

---

# MATH 251 CLASS NOTES

McGILL UNIVERSITY

NICHOLAS HAYEK

*Based on lectures by Prof. Anush Tserunyan*

---

## CONTENTS

|           |  |           |
|-----------|--|-----------|
| <b>I</b>  | <b>Vector Spaces</b>                             | <b>3</b>  |
|           | Axioms   | 3         |
|           | Further Constructions                            | 4         |
|           | Subspaces . . . . .                              | 4         |
|           | Linear Combinations                              | 5         |
|           | Linear (In)dependence                            | 7         |
|           | Bases  | 8         |
|           | Zorn's Lemma . . . . .                           | 10        |
|           | Steinitz Substitution Lemma . . . . .            | 11        |
| <br>      |  |           |
| <b>II</b> | <b>Linear Transformations</b>                    | <b>13</b> |
|           | Axioms and Initial Properties                    | 13        |
|           | Isomorphisms                                     | 15        |
|           | The Space $\text{Hom}(V, W)$                     | 18        |
|           | Matrices   | 19        |
|           | Matrix Representations in Generality . . . . .   | 21        |
|           | Compositions and Matrix Multiplication . . . . . | 23        |
|           | Invariants and Nilpotent Transformations         | 23        |
|           | Preliminaries . . . . .                          | 23        |
|           | T-Invariants . . . . .                           | 24        |

# I Vector Spaces

## AXIOMS

In previous calculus courses, we've considered the set of tuples  $\langle a_1, \dots, a_n \rangle$ , where, in particular,  $a_i \in \mathbb{R}$ . These are examples of *vector spaces*, the construction which we will primarily study in this course.

Define a *vector space*  $V$  over the field  $\mathbb{F}$  to be an abelian group under the operation  $+$  and an identity element  $\mathbf{0}_V$ , which one calls the *zero vector*. Members of  $V$  are called *vectors*. Finally,  $V$  is equipped with scalar multiplication by members of  $\mathbb{F}$ , and satisfy the following axioms:

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

Recall also the properties of abelian groups from MATH 235, which apply to  $V$ :

$$u(vw) = (uv)w \quad v + \mathbf{0}_V = v \quad \exists(-v) \text{ s.t. } v + (-v) = \mathbf{0}_V \quad uv = vu$$

for all  $u, v, w \in V$ .

Some formal consequences of the vector space axioms:

PROPOSITION 1.1

$$\mathbf{0}_{\mathbb{F}}v = \mathbf{0}_V \text{ for all } v \in V \quad -\mathbf{1}_{\mathbb{F}}v = -v \quad \alpha\mathbf{0}_V = \mathbf{0}_V$$

$$(1): \quad \mathbf{0}_{\mathbb{F}}v = (\mathbf{0}_{\mathbb{F}} + \mathbf{0}_{\mathbb{F}})v = \mathbf{0}_{\mathbb{F}}v + \mathbf{0}_{\mathbb{F}}v \implies \mathbf{0}_V = \mathbf{0}_{\mathbb{F}}v$$

$$(2): \quad -\mathbf{1}_{\mathbb{F}}v + v = (-\mathbf{1}_{\mathbb{F}} + \mathbf{1}_{\mathbb{F}})v = \mathbf{0}_{\mathbb{F}}v = \mathbf{0}_V$$

$$(3): \quad \alpha\mathbf{0}_V = \alpha(\mathbf{0}_V + \mathbf{0}_V) \implies \alpha\mathbf{0}_V = \mathbf{0}_V \quad \square$$

PROOFS.

**Examples:** Most of the pedagogical examples of vector spaces we'll see do not bear much resemblance to the  $\mathbb{R}^n$ ,  $\langle x, y, z \rangle$ -like form we are familiar with:

1. The set of real, continuous functions, denoted  $C[\mathbb{R}] := \{f : \mathbb{R} \rightarrow \mathbb{R}\}$ , is a vector space over  $\mathbb{R}$ .
2.  $\mathbb{F}[t]$ , the set of polynomials with coefficients in  $\mathbb{F}$ , where addition and scalar multiplication are defined as usual, is a vector space over  $\mathbb{F}$ .

## FURTHER CONSTRUCTIONS

Define a *product*, sometimes called the *direct sum*, of two vector spaces  $U, V$  over the same field  $\mathbb{F}$  to be the Cartesian product  $U \times V$  equipped with the following:

$$(u_1, v_1) + (u_2, v_2) = (u_1 + u_2, v_1 + v_2) \quad \text{and} \quad \lambda(u, v) = (\lambda u, \lambda v)$$

$\forall u_1, u_2 \in U, v_1, v_2 \in V, \lambda \in \mathbb{F}$ . One notates this as  $U \oplus V$ . This is itself a vector space. Note that the coordinate-wise addition and scalar multiplication are defined as in the original vector spaces.

A good exercise to prove.

For example, consider  $\mathbb{F}^2$  over the field  $\mathbb{F}$ . One can conceptualize  $\mathbb{F}$  as a vector space over  $\mathbb{F}$ , and thus the direct product of  $\mathbb{F}$  with itself is a vector space.

## Subspaces

We have constructed from a vector space one larger than it. Here is one smaller: define a *subspace* to be a set  $W \subseteq V$  satisfying the following conditions

$$0_V \in W \quad u + v \in W \quad \forall u, v \in W \quad \alpha u \in W \quad \forall u \in W, \alpha \in \mathbb{F}$$

If  $W$  were non-empty, then choose  $u \in W$ . Then  $0_{\mathbb{F}}u \in W$ , so  $0_V \in W$  as required.

There are a few equivalent characterizations of subspaces:  $W \subseteq V$  is a vector space; or,  $W \subseteq V$  is non-empty and satisfies the latter two conditions from above.

**Examples:** Consider  $\mathbb{F}^n$  over the field  $\mathbb{F}$ . This is a vector space. The following are subspaces of  $\mathbb{F}^n$  :

1.  $\{(0, x_2, \dots, x_n) \in \mathbb{F}^n : x_i \in \mathbb{F}\}$ .
2.  $W = \{(x_1, \dots, x_n) \in \mathbb{F}^n : x_1 + 2x_2 = 0\}$ . One can choose  $x_3, \dots, x_n$  all 0, and since  $x_1 = x_2 = 0$  satisfy  $x_1 + 2x_2 = 0$ , one sees that  $0_V \in W$ . If  $x_1 + 2x_2 = 0$ , then  $\lambda x_1 + 2\lambda x_2 = 0$  as well, so  $W$  is closed under scalar multiplication. Lastly, if  $x_1 + 2x_2 = 0$  and  $x'_1 + 2x'_2 = 0$ , then  $(x_1 + x'_1) + 2(x_2 + x'_2) = 0$ , so  $W$  is closed under addition.
3. *Generally*, though it is not a fact we can prove now,  $W \subseteq \mathbb{F}^n :=$

$$\left\{ (x_1, \dots, x_n) \in \mathbb{F}^n \text{ s.t. } \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = 0 \end{cases} \right\}$$

i.e. a subset of  $\mathbb{F}^n$  where a system of at least one linear equation is homogeneous. As a counter-example to this construction, see that  $\{(x_1, \dots, x_n) \in \mathbb{F}^n : x_1 + x_2 = 1\}$  is not a subspace of  $\mathbb{F}^n$  (it violates all 3 conditions).

4. Let  $\mathbb{F}[t]_n$  denote the space of polynomials whose degree is at most  $n \in \mathbb{N}$ . Then  $\mathbb{F}[t]_n \subseteq \mathbb{F}[t]$  is a subspace. However, the set of polynomials whose degree is *exactly*  $n$ , for some positive  $n \in \mathbb{N}$ , is *not* a subspace. The following are further subspaces of  $\mathbb{F}[t]_n$ , where  $p''(t)$  is defined as usual (notice the similarity to 3.).

- (a)  $\{p(t) \in \mathbb{F}[t]_n : p(1) = 0\}$  or even  $\{p(t) \in \mathbb{F}[t]_n : p(\alpha) = 0 \text{ with } \alpha \in \mathbb{F}\}$
- (b)  $\{p(t) \in \mathbb{F}[t]_n : p''(t) + 2p'(t) - p(t) = 0\}$

5. For  $C[\mathbb{R}]$ , which, as noted above, is a vector space, the following are subspaces:

- (a)  $\{f \in C[\mathbb{R}] : f(\pi) + 7f(\sqrt{2}) = 0\}$
- (b)  $\{f \in C[\mathbb{R}] \text{ differentiable everywhere}\}$
- (c)  $\left\{f \in C[\mathbb{R}] : \int_0^1 f dx = 0\right\}$ . The proof of this follows from linearity of the integral (MATH 255). In truth, the bounds for the integral can be arbitrary, though see this integral cannot be set arbitrarily.

---

If  $W_1, W_2$  are subspaces of some common VS over the field  $\mathbb{F}$ , then

PROPOSITION 1.2

$$W_1 + W_2 := \{w_1 + w_2 : w_1 \in W_1, w_2 \in W_2\} \quad \text{and} \quad W_1 \cap W_2$$

are both subspaces. The proofs for these are left to the reader.

### LINEAR COMBINATIONS

Define a *linear combination* of vectors  $v_1, \dots, v_n \in V$ , where  $V$  is a vector space over  $\mathbb{F}$ , to be  $\sum_{i=1}^n a_i v_i$ , where  $a_i \in \mathbb{F}$ . So long as one  $a_i$  is non-zero, one calls this a *non-trivial* LC. Otherwise (i.e. all  $a_i = 0$ ), we have a *trivial* LC.

When we deal with a possibly infinite set of vectors,  $S \subseteq V$ , we will only take *finite* linear combinations, for a subset  $\{v_1, \dots, v_n\} \subseteq S$ . Never will we compute infinite sums in this course.

Define the *span* of  $S \subseteq V$  to be the set of all possible linear combinations of  $S$ ,  $\{a_1 v_1 + \dots + a_n v_n : a_i \in \mathbb{F}, v_i \in S\}$ . By convention, we say that  $\text{Span}(\emptyset) = \{0\}_V$ .

---

**Example:** Let  $S := \{(1, 0, -1), (0, 1, -1), (1, 1, -2)\} \subseteq \mathbb{R}^3$ . Then  $\mathbf{0}_{\mathbb{R}^3} = (0, 0, 0) = 0(1, 0, -1) + 0(0, 1, -1) + 0(1, 1, -2)$  is a trivial linear combination. However, we can get to 0 non-trivially:  $(1, 0, -1) + (0, 1, -1) - (1, 1, -2) = 0$ .

What about  $\text{Span}(S)$ ? This is the set  $\{a(1, 0, -1) + b(0, 1, -1) + c(1, 1, -2)\} = \{(a + c, b + c, -a - b - 2c)\}$ . Clearly this is a subset of  $\{(a, b, c) : a + b + c = 0\}$ , since, indeed,

$a + c + b + c - a - b - 2c = 0$ . The converse is also true: suppose  $(x, y, z)$  is such that  $x + y + z = 0$ . Then  $z = -x - y$ , and one writes  $(x, y, -x - y) = x(1, 0, -1) + y(0, 1, -1)$ . It follows that  $\text{Span}(S) = \{(x, y, z) : x + y + z = 0\}$ .

PROPOSITION 1.3

Let  $V$  be a VS over a field  $\mathbb{F}$ , and  $S$  be some subspace of it. Then  $\text{Span}(S)$  is a subspace of  $V$  containing  $S$ , and furthermore is the smallest such subspace containing  $S$ .

PROOF.

Adding and scalar multiplying a linear combination of vectors produces a further linear combination, so  $\text{Span}(S)$  is closed under these operations. Furthermore,  $0_V \in \text{Span}(S)$  by taking a trivial combination of vectors  $\implies \text{Span}(S)$  is a subspace.

If  $U \supseteq S$  is a subspace, then  $U$  is closed under addition and scalar multiplication, so it contains all linear combinations of  $S$ , i.e.  $U \supseteq \text{Span}(S)$   $\square$

PROPOSITION 1.4

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

PROOF.

( $\implies$ ) If  $v \in \text{Span}(S)$ , then  $v$  is some linear combination of vectors in  $S$ , so  $v = a_1 v_1 + \dots + a_n v_n$ . Let  $u \in \text{Span}(S \cup \{v\})$ . Then  $u = a'_1 v'_1 + \dots + a'_m v'_m + av$ , where  $a$  may be 0, and  $v'_i \in S$ . One rewrites  $u = a'_1 v'_1 + \dots + a'_m v'_m + a(a_1 v_1 + \dots + a_n v_n)$  from above. Thus,  $\text{Span}(S \cup \{v\}) \subseteq \text{Span}(S)$ . Trivially,  $\text{Span}(S) \subseteq \text{Span}(S \cup \{v\})$ , so  $\text{Span}(S) = \text{Span}(S \cup \{v\})$ .

( $\impliedby$ ) Assume  $\text{Span}(S) = \text{Span}(S \cup \{v\})$ . Clearly,  $v \in \text{Span}(S \cup \{v\})$ , so  $v \in \text{Span}(S)$  as well.  $\square$

For a VS over a field  $\mathbb{F}$ , call  $S \subseteq V$  a *spanning set* of  $V$  if  $\text{Span}(S) = V$ . One calls a spanning set *minimal* if no proper subset of  $S$  is spanning, i.e.  $\text{Span}(S \setminus v) \neq V$  for all  $v \in S$ .

**Example:** For  $S := \{(1, 0, -1), (0, 1, -1), (1, 1, -2)\}$ , we have from Proposition 1.4 that  $\text{Span}(S) = \text{Span}(\{(1, 0, -1), (0, 1, -1)\})$ , as  $(1, 1, -2) \in \text{Span}(\{(1, 0, -1), (0, 1, -1)\})$ .

Thus, it follows that  $S$  is not a minimal spanning set over itself.

For the VS  $\mathbb{F}^n$  over  $\mathbb{F}$ , define the *standard spanning set*:

$$\text{St}_n := \{(\underbrace{1, 0, 0, \dots, 0}_{n-1 \text{ times}}), (0, 1, 0, \dots, 0), \dots, (0, \dots, 0, 1)\}$$

This is indeed spanning for  $\mathbb{F}^n$ , and is minimal.

### LINEAR (IN)DEPENDENCE

Let  $V$  be a vs and  $S \subseteq V$  a subspace.  $S$  is called *linearly dependent* if there exists a non-trivial linear combination equal to  $\mathbb{0}_V$ . Otherwise  $S$  is called *linearly independent*.

#### Examples:

1. The empty set, by vacuous implication, is linearly independent.
2. For  $v \in V$ ,  $v$  is linearly dependent  $\iff v = \mathbb{0}_V$
3.  $S := \{(1, 0, -1), (0, 1, -1), (1, 1, -2)\}$  is linearly dependent
4.  $S \subseteq \mathbb{F}^3 = \{(1, 0, -1), (0, 1, -1), (0, 0, 1)\}$  is linearly dependent. We argue by contradiction: let  $(0, 0, 0) = a(1, 0, -1) + b(0, 1, -1) + c(0, 0, 1) = (a, b, c - a - b)$ . Then  $a = b = 0$  by necessity, and it follows that  $c - a - b = c = 0$ . Thus, only a trivial linear combination equals the zero vector.
5.  $\text{St}_n \subseteq \mathbb{F}^n$  is linearly independent

Let  $V$  be a vs over  $\mathbb{F}$ ,  $S \subseteq V$  (possibly infinite). Then:

PROPOSITION 1.5

- (a)  $S$  is linearly dependent  $\iff$  there exists a finite  $S_0 \subseteq S$  which is linearly dependent
- (b)  $S$  is linearly independent  $\iff$  all finite  $S_0 \subseteq S$  are linearly independent

Note that (b) is simply the negation of (a), so only (a) requires a proof.

PROOF.

( $\implies$ ) Suppose  $S$  is linearly dependent. Then  $a_1 v_1 + \dots + a_n v_n = \mathbb{0}_V$ , where, WLOG, we assume that  $a_i \neq \mathbb{0}_{\mathbb{F}}$ . The set  $\{v_1, \dots, v_n\} \subseteq S$  is clearly linearly dependent.

( $\impliedby$ ) If  $S_0 \subseteq S$  is linearly dependent, then clearly  $S$  is too □

For  $S \subseteq V$  over  $\mathbb{F}$ , we have

PROPOSITION 1.6

- (a)  $S$  is linearly dependent  $\iff$  there exists  $v \in S$  with  $v \in \text{Span}\{S \setminus v\}$
- (b)  $S$  is linearly independent  $\iff$  for all  $v \in S$ ,  $v \notin \text{Span}\{S \setminus v\}$

Once again, only (a) requires proof.

PROOF.

( $\implies$ ) Let  $S$  be linearly dependent. Then  $a_1 v_1 + \dots + a_n v_n = \mathbb{0}_V$ , and WLOG we assume all  $a_i$  are non-zero. Since  $\mathbb{F}$  is a field, we may write  $v_1 = -a_1^{-1} a_2 v_2 - \dots - a_1^{-1} a_n v_n$ . Thus,  $v_1 \in \text{Span}\{S \setminus v_1\}$ , and we are done.

( $\Leftarrow$ ) Suppose  $v \in S$  is such that  $v \in \text{Span}(S \setminus v)$ . Then  $v = a_1 v_1 + \dots + a_n v_n$ , where  $v_i \in S \setminus v$ . It follows that  $0_V = a_1 v_1 + \dots + a_n v_n - v$ . As  $-1 \neq 0$ , this is non-trivial, and we are done.  $\square$

**COROLLARY**

Clearly,  $\text{Span}(S) = \text{Span}(S)$ . However,  $v \in S \implies v \notin \text{Span}(S \setminus v)$ . We know  $v \in \text{Span}(S)$ , so  $\text{Span}(S) \neq \text{Span}(S \setminus v)$ .

$S \subseteq V$  is linearly independent  $\iff S$  is a minimal spanning set for  $\text{Span}(S)$

For a vector space  $V$  over  $\mathbb{F}$ ,  $S \subseteq V$  is called *maximally independent* if  $S$  is linearly independent AND there does not exist  $v \in V \setminus S$  s.t.  $S \cup \{v\}$  is linearly independent. In other words,  $S$  is independent, and adding *any* new vectors will break this independence.

**PROPOSITION 1.7**

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

**PROOF.**

Let  $S$  be maximally independent. Then for any  $v \in V \setminus S$ , the set  $S \cup \{v\}$  is linearly dependent, i.e.  $av + a_1 v_1 + \dots + a_n v_n = 0_V$  for all non-zero  $a_i$ . In particular,  $a \neq 0$ , or else we would yield a non-trivial linear combination for only vectors in  $S$ , which violates our independence condition.

Thus, write  $v = -a^{-1}a_1 v_1 - \dots - a^{-1}a_n v_n$ , and conclude that  $v \in \text{Span}(S)$ . Then  $V \subseteq \text{Span}(S)$ . Clearly,  $\text{Span}(S) \subseteq V$ , so we conclude that  $\text{Span}(S) = V$ .  $\square$

## BASES

### 1.1 Characterization of a Basis

Let  $V$  be a VS over  $\mathbb{F}$  and  $S \subseteq V$ . The following are then equivalent:

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

**PROOFS.**

(1)  $\implies$  (2) Let  $S \subseteq V$  be a minimal spanning set for  $V$ . Then, especially,  $S$  is a minimal spanning set for  $\text{Span}(S)$ , and by the corollary above,  $S$  is linearly independent.

(3)  $\implies$  (1) Let  $S \subseteq V$  be maximally independent. By proposition 1.7,  $S$  is spanning for  $V$ . By the corollary,  $S$  is also minimally spanning for  $\text{Span}(S)$ . Combining, we see that  $S$  is minimally spanning for  $V$ .

(2)  $\implies$  (4) Let  $S \subseteq V$  be linearly independent and spanning for  $V$ . Then, clearly,  $l \in V \in \text{Span}(S)$  means that it can be written as a linear combination



of vectors in  $S$ . We need this combination to be unique: let  $a_1 v_1 + \dots + a_n v_n = l$  and  $b_1 v_1 + \dots + b_n v_n = l$ , where  $S = \{v_i\}_{1 \leq i \leq n}$ . One uses the same vectors, noting that some coefficients may be 0, as needed.

$a_1 v_1 + \dots + a_n v_n = b_1 v_1 + \dots + b_n v_n \implies a_1 v_1 + \dots + a_n v_n - b_1 v_1 - \dots - b_n v_n = 0$ . We can thus combine  $a_i - b_i = c_i$ , and write  $c_1 v_1 + \dots + c_n v_n = 0$ . Since  $S$  is linearly independent, we require that all  $c_i = 0$ , i.e.  $a_i = b_i \forall i$ .

(4)  $\implies$  (2) This result is immediate, as  $V \subseteq \text{Span}(S)$ ,  $\text{Span}(S) \subseteq V \implies \text{Span}(S) = V$ . Since all vectors in  $v$  have a *unique* representation, consider  $v = 0_V$ . A trivial combination produces the zero vector, and by uniqueness this must be the *only* such combination, and we conclude that  $S$  is linearly independent.  $\square$

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

With respect to (4), the unique combination is called a *unique representation* of  $v$  in  $S$ . The associated coefficients are called the *Fourier coefficients* of  $v$  in  $S$ .

### Examples:

1. Consider  $\text{St}_n$ , the standard basis for  $\mathbb{F}^n$  (notice the terminology). This is, of course, a basis
2.  $\mathbb{F}[t]_n$ , the space of polynomials with degree at most  $n$ , has a basis  $\{1, t, t^2, \dots, t^n\}$ .
3. In  $\mathbb{F}^3$ ,  $\{1, 0, -1\}, (0, 1, -2), (0, 0, 1)\}$  is a basis.
4. The standard basis of  $\mathbb{F}[t]$ , the space of *all* polynomials, is  $\{1, t, t^2, \dots\} = \{t^n : n \in \mathbb{N}\}$ . One checks linear independence of this space by considering all finite subsets (remember, we do not take infinite sums). Note: in this case,  $0 \in \mathbb{N}$
5. Define  $\mathbb{F}[[t]]$  to be the set of all power series, i.e.  $\left\{ \sum_{n \in \mathbb{N}} a_n t^n : a_n \in \mathbb{F} \right\}$ . In the bullet above, we consider the space of polynomials, i.e. formal power series with *finitely* many non-zero terms. Not so for  $\mathbb{F}[[t]]$ , in generality. We ask: does this have a basis?

## 1.2 Every vector space has a basis

Since  $V$  may be infinite, we will have to rely on some non-rigorous notions to prove this in any short form. Suppose  $V$  is a vector space over  $\mathbb{F}$ , and let  $S_0 = \emptyset$  be a trivial, independent subspace. If  $S_0$  is maximally independent, then we are done. Otherwise, there exists  $v_1 \in V$  such that  $S_1 := S_0 \cup \{v_1\}$  is also independent. If  $S_1$  is maximal, then we are done. Otherwise, choose

PROOF ATTEMPT.

$v_2 \in V$ , define  $S_2 := S_1 \cup \{v_2\}$ , and so on and so on. This last notion ("and so on and so on") is problematic when  $V$  is not finite. To resolve this, we'll need to learn and understand *Zorn's Lemma*  $\square$

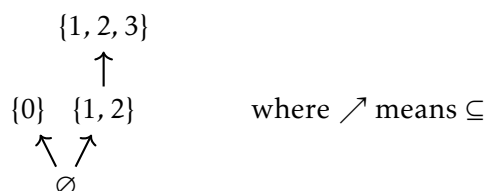
### Zorn's Lemma

The definition of ambient sets is not necessary to understand Zorn's lemma, but you can read about it here

Let  $X$  be some ambient set and  $\mathcal{I}$  be a collection of subsets of  $X$ . In other words,  $\mathcal{I} \subseteq \mathcal{P}(X)$ , the powerset of  $X$ . Call a set  $S \in \mathcal{I}$  an *inclusion-maximal* element if  $\nexists$  any strict superset  $S' \supsetneq S$  such that  $S' \in \mathcal{I}$ . Call a collection of sets  $C \subseteq \mathcal{P}(X)$  a *chain* if, for any two sets  $A, B \in C$ , one has  $A \subseteq B$  or  $B \subseteq A$ .

To demonstrate these definitions, let  $X := \mathbb{N}$  and  $\mathcal{I} := \{\emptyset, \{0\}, \{1, 2\}, \{1, 2, 3\}\} \subseteq \mathcal{P}$ .

Both  $\{0\}$  and  $\{1, 2, 3\}$  are inclusion-maximal in  $\mathcal{I}$ : adding any element to either of these sets will land you outside of  $\mathcal{I}$ .  $C_1 = \{\emptyset, \{1, 2\}, \{1, 2, 3\}\}$  is a chain, but  $C_2 = \{\emptyset, \{1, 2\}, \{0\}\}$  is *not* a chain.



Lastly, define an *upper bound* of  $\mathcal{J} \subseteq \mathcal{P}(X)$  to be a set  $U \subseteq X$  such that  $U \supseteq J$  for all sets  $J \in \mathcal{J}$ .

### 1.3 Zorn's Lemma

Let  $X$  be a set,  $\mathcal{I} \subseteq \mathcal{P}(X)$  non-empty. If every chain  $C \subseteq \mathcal{I}$  has an upper bound in  $\mathcal{I}$ , then  $\mathcal{I}$  has a maximal element.

The proof for this is statement beyond this course (see MATH 488).

Let's revisit the statement that every vector space has a basis, now equipped with Zorn's lemma:

PROOF.

Let  $\mathcal{I}$  be the collection of linearly independent subspaces in  $V$ . This is non-empty, since at least the empty set is linearly independent. If one can show that  $\mathcal{I}$  has a maximal element, in the sense of Zorn's lemma, then this element is also maximally independent.

Consider a chain  $C \subseteq \mathcal{I}$ , and let  $S := \cup C$  be the union of all sets in  $C$ . This is clearly an upper bound of  $C$ , so we want to show that it is linearly independent. However,  $S$  may be infinite, so consider an arbitrary subset  $\{v_1, \dots, v_n\} \subseteq S$ .

Let  $S_i \in C$  be some set that contains  $v_i$ , from the set described above. Since  $C$  is a chain, for any  $i, j$ , we have  $S_i \subseteq S_j$  or  $S_j \subseteq S_i$ , so WLOG we can order these sets as follows:

$$S_1 \subseteq S_2 \subseteq \dots \subseteq S_n$$

Thus,  $v_1, \dots, v_n \in S_n$ , and since  $S_n \in C \subseteq \mathcal{I}$  (recall the definition of  $\mathcal{I}$ ),  $S_n$  is linearly independent. Thus,  $\{v_1, \dots, v_n\} \subseteq S_n$  is linearly independent  $\implies S$  is linearly independent  $\implies S \in \mathcal{I}$  is an upper-bound of  $\mathcal{I}$ .

Zorn's lemma is satisfied, so  $\mathcal{I}$  has a maximal element, and we are done.  $\square$

### Steinitz Substitution Lemma

#### 1.4 Cardinality of Bases

For a vector space  $V$  over  $\mathbb{F}$ , any two bases have the same cardinality.

We'll require another lemma to prove this statement:

#### 1.5 Steinitz Substitution Lemma

Let  $V$  be a vector space over  $\mathbb{F}$ . Let  $Y \subseteq V$  be a finite, linearly independent set and  $Z \subseteq V$  be a finite spanning set. Then the following hold:

- (a)  $|Y| \leq |Z|$
- (b)  $\exists Z' \subseteq Z$  such that  $Y \cup Z'$  still spans  $V$ , where  $|Z'| = |Z| - |Y|$

*Proof TBD*

Now we'll show theorem 1.4:

Let  $Y$  and  $Z$  be two finite bases for  $V$ . Then  $Y$  is linearly independent and  $Z$  is spanning. Thus,  $|Y| \leq |Z|$  by Steinitz. However,  $Y$  is also spanning, and  $Z$  is linearly independent, so  $|Z| \leq |Y| \implies |Y| = |Z|$ .  $\square$

PROOF.

For a vector space  $V$  over  $\mathbb{F}$ , define the *dimension* of  $V$ , denoted by  $\dim(V)$ , to be the cardinality of its (i.e. any) basis. We call  $V$  a *finite dimensional vector space* if  $\dim(V)$  is a natural number, otherwise its *infinite dimensional*.

Let  $V$  have  $\dim(V) = n$ . Then the following hold by Steinitz:

PROPOSITION 1.8

- (a) For every linearly independent set  $I \subseteq V$ ,  $|I| \leq n$ . If  $|I| = n$ , then  $I$  is a basis.
- (b) For every spanning set  $S \subseteq V$ ,  $|S| \geq n$ . If  $|S| = n$ , then  $S$  is a basis.
- (c) Every linearly independent set  $I$  can be completed to a basis for  $V$ , i.e.  $\exists$  a basis  $B$  for  $V$  which contains  $I$

PROOFS.

(a) Since a basis  $B$  is spanning, one has  $|I| \leq |B| = n$

(b) Since a basis  $B$  is independent,  $|B| \leq |S|$ , i.e.  $|S| \geq n$

(c) Let  $I$  be independent and  $B$  be a basis. Then  $\exists B' \subseteq B$  with  $I \cup B'$  spanning.  $I \cup B'$  is also independent: we know that  $|I \cup B'| \geq n$ . However,  $|I \cup B'| \leq |I| + |B'| = |B| = n$ . Thus,  $|I \cup B'| = n$ . It follows that this set is minimally spanning, since  $|I \cup B'| = n - 1$  is a contradiction of (b).  $\implies |I \cup B'|$  is a basis.  $\square$

### 1.6 Monotonicity of Dimension

Let  $V$  be a finite dimensional vector space. Then for any subspace  $W \subseteq V$ ,  $\dim(W) \leq \dim(V)$  and  $\dim(W) = \dim(V) \iff W = V$ .

PROOF.

Let  $B$  be a basis for  $W$ . Since  $B$  is independent and  $W \subseteq V$ ,  $|B| \leq \dim(V)$  by proposition 1.8, so  $\dim(W) \leq \dim(V)$ .

( $\implies$ ) If  $|B| = \dim(V)$ , then  $B$  is a basis for  $V$  by 1.8, so  $\text{Span}(B) = V$ , or  $W = V$ . The ( $\impliedby$ ) direction is trivial.  $\square$

## II Linear Transformations

### AXIOMS AND INITIAL PROPERTIES

Let  $V, W$  be vector spaces over  $\mathbb{F}$ . One calls a mapping  $T : V \rightarrow W$  a *linear transformation* if it preserves vector space structure, i.e.

$$1. T(v_1 + v_2) = T(v_1) + T(v_2) \quad \forall v_1, v_2 \in V$$

$$2. T(\alpha v) = \alpha T(v) \quad \forall v \in V, \alpha \in \mathbb{F}$$

Immediately, we have that  $T(0_V) = 0_W$  and  $T(-v) = -T(v)$ .

#### Examples:

1. Consider  $T : \mathbb{F}^2 \rightarrow \mathbb{F}^2 : T(a_1, a_2) = (a_1 + 2a_2, a_1)$ . This is a linear transformation. Checking the axioms:  $T(a_1 + b_1, a_2 + b_2) = (a_1 + b_1 + 2(a_2 + b_2), a_1 + b_1) = (a_1 + 2a_2 + b_1 + 2b_2, a_1 + b_1) = T(a_1, a_2) + T(b_1, b_2)$ . Also,  $T(\alpha a_1, \alpha a_2) = (\alpha a_1 + 2\alpha a_2, \alpha a_1) = \alpha(a_1 + 2a_2, a_1) = \alpha T(a_1, a_2)$ .
2. Let  $\theta$  be an angle, and  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  be the rotation of a vector by  $\theta$ . This is a linear transformation.
3.  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , the reflection transformation defined by  $T(a_1, a_2) = T(a_1, -a_2)$
4. The transpose  $M_n(\mathbb{F}) \rightarrow M_n(\mathbb{F}) : A \rightarrow A^T$
5.  $\mathcal{D}$ , the derivative of finite polynomials.

#### 2.1 Linear transformations are completely determined by values on a basis

Let  $B := v_1, \dots, v_n$  be a basis for a vector space  $V$ . Let  $W$  be a vector space over a common field  $\mathbb{F}$ , and  $w_1, \dots, w_n \in W$ . Then there exists a unique linear transformation  $T : V \rightarrow W$  which sends  $T(v_i) = w_i \quad \forall i \in [1, n]$ .

*Existence:* Let  $v \in V$ ,  $B \subseteq V$  a basis for  $V$ , and consider some transformation  $T(v)$ . We write  $v = a_1 v_1 + \dots + a_n v_n$ ,  $v_i \in B$ , the unique representation of  $v$  in  $B$ . Now, define  $T(v) = a_1 w_1 + \dots + a_n w_n$  for fixed  $w_i \in W$ . This will indeed send  $T(v_i) = w_i$  as desired. To show that  $T$  is linear, one checks the axioms:

For  $u, v \in V$ , let  $v = a_1 v_1 + \dots + a_n v_n$  and  $u = b_1 v_1 + \dots + b_n v_n$  be the unique representations of  $u, v$  in  $B$ . Then  $u + v = (a_1 + b_1)v_1 + \dots + (a_n + b_n)v_n$ , so  $T(u + v) = (a_1 + b_1)w_1 + \dots + (a_n + b_n)w_n = a_1 w_1 + \dots + a_n w_n + b_1 w_1 + \dots + b_n w_n = T(u) + T(v)$ .

PROOF.

$T(\alpha v) = \alpha T(v)$  follows immediately from its definition.

*Uniqueness:* Suppose  $T_1, T_2$  are both such that  $T_1(v_i) = w_i = T_2(v_i)$  for all  $i$ . One shows that  $T_1(v) = T_2(v) \forall v \in V$ . Let  $v = a_1 v_1 + \dots + a_n v_n$  be the unique representation of  $v$  in  $B$ . By linearity,  $T(v) = a_1 T(v_1) + \dots + a_n T(v_n) = a_1 w_1 + \dots + a_n w_n$  for both  $T_1$  and  $T_2$ . Since  $a_i$  and  $w_i$  are all fixed, we see that  $T_1(v) = T_2(v)$ .  $\square$

## 2.2 Extension of Functions on Basis

Let  $V, W$  be vector spaces, possibly infinite, over  $\mathbb{F}$ , and let  $\beta$  be a basis for  $V$ . Every function  $T : \beta \rightarrow W$  can be extended to a linear transformation  $\hat{T} : V \rightarrow W$ .

PROOF.

This is essentially the infinite case of theorem 1.2:

*Existence:* Let  $T : \beta \rightarrow \gamma$  be (any) function between the bases of  $V$  and  $W$ . For  $v \in V$ , let  $v = a_1 v_1 + \dots + a_n v_n$  be its unique representation in  $\beta$ , where  $v_i \in \beta$ . Define the function

$$\hat{T}(v) = a_1 T(v_1) + \dots + a_n T(v_n)$$

We'll show that this is linear. Let  $x, y \in V$ . Without loss of generality, we can write  $x = a_1 v_1 + \dots + a_m v_m$  and  $y = b_1 v_1 + \dots + b_m v_m$  as their unique representations, where  $a_i, b_i$  may be zero. We thus have

$$\begin{aligned} \hat{T}(x + y) &= (a_1 + b_1)T(v_1) + \dots + (a_m + b_m)T(v_m) \\ &= a_1 T(v_1) + b_1 T(v_1) + \dots + a_m T(v_m) + b_m T(v_m) \\ &= \hat{T}(x) + \hat{T}(y) \end{aligned}$$

$$\begin{aligned} \hat{T}(\alpha x) &= \alpha a_1 T(v_1) + \dots + \alpha a_m T(v_m) \\ &= \alpha [a_1 T(v_1) + \dots + a_m T(v_m)] = \alpha \hat{T}(x) \end{aligned}$$

*Uniqueness:* Let  $\hat{T}$  be as defined, and let  $\tilde{T} : V \rightarrow W$  be another transformation which also sends  $\beta \rightarrow \gamma$  according to  $T : \beta \rightarrow \gamma$ . Fix  $v \in V$ , and let  $a_1 v_1 + \dots + a_n v_n$  be its unique representation in  $\beta$ .

$$\hat{T}(v) = a_1 T(v_1) + \dots + a_n T(v_n) = a_1 \tilde{T}(v_1) + \dots + a_n \tilde{T}(v_n) = \tilde{T}(a_1 v_1 + \dots + a_n v_n) = \tilde{T}(v)$$

$\square$

$\hat{T}$  is indeed an extension of  $T$ . See that  $\hat{T}(v_i) = T(v_i)$ , since  $v_i$  is its own representation in  $\beta$ .

## ISOMORPHISMS

Define an *isomorphism*  $T : V \rightarrow W$ , for two vector space  $V, W$  over  $\mathbb{F}$ , to be a linear transformation which admits a linear inverse.

If there exists an isomorphism between  $V$  and  $W$ , one says that  $V$  and  $W$  are *isomorphic* (to eachother). Write  $V \cong W$ .

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

PROPOSITION 2.1

This may seem trivial, and the  $(\implies)$  direction is. However, we need to show that, for  $T$  linear and bijective, its inverse is linear:

PROOF.

We know  $T^{-1}$  exists, since  $T$  is bijective. Let  $w_1, w_2 \in W$  and  $a_1, a_2 \in \mathbb{F}$ :

$$\begin{aligned} T^{-1}(a_1 w_1 + a_2 w_2) &= T^{-1}[a_1 T(T^{-1}(w_1)) + a_2 T(T^{-1}(w_2))] \\ &= T^{-1}[T(a_1 T^{-1}(w_1)) + T(a_2 T^{-1}(w_2))] \\ &= T^{-1}[T(a_1 T^{-1}(w_1) + a_2 T^{-1}(w_2))] \\ &= a_1 T^{-1}(w_1) + a_2 T^{-1}(w_2) \quad \square \end{aligned}$$

**2.3 Freeness of Vector Spaces**

All bijections from  $\beta \rightarrow \gamma$  can be extended to a unique isomorphism between  $V$  and  $W$ . This follows from Theorem 2.2.

**2.4 Isomorphism with Same Dimension**

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

Fix a basis  $B := \{v_1, \dots, v_n\}$  for  $V$ . Let  $V \rightarrow \mathbb{F}^n$  be the unique transformation which sends  $T(v_i) = e_i$ , where  $e_i = \{0, \dots, 0, 1, 0, \dots, 0\}$ , with 1 in the  $i^{th}$  position.

PROOF.

$T$  is injective: let  $T(x) = T(y)$  for  $x, y \in V$ , and write  $x = a_1 v_1 + \dots + a_n v_n$ ,  $y = b_1 v_1 + \dots + b_n v_n$ , the unique representations of  $x, y$  in  $B$ .

Then  $T(x) = T(y) \implies a_1 e_1 + \dots + a_n e_n = b_1 e_1 + \dots + b_n e_n$ , since  $T$  sends  $v_i \rightarrow e_i$ . By uniqueness of representation in a basis, one has  $a_i = b_i$ .

$T$  is surjective: let  $w \in \mathbb{F}^n$ . Then let  $w = a_1 e_1 + \dots + a_n e_n$  be its unique representation in  $\text{St}_n$ . Then  $T(a_1 v_1 + \dots + a_n v_n) = w$ .

Recall that  $\{e_i\}_{i \leq n}$  is the standard basis,  $\text{St}_n$ , of  $\mathbb{F}^n$ .

Thus,  $V \cong \mathbb{F}^n$ , and so all  $n$ -dim vector spaces are isomorphic to each other.  $\square$

For a linear transformation  $T : V \rightarrow W$ , define its *image*, notated  $\text{Im}(T)$  or  $T(V)$ , to be the set  $\{T(v) : v \in V\}$ . Similarly, define its *kernel*, notated  $\ker(T)$  or  $T^{-1}(\mathbf{0}_W)$ , to be  $\{v \in V : T(v) = \mathbf{0}_W\}$ .

PROPOSITION 2.2

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

PROOF.

For  $\ker(T)$  : Let  $v_1, v_2 \in \ker(T)$  and  $a_1, a_2 \in \mathbb{F}$ . Then  $T(a_1 v_1 + a_2 v_2) = a_1 T(v_1) + a_2 T(v_2) = a_1 \mathbf{0}_W + a_2 \mathbf{0}_W = \mathbf{0}_W$ , so  $a_1 v_1 + a_2 v_2 \in \ker(T)$ .

For  $\text{Im}(T)$  : Let  $w_1, w_2 \in \text{Im}(T)$ . Then  $w_i = T(v_i)$  for some  $v_i \in V$ , so  $a_1 w_1 + a_2 w_2 = a_1 T(v_1) + a_2 T(v_2) = T(a_1 v_1 + a_2 v_2)$ , so  $a_1 w_1 + a_2 w_2 \in \text{Im}(T)$ .  $\square$

PROPOSITION 2.3

Let  $T : V \rightarrow W$  be a linear transformation. Let  $B \subseteq V$  be a basis for  $V$ . Then  $T(B)$  spans  $\text{Im}(T)$ . In particular,  $T$  is surjective  $\iff T(B)$  spans  $W$ .

PROOF.

Let  $w \in \text{Im}(T)$ , so  $w = T(v)$  for some  $v \in V$ . Write  $v = a_1 v_1 + \dots + a_n v_n$  to be the unique representation of  $v$  in  $B$ . Then  $w = T(v) = a_1 T(v_1) + \dots + a_n T(v_n) \in \text{Span}(\{T(v_1), \dots, T(v_n)\})$ , so  $T(B)$  spans  $\text{Im}(T)$ .

If  $T$  is surjective, then  $\text{Im}(T) = W$ , and vice-versa.  $\square$

PROPOSITION 2.4

Let  $T : V \rightarrow W$  be a linear transformation. Then the following are equivalent:

1.  $T$  is injective
2.  $\ker(T) = \{0_V\}$
3.  $T(B)$  is independent for all bases  $B \subseteq V$
4.  $T(B)$  is independent for some basis  $B \subseteq V$

PROOF.

(1)  $\iff$  (2).  $\implies$  direction trivial. ( $\impliedby$ ) Let  $\ker(T) = \{0_V\}$ , and  $T(x) = T(y)$  for some  $x, y \in V$ . Then  $T(x) - T(y) = \mathbf{0}_W = T(x - y)$ , so  $x - y \in \ker(T)$ . But then  $\mathbf{0}_V = x - y$ , so  $x = y$ .

(2)  $\implies$  (3) Fix a basis  $B := \{v_1, \dots, v_n\} \subseteq V$ . To show that  $T(B)$  is linearly independent, take a combination  $a_1 w_1 + \dots + a_n w_n$ , where  $T(v_i) = w_i$ . These  $w_i$  are distinct, since  $T$  is injective by (2)  $\implies$  (1).

Suppose  $a_1 w_1 + \dots + a_n w_n = 0$ . Then  $T(a_1 v_1 + \dots + a_n v_n) = 0$ , so  $a_1 v_1 + \dots + a_n v_n \in \ker(T)$ . Thus, by (2),  $a_1 v_1 + \dots + a_n v_n = 0$ , but  $v_i \in B$  are linearly independent, so  $a_i = 0$ .

(3)  $\implies$  (4) trivial.



(4)  $\implies$  (2) Fix  $B \subseteq V$  and let  $T(B)$  be linearly independent. Suppose  $T(v) = 0$ , and write  $v = a_1 v_1 + \dots + a_n v_n$  for  $v_i \in B$ . Then  $a_1 T(v_1) + \dots + a_n T(v_n) = 0$ , but  $T(B)$  is linearly independent, so  $a_i = 0$   $\square$

If  $V$  and  $W$  are isomorphic, they have the same dimension.

PROPOSITION 2.5

This follows directly from propositions 2.3 and 2.4: if  $V$  and  $W$  are isomorphic, then  $\exists T : V \rightarrow W$  which is bijective. Let  $B$  be a basis for  $V$ . Then  $T$  surjective  $\implies T(B)$  is spanning for  $W$  by 2.3.  $T$  injective  $\implies T(B)$  independent by 2.4. Thus,  $T(B)$  is a basis for  $W$ . But  $T$  is a bijection, so  $|T(B)| = |B|$ , and we conclude that  $\dim(V) = \dim(W)$ .  $\square$

PROOF.

If  $T : V \rightarrow W$  is an injective linear transformation, then  $\dim(W) \geq \dim(V)$ . This is something along the lines of a pigeonhole principle for vector spaces.

PROPOSITION 2.6

Since  $\text{Im}(T) \subseteq W$ , we know  $\dim(\text{Im}(T)) \leq \dim(W)$ . Thus, we show that  $\dim(\text{Im}(T)) = \dim(V)$ . But since  $T$  is injective, it is an extension of  $\hat{T} : V \rightarrow \text{Im}(T)$  which is surjective, and thus bijective. We conclude that  $V$  and  $\text{Im}(T)$  are isomorphic to each other, so they have the same dimension.  $\square$

PROOF.

For vector spaces  $V, W$  over  $\mathbb{F}$ , define the *rank* of  $T$ , denoted  $\text{rank}(T)$ , to be  $\dim(\text{Im}(T))$ . Similarly, define the *nullity* of  $T$ , denoted  $\text{null}(T)$ , to be  $\dim(\ker(T))$ .

## 2.6 Rank-Nullity (or Dimension) Theorem

Let  $V$  be finite dimensional, and  $W$  any v.s. over a common field  $\mathbb{F}$ . If  $T : V \rightarrow W$  is a linear transformation, then  $\text{null}(T) + \text{rank}(T) = \dim(V)$

This follows directly from the 1<sup>st</sup> isomorphism theorem for vector space (to be seen), along with the fact that  $\dim(V/\ker(T)) = \dim(V) - \dim(\ker(T))$ . A more manual proof is as follows:

PROOF.

Let  $\{v_1, \dots, v_k\}$  be a basis for  $\ker(T)$ . By Steinitz' Lemma, this can be completed to a basis for  $V$ , say  $\{v_1, \dots, v_k, u_1, \dots, u_{n-k}\}$ , where  $n = \dim(V)$ . If we show  $\dim(\text{Im}(T)) = n - k$ , then the theorem follows.

Recall that  $T(B)$  spans  $\text{Im}(T)$  for any basis  $B \subseteq V$ . Thus,

$$\text{Span}(\{T(v_1), \dots, T(v_k), T(u_1), \dots, T(u_{n-k})\}) = \text{Im}(T)$$

However,  $v_i \in \ker(T)$ , so  $T(v_i) = 0$ , and we conclude that  $\{T(u_1), \dots, T(u_{n-k})\}$  is spanning for  $\text{Im}(T)$ .

This set (of  $n - k$  elements) is linearly independent as well:

$$\begin{aligned} a_1 T(u_1) + \dots + a_{n-k} T(u_{n-k}) &= 0_W \\ \implies a_1 u_1 + \dots + a_{n-k} u_{n-k} &\in \ker(T) \\ \implies a_1 u_1 + \dots + a_{n-k} u_{n-k} &= b_1 v_1 + \dots + b_n v_n \quad (\text{uniquely}) \\ \implies a_1 u_1 + \dots + a_{n-k} u_{n-k} - b_1 v_1 - \dots - b_n v_n &= 0_V \\ \implies a_i &= 0 \quad \forall i \quad \text{by linear independence of basis of } V \quad \square \end{aligned}$$

PROPOSITION 2.7

Let  $V, W$  be  $n$ -dimensional vector spaces over  $\mathbb{F}$ . For a linear transformation  $T : V \rightarrow W$ , the following are equivalent:

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

PROOF.

$$\begin{aligned} T \text{ surjective} &\iff \text{rank}(T) = \dim(\text{Im}(T)) = \dim(W) = n \\ T \text{ injective} &\implies \text{null}(T) = 0, \text{ so } \text{rank}(T) = \dim(V) = n \\ \text{rank}(T) = n &\implies \text{null}(T) = 0, \text{ so } \ker(T) = \{0_V\}, \text{ so } T \text{ injective} \quad \square \end{aligned}$$

## 2.7 First Isomorphism Theorem

Let  $V, W$  be vector spaces over  $\mathbb{F}$ . Let  $T : V \rightarrow W$  be a linear transformation. Then  $V/\ker(T)$  is isomorphic to  $\text{Im}(T)$  through the isomorphism  $\bar{v} \rightarrow T(v)$ , where  $\bar{v} := v + \ker(T)$  (as in quotient groups).

PROOF.

We know that  $\hat{T} : V/\ker(T) \rightarrow \text{Im}(T)$  given by  $\hat{T}(\bar{v}) = T(v)$  is a well-defined group isomorphism. Thus, we need to check that  $\hat{T}$  is linear. In particular, we need to check scalar multiplication, since group homomorphisms obey  $T(x + y) = T(x) + T(y)$ .

$$\hat{T}(a\bar{v}) = \hat{T}(\overline{av}) = T(av) = aT(v) = a\hat{T}(\bar{v}) \quad \square$$

## THE SPACE $\text{Hom}(V, W)$

For vector spaces  $V, W$  over  $\mathbb{F}$ , define  $\text{Hom}(V, W)$  to be the set of all linear transformations from  $V \rightarrow W$ .

PROPOSITION 2.8

$\text{Hom}(V, W)$  is a vector space over  $\mathbb{F}$ , equipped with the following:

*Addition*  $T_0 + T_1$  defines the function  $T_0 + T_1 : V \rightarrow W$ , where  $(T_0 + T_1)(v) = T_0(v) + T_1(v)$ , where  $T_0, T_1 \in \text{Hom}(V, W)$ .

**Scalar Multiplication** For  $T \in \text{Hom}(V, W)$  and  $a \in \mathbb{F}$ ,  $a \times T$  defines the function  $(aT) : V \rightarrow W$ , where  $(aT)(v) = aT(v)$ .

**Zero Vector**  $\mathbb{0}_{\text{Hom}(V, W)}$  is the function which takes  $v \rightarrow \mathbb{0}_W$

### 2.8 Basis for Hom

Let  $V, W$  be vector spaces over  $\mathbb{F}$ , which have bases  $\beta, \gamma$ , respectively. The set  $\{T_{v,w} : v \in \beta, w \in \gamma\}$  is a basis for  $\text{Hom}(V, W)$ , where  $T_{v,w}$  is the unique transformation such that  $T_{v,w}(v') = w$  for all  $v' \in \beta \setminus \{v\}$ .

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

PROPOSITION 2.9

Let  $\beta = \{v_1, \dots, v_n\}$ ,  $\gamma = \{w_1, \dots, w_m\}$ . Then  $\{T_{v_i, w_j} : i \in [1, n], j \in [1, m]\}$  is a basis for  $\text{Hom}(V, W)$  by the theorem above. This has  $n \cdot m$  elements.  $\square$

PROOF.

## MATRICES

We wish to characterize a transformation  $T : \mathbb{F}^n \rightarrow \mathbb{F}^m \in \text{Hom}(\mathbb{F}^n, \mathbb{F}^m)$  in matrix form. Given  $T$ , we know it's uniquely determined by its values on the standard basis for  $\mathbb{F}^n$ ,  $\text{St}_n = \{e_1, \dots, e_n\}$ . Thus,  $T$  is uniquely determined by the ordered set

$$\{T(e_1), \dots, T(e_n)\} \subseteq \mathbb{F}^m$$

Each  $T(e_i)$  is a vector in  $\mathbb{F}^m$ , so we can represent it as  $\langle a_{i1}, \dots, a_{im} \rangle$ , where  $a_{ij} \in \mathbb{F}$ . Thus, form the following matrix of column vectors:

$$[T] := \begin{bmatrix} | & | & & | \\ T(e_1) & T(e_2) & \cdots & T(e_n) \\ | & | & & | \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1m} & \cdots & a_{nm} \end{bmatrix}$$

We call this the *matrix representation* of  $T$  in the standard basis.

$T(v) = [T] \cdot v$ , where  $v$  is represented as a column vector,  $\langle b_1, \dots, b_n \rangle$ , for  $b_i \in \mathbb{F}$ .

PROPOSITION 2.10

We have  $v = b_1 e_1 + \dots + b_n e_n$  in the standard basis. Then,  $T(v) = b_1 T(e_1) + \dots +$

PROOF.

$b_n T(e_n)$ , where  $T(e_i) = \langle a_{i1}, \dots, a_{im} \rangle \subseteq \mathbb{F}^m$ . In column-vector form, this is:

$$T(v) = \begin{bmatrix} b_1 a_{11} + \dots + b_n a_{n1} \\ b_1 a_{12} + \dots + b_n a_{n2} \\ \vdots \\ b_1 a_{1m} + \dots + b_n a_{nm} \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1m} & \cdots & a_{nm} \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = [T]v$$

□

In this way, matrices can act as linear transformations, but we would like to formalize this idea.

For a given  $m \times n$  matrix  $A$ , define the function  $L_A : \mathbb{F}^n \rightarrow \mathbb{F}^m$  by  $v \rightarrow A \cdot v$ , where  $v \in \mathbb{F}^n$ . This is a linear transformation by the proposition above.

The function from  $\text{Hom}(\mathbb{F}^n, \mathbb{F}^m) \rightarrow M_{m \times n}(\mathbb{F})$  defined by  $T \rightarrow [T]$  is an isomorphism. Furthermore, its inverse  $M_{m \times n}(\mathbb{F}) \rightarrow \text{Hom}(\mathbb{F}^n, \mathbb{F}^m)$  is  $A \rightarrow L_A$ .

*Linearity:* We need to first show  $[T_1 + T_2] = [T_1] + [T_2]$  and  $[aT] = a[T]$  for  $a \in \mathbb{F}$ ,  $T \in \text{Hom}(\mathbb{F}^n, \mathbb{F}^m)$ . Consider the standard basis for  $\mathbb{F}^n$ ,  $\text{St}_n = \{e_1, \dots, e_n\}$ . We have that

$$\begin{aligned} [T_1 + T_2] &= \begin{bmatrix} | & | & & | \\ (T_1 + T_2)(e_1) & (T_1 + T_2)(e_2) & \cdots & (T_1 + T_2)(e_n) \\ | & | & & | \end{bmatrix} \\ &= \begin{bmatrix} | & | & & | \\ T_1(e_1) + T_2(e_1) & T_1(e_2) + T_2(e_2) & \cdots & T_1(e_n) + T_2(e_n) \\ | & | & & | \end{bmatrix} \\ &= \begin{bmatrix} | & | & & | \\ T_1(e_1) & T_1(e_2) & \cdots & T_1(e_n) \\ | & | & & | \end{bmatrix} + \begin{bmatrix} | & | & & | \\ T_2(e_1) & T_2(e_2) & \cdots & T_2(e_n) \\ | & | & & | \end{bmatrix} \\ &= [T_1] + [T_2] \end{aligned}$$

$a[T] = [aT]$  is shown similarly.

*Inverse:* If an inverse exists for  $T \rightarrow [T]$ , then it is bijective; as linearity has

Proposition 2.10 established that every transformation  $T$  can be represented in matrix form. One can work backwards, too: given a matrix  $A$ , one forms the unique transformation that sends  $e_i \rightarrow A^{(j)}$ , the  $j^{\text{th}}$  column of  $A$ .

PROPOSITION 2.11

PROOF.

been shown, this is sufficient to show isomorphism by Prop. 2.1.

Consider the composition  $T \rightarrow [T] \rightarrow L_{[T]}$ . One sees  $L_{[T]}(v) = [T] \cdot v = T(v)$  by definition, so this is precisely the identity on  $\text{Hom}(\mathbb{F}^n, \mathbb{F}^m)$ .

Now we need to show  $A \rightarrow L_A \rightarrow [L_A]$  is the identity on  $M_{m \times n}(\mathbb{F})$ . Consider the  $j^{\text{th}}$  column of  $[L_A]$ . This is the result of  $L_A(e_j)$ , which is  $A \cdot e_j$ . Thus:

$$[L_A] = \begin{bmatrix} | & | & & | \\ A \cdot e_1 & A \cdot e_2 & \cdots & A \cdot e_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ A^{(1)} & A^{(2)} & \cdots & A^{(n)} \\ | & | & & | \end{bmatrix} = A$$

□

As a corollary, we get that  $\dim(\text{Hom}(\mathbb{F}^n, \mathbb{F}^m)) = \dim(M_{m \times n}(\mathbb{F}))$

PROPOSITION 2.12

### Matrix Representations in Generality

Thus far we've considered matrix representations in  $\mathbb{F}^n, \mathbb{F}^m$ , but we can do so for general vector spaces  $V, W$ .

Let  $V$  be finite dimensional over  $\mathbb{F}$ , and  $\beta = \{v_1, \dots, v_n\}$  be a basis for  $V$ . Recall the set  $\{a_1, \dots, a_n\}$  for which  $a_1 v_1 + \dots + a_n v_n = v$  is the unique representation of  $v$  in  $\beta$ . We call this set the *coordinates* of  $v$  in  $\beta$ . Represented as a column vector, define

$$[v]_{\beta} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{F}^n$$

to be the *coordinate vector* of  $v$  in  $\beta$ .

Recall that, in the proof that all  $n$ -dimensional vector spaces  $V$  are isomorphic to  $\mathbb{F}^n$ , we used the transformation  $T(v_i) = e_i$ . We denote this function by  $I_{\beta} : V \rightarrow \mathbb{F}^n$ . For any  $v \in V$ , we have

$$I_{\beta}(v) = I_{\beta}(a_1 v_1 + \dots + a_n v_n) = a_1 I(v_1) + \dots + a_n I(v_n) = a_1 e_1 + \dots + a_n e_n = [v]_{\beta}$$

Thus,  $I_{\beta} : V \rightarrow \mathbb{F}^n$  which sends  $v \rightarrow [v]_{\beta}$  is an isomorphism.

Suppose we are given  $T : V \rightarrow W$ , where  $V$  and  $W$  are both finite dimensional. Let  $\beta = \{v_1, \dots, v_n\}$  and  $\gamma = \{w_1, \dots, w_m\}$  be bases of  $V$  and  $W$ , respectively. We know that  $T$  is determined by its values on  $\beta$ . Thus, we can encode  $T$  in matrix-form,

where the  $i^{\text{th}}$  column corresponds to  $[T(v_i)]_\gamma \in \mathbb{F}^m$ , as follows:

$$[T]_\beta^\gamma := \begin{bmatrix} | & | & & | \\ [T(v_1)]_\gamma & [T(v_2)]_\gamma & \cdots & [T(v_n)]_\gamma \\ | & | & & | \end{bmatrix}$$

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

### 2.9 Relation Between $V, W, \mathbb{F}^n$ , and $\mathbb{F}^m$

Let  $V, W$  be of dimension  $n$  and  $m$  with bases  $\beta$  and  $\gamma$ , respectively. Let  $T : V \rightarrow W$  be a linear transformation. Then the following diagram commutes:

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

Furthermore, the function  $\text{Hom}(V, W) \rightarrow M_{m \times n}(\mathbb{F})$  that maps  $T \rightarrow [T]_\beta^\gamma$  is an isomorphism whose inverse is the map  $M_{m \times n}(\mathbb{F}) \rightarrow \text{Hom}(V, W)$  which maps  $A \rightarrow I_\gamma^{-1} \circ L_A \circ I_\beta$

PROOF.

To show the diagram commutes, we essentially prove  $I_\gamma \circ T = L_{[T]_\beta^\gamma} \circ I_\beta$ . We have  $I_\gamma \circ T(v) = [T(v)]_\gamma$ , applying definitions. On the other hand,

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

Thus, we need to show that  $[T]_\beta^\gamma \cdot [v]_\beta = [T(v)]_\gamma$ . To do so, write  $[v]_\beta = \langle a_1, \dots, a_n \rangle \in \mathbb{F}^n$ , and recall that

$$[T]_\beta^\gamma := \begin{bmatrix} | & | & & | \\ [T(v_1)]_\gamma & [T(v_2)]_\gamma & \cdots & [T(v_n)]_\gamma \\ | & | & & | \end{bmatrix}$$

Then we can write

$$\begin{aligned}
 [T]_\beta^\gamma \cdot [v]_\beta &= a_1[T(v_1)]_\gamma + \dots + a_n[T(v_n)]_\gamma \\
 &= [a_1 T(v_1) + \dots + a_n T(v_n)]_\gamma && \text{by linearity of } I_\gamma \\
 &= [T(a_1 v_1 + \dots + a_n v_n)]_\gamma && \text{by linearity of } T \\
 &= [T(v)]_\gamma && \text{and we are done } \square
 \end{aligned}$$

### Compositions and Matrix Multiplication

Recall the composition of functions  $T : V \rightarrow W$ ,  $S : W \rightarrow X$ , written as  $S \circ T(v) = S(T(v))$ . Compositions are associative: for functions  $T \rightarrow S \rightarrow R$ , we have  $(R \circ S) \circ T = (R \circ S)(T(v)) = R(S(T(v))) = R(S \circ T(v)) = R \circ (S \circ T(v))$ .

Consider the two linear maps  $L_A : \mathbb{F}^n \rightarrow \mathbb{F}^m$ ,  $L_B : \mathbb{F}^m \rightarrow \mathbb{F}^l$ . Then the composition  $L_B \circ L_A$  is itself a linear combination, and is equal to  $L_C : \mathbb{F}^n \rightarrow \mathbb{F}^l$  for some matrix  $C \in M_{l \times n}(\mathbb{F})$ . This unknown  $C$  is precisely  $[L_B \circ L_A]$ , by definition.

For two matrices  $A$  and  $B$ , define their product  $B \cdot A$  to be  $[L_B \circ L_A]$ .

$L_B \circ L_A = L_{B \cdot A}$ . The proof for this follows immediately from  $[L_B \circ L_A] = B \cdot A$ .

Matrix multiplication is associative.

$$C(BA) = C \cdot [L_B \circ L_A] = [L_C \circ (L_B \circ L_A)] = [(L_C \circ L_B) \circ L_A] = (CB)A \quad \square$$

For  $V, W, U$  finite-dimensional, with bases  $\alpha, \beta, \gamma$ , respectively, and transformations  $T : V \rightarrow W$ ,  $S : W \rightarrow U$ , we have the similar statement  $[S \circ T]_\alpha^\gamma = [S]_\beta^\gamma \cdot [T]_\alpha^\beta$ .

## INVARIANTS AND NILPOTENT TRANSFORMATIONS

### Preliminaries

For a function  $f : X \rightarrow Y$ , we call  $g : Y \rightarrow X$

1. a *left inverse* if  $g \circ f = I_X$ , the identity on  $X$
2. a *right inverse* if  $f \circ g = I_Y$ , the identity on  $Y$
3. an *inverse* if  $g$  is both a left and right inverse

Also consider the following facts, whose proofs are good exercise:

1.  $f$  has a left inverse  $\iff f$  is injective
2.  $f$  has a right inverse  $\iff f$  is surjective

By function, we don't just mean linear transformations

One can work out, explicitly, what  $[L_B \circ L_A]$  is, and see that it agrees with our usual notion for  $B \cdot A$

PROPOSITION 2.13

PROPOSITION 2.14

PROOF.

PROPOSITION 2.15

Where  $T := L_A$  and  $S := L_B$  as above, this is equivalent to saying  $[L_B \circ L_A] = B \cdot A$ , which has been shown.

i.e. takes  $x \rightarrow x$

Sometimes called a two-sided inverse

3.  $f$  has an inverse  $\iff f$  is bijective

### Examples:

1.  $\delta : \mathbb{F}[t]_{n+1} \rightarrow \mathbb{F}[t]_n$ , the derivative of polynomials, has a right inverse, namely the anti-derivative.
2. Let  $f : \mathbb{F}[[t]] \rightarrow \mathbb{F}[[t]]$  be the left shift map of coefficients, i.e.  $\sum_0^\infty a_n t^n \rightarrow \sum_1^\infty a_n t^{n-1}$ . This has a right inverse, namely the right shift map of coefficients,  $\sum_0^\infty a_n t^n \rightarrow \sum_0^\infty a_n t^{n+1}$ . Recall that  $\mathbb{F}[[t]]$  is the set of formal power series.

PROPOSITION 2.16

Let  $T : V \rightarrow W$  be a transformation of vector spaces of the same (finite) dimension. Then TFAE:

$T$  has a right inverse       $T$  has a left inverse       $T$  has an inverse

PROOF.

This follows directly from Prop. 2.7, which states that transformations over  $n$  dimensional spaces are surjective IFF injective  $\square$

Recall that an  $n \times n$  dimensional matrix  $A$  is called invertible IFF there exists  $B$  such that  $A \cdot B = B \cdot A = I$ , the identity matrix. One notates  $B = A^{-1}$ .

PROPOSITION 2.17

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  $\iff$  it has a right inverse.

PROOF.

$L_A$  is invertible  $\iff$  there exists  $T : \mathbb{F}^n \rightarrow \mathbb{F}^n$  such that  $L_A \circ T = T \circ L_A = I_{\mathbb{F}^n} \iff \exists B \in M_n(\mathbb{F})$  with  $L_A \circ L_B = L_B \circ L_A = I_{\mathbb{F}^n} \iff \exists B$  s.t.  $L_{AB} = L_{BA} = I_{\mathbb{F}^n} \iff \exists B$  s.t.  $AB = BA = [I]$ , and  $[I]$  is the identity matrix (this last bit has not been previously shown, but the verification is easy).

This shows (1), and (2) follows directly.  $\square$

### $T$ -Invariants

Let  $T : V \rightarrow V$  be a linear transformation over a vector space  $V$ . Transformations of this form are sometimes called *linear operators*. A subspace  $W \subseteq V$  is called  *$T$  invariant* if  $T(W) \subseteq W$ .

i.e., you can apply  $T$  to  $W$  an indeterminate amount of times, and it will always remain as a subset of itself.

### Examples:

1. For  $T : V \rightarrow V$ , both  $\ker(T)$  and  $\text{Im}(T)$  are  $T$ -invariant

For (1), note that  $T(\text{Im}(T)) \subseteq \text{Im}(T)$  by definition, and  $T(\ker(T)) = \{0_V\} \in \ker(T)$



2. For any  $n \in \mathbb{N}$ , where  $T^n := \underbrace{T \circ T \circ \dots \circ T}_{n \text{ times}}$ ,  $\ker(T^n)$  is  $T$ -invariant.
- 

Proof to come