■