Honours
Linear Algebra II

Contents

1. Some Algebraic Background

In MAT2042, one has studied vector spaces over $\mathbb{R}$ or $\mathbb{C}$. One main goal of this course is to generalize the theories to other algebraic structures other than $\mathbb{R}$ or $\mathbb{C}$. We will roughly go through the very basics of abstract algebras, and give a definition of a field.

1.1. Groups.

When one talks about algebraic structure, we would think of addition $a+b$ and multiplication $a \cdot b$. In general, we make the following definition:

Definition 1.1.1. *Let S be a set. A* **binary operation** *S is a map*

$$* : S \times S \to S.$$

A subset $T \subseteq S$ is **closed under** $*$ *if for all $a, b \in T$, $a * b \in T$.*

Definition 1.1.2. *A group G is a set along with a binary operation $* : G \times G \to G$ satisfying:*

- $(a * b) * c = a * (b * c)$;
- *There exists $e \in G$ such that $e * g = g * e = g$ for all $g \in G$.*
- *For all $g \in G$, there exists $g^{-1} \in G$ such that $g * g^{-1} = g^{-1} * g = e$.*

Example 1.1.3. (1) $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$ *are groups.*
(2) $(\mathbb{R}[x], +)$ *is a group.*
(3) *(Modular arithmetic)* $(\mathbb{Z}_n, +)$ *is a group.*
(4) $(\mathbb{Q}\backslash\{0\}, \cdot)$ *is a group.*
(5) $(GL_n(\mathbb{R}), \cdot)$ *is a group.*

Definition 1.1.4. *A group $(G, *)$ is called* **abelian/commutative** *if*

$$a * b = b * a$$

for all $a, b \in G$.

Example 1.1.5. $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$, $(\mathbb{R}[x], +)$, $(\mathbb{Z}_n, +)$, $(\mathbb{Q}\backslash\{0\}, \cdot)$ *are commutative, but $(GL_n(\mathbb{R}), \cdot)$ is not commutative.*

1.2. Rings.

Now we study algebraic structure with both addition and multiplication structures:

Definition 1.2.1. *A* **ring** $(R, +, \cdot)$ *is a set with two binary operations $+, \cdot : R \times R \to R$ such that:*

- $(R, +)$ *is an abelian group;*
- $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ *for all $a, b, c \in R$;*
- $(a + b) \cdot c = a \cdot c + b \cdot c$ *and $c \cdot (a + b) = c \cdot a + c \cdot b$ for all $a, b, c \in R$;*

Example 1.2.2. (1) $(R, +, \cdot)$ *with $R = \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}_n, \mathbb{R}[x]$ are rings.*

(2) $(2\mathbb{Z}, +, \cdot)$ *is a ring.*
(3) *For all ring R, $(M_{n\times n}(R), +, \cdot)$ (here $\cdot$ is matrix multiplication) is a ring.*

Note that $(R, \cdot)$ is not necessarily a group. For instance, $2\mathbb{Z}$ does not have a multiplicative identity $1 \notin 2\mathbb{Z}$.

In this course, we will focus on the following kind of rings:

Definition 1.2.3. *Let $(R, +, \cdot)$ is a ring. We say R is*

- **unital** *if there exists $1_R \in R$ such that 1;*
- **commutative** *if $a \cdot b = b \cdot a$ for all $a, b \in R$;*

Example 1.2.4. (1) *$(R, +, \cdot)$ with $R = \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}_n, \mathbb{R}[x]$ are unital and commutative.*
(2) *$2\mathbb{Z}$ is not unital but commutative.*
(3) *$M_{n\times n}(R)$ is unital if R is unital. However, it is always not commutative for $n > 1$.*

Note that if $R \neq \{0\}$, then $(R, \cdot)$ is never a group: To begin with, since $(R, +)$ is a(n abelian) group, it must have the additive identity element $0 \in R$. On the other hand, suppose on contrary that $(R, \cdot)$ is a group, then it must be unital and contain the multiplicative identity $1 \in R$. However, there exists no $a \in R$ such that

$$0 = 0 \cdot a = a \cdot 0 = 1$$

otherwise one has $0 = r \cdot 0 = r \cdot 1 = r$ for all $r \in R$.

1.3. **Fields.** Now we can make precise which algebraic objects we can generalize from $\mathbb{R}$ and $\mathbb{C}$ for vector spaces:

Definition 1.3.1. *Let $(\mathbb{F}, +, \cdot)$ be a unital commutative ring. We say $\mathbb{F}$ is a* **field** *if for all $a \in \mathbb{F}\backslash\{0\}$, there exists $a^{-1} \in \mathbb{F}\backslash\{0\}$ such that*

$$a \cdot a^{-1} = 1.$$

The **characteristic** *of $\mathbb{F}$ is the smallest positive number $\mathrm{char}(\mathbb{F}) = p$ such that*

$$\overbrace{1 + 1 + \cdots + 1}^{p\ terms} = 0$$

(if no such number exists, then we let $\mathrm{char}(\mathbb{F}) = 0$.

Example 1.3.2. (1) *$\mathbb{Z}$ is not a field, but $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are fields.*
(2) *$\mathbb{Z}_6$ is not a field, since 2^{-1} does not exist.*
(3) *In general, $\mathbb{Z}_p$ is a field if and only if p is a prime number. In such a case, $\mathrm{char}(\mathbb{Z}_p) = p$.*

In the Homework set, you will construct a field $\mathbb{F}$ with 9 elements.

2. Vector Spaces

2.1. Definition of vector space.

Definition 2.1.1 (Vector Space)**.** *Let $\mathbb{F}$ be a field. A vector space V over $\mathbb{F}$ is a set equipped with two operations*

$$+ : V \times V \to V \text{ (addition)} \quad \text{and} \quad \cdot : \mathbb{F} \times V \to V \text{ (scalar multiplication)}$$

such that the following holds:

- *Additive axioms - For every $x, y, z \in V$, we have*
  1. $x + y = y + x$.
  2. $(x + y) + z = x + (y + z)$.
  3. *There exists $\mathbf{0} \in V$ such that $\mathbf{0} + x = x + \mathbf{0} = x$.*
  4. *There exists $-x \in V$ such that $(-x) + x = x + (-x) = \mathbf{0}$.*
- *Multiplicative axioms - For every $x \in V$ and $\alpha, \beta \in \mathbb{F}$, we have*
  1. $0 \cdot x = \mathbf{0}$
  2. $1 \cdot x = x$
  3. $(\alpha\beta) \cdot x = \alpha \cdot (\beta \cdot x)$
- *Distributive axioms - For every $x, y \in V$ and $\alpha, \beta \in \mathbb{F}$, we have*
  1. $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$.
  2. $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$.

Example 2.1.2. (1) $\mathbb{F}^n$ *(all n-vectors) is a vector space over $\mathbb{F}$;*
(2) $M_{m\times n}(\mathbb{F})$ *(all $m \times n$-matrices) is a vector space over $\mathbb{F}$;*
(3) $\mathbb{R}[x]$, $C^\infty(\mathbb{R})$ *are vector spaces over $\mathbb{R}$;*
(4) $V_1 := \{(\alpha_1, \alpha_2, \dots) \mid \alpha_i \in \mathbb{R}\}$ *is a vector space over $\mathbb{R}$.*
(5) $V_2 := \{(\alpha_1, \alpha_2, \dots) \in V_1 \mid \lim_{i\to\infty} \alpha_i = 0\}$ *is a vector space over $\mathbb{R}$.*
(6) $V_3 := \{(\alpha_1, \alpha_2, \dots) \in V_2 \mid$ *only finitely many α_i are nonzero*$\}$ *is a vector space over $\mathbb{R}$.*

Definition 2.1.3 (Vector Subspace)**.** *Let V be a vector space over $\mathbb{F}$ with addition $+$ and scalar multiplication $\cdot$. A subset $W \leq V$ is a vector subspace if the operations satisfy:*

$$+|_{W\times W} : W \times W \to W, \qquad \cdot|_{\mathbb{F}\times W} : \mathbb{F} \times W \to W,$$

i.e. for all $w_1, w_2 \in W$ and $\alpha \in \mathbb{F}$, $w_1 + w_2 \in W$ and $\alpha \cdot w_1 \in W$.

The following proposition, proved in MAT2042, gives a necessary and sufficient condition for any subset of a vector space is a vector subspace:

Proposition 2.1.4. *Let V be a vector space over $\mathbb{F}$, a subset W of V is a vector subspace if and only if*

$$\textit{for all } a, b \in \mathbb{F} \textit{ and all } w_1, w_2 \in W,\ aw_1 + bw_2 \in W.$$

Example 2.1.5. (1) *(MAT2042) In $V = \mathbb{F}^n$, all vector subspaces are of the form $W = \text{Span}_{\mathbb{F}}\{v_1, \dots, v_k\}$ for some vectors $v_i \in \mathbb{F}^n$.*

(2) *In $V = \mathbb{F}[x]$, the subset*

$$W = \{p(x) \in \mathbb{F}[x] \mid p(x) = p(-x)\}$$

is a vector subspace.

(3) *In $V = C^\infty(\mathbb{R})$, the subset*

$$W = \{f(x) \in C^\infty(\mathbb{R}) \mid f'(1) = f''(2) = 0\}$$

is a vector subspace.

(4) *$V_3 \leq V_2 \leq V_1$ in the Example 2.1.2 above.*

(5) *If $\{W_i \mid i \in I\}$ is a collection of vector subspaces of V, then*

$$\bigcap_{i \in I} W_i \leq V$$

is also a vector subspace.

2.2. **Basis and Dimension.** We will briefly go through some well-known notions in MAT2042.

Definition 2.2.1. *Let V be a vector space over $\mathbb{F}$, and $S \subseteq V$ is a (not necessarily finite) subset. We say*

- *$v \in V$ is a* **linear combination of** *S if v can be expressed as*

$$v = \alpha_1 s_1 + \dots \alpha_k s_k$$

for some $\alpha_1, \dots, \alpha_k \in \mathbb{F}$, $s_1, \dots, s_k \in S$ a $k \in \mathbb{N}$. In particular, the above sum must be finite even when S is infinite.

- *The* **span** *of S is the collection of linear combinations of S:*

$$\text{Span}_{\mathbb{F}}(S) = \{\alpha_1 s_1 + \dots \alpha_k s_k \mid \alpha_i \in \mathbb{F},\ s_i \in S,\ k \in \mathbb{N}\}$$

- *S* **spans** *V (or S* **is a spanning set of** *V) if $V = \text{Span}(S)$.*
- *S is* **linear independent** *if for any finite subset $\{s_1, \dots, s_n\} \subseteq S$,*

$$\alpha_1 s_1 + \dots + \alpha_n s_n = \mathbf{0} \quad \Leftrightarrow \quad \alpha_1 = \dots = \alpha_n = 0.$$

- *S is a* **basis** *of V if S is a spanning set of V and S is linearly independent.*

Note that in the definition of linear combination of S, the sum must be finite.

Example 2.2.2. (1) *$V = \mathbb{F}[x]$, and $S = \{x^i \mid i \in \mathbb{N} \cup \{0\}\}$, then*

$$e + \frac{1}{\pi}x^3 + \pi^e x^5 \in \text{Span}_{\mathbb{F}}(S), \qquad 1 + x + x^2 + \dots \notin \text{Span}_{\mathbb{F}}(S).$$

(2) *$V = V_1$ be given in Example 2.1.2, and $\mathcal{B} = \{e_i \mid i \in \mathbb{N}\}$ where*

$$e_i := (0, \dots, 0, \overbrace{1}^{i \text{ entry}}, 0, \dots).$$

Then S is linearly independent in V_1, and $\mathrm{Span}(S) = V_3 < V_1$.

Here are some examples of basis of some vector spaces V:

Example 2.2.3. (1) *For $V = \mathbb{F}^n$, $\{e_1, \dots, e_n\}$ is a basis of V.*
(2) *For $V = M_{m\times n}(\mathbb{F})$, $\{E_{ij} \mid 1 \le i \le m,\ 1 \le j \le n\}$ is a basis of V.*
(3) *For $V = \mathbb{F}[x]$, $\{x^i \mid i \in \mathbb{N} \cup \{0\}\}$ is a basis of V.*
(4) *As in Example 2.1.2(2), $\mathcal{B}$ is a basis of V_3 but not a basis of V_1.*

The importance of basis is given by the following:

Theorem 2.2.4. *Let V be a vector space over $\mathbb{F}$, and $\mathcal{B} \subseteq V$ is a subset. Then*

$\mathcal{B}$ is a basis of V iff all $v \in V$ can be **uniquely** *expressed by a element in* $\mathrm{Span}(\mathcal{B})$.

Proof. We first explain what it means for a vector $v \in V$ to be **uniquely** expressed by a element in $\mathrm{Span}(\mathcal{B})$: Namely, if

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n = \beta_1 v_1' + \dots + \beta_m v_m' \tag{1}$$

for some *nonzero* $\alpha_i, \beta_j \in \mathbb{F}$ and $v_i, v_j \in \mathcal{B}$, then one has (1) $n = m$, and (2) after some reordering of the indices, one has $v_i = v_i'$ for all $1 \le i \le n = m$.

($\Rightarrow$)If $\mathcal{B}$ is a basis of V, then by definition all $v \in V$ is in $\mathrm{Span}(\mathcal{B})$. As for uniqueness, suppose (1) holds. Then by reordering the indices, we assume that

$$\{v_1, \dots, v_n\} \cap \{v_1', \dots, v_m'\} = \{v_1 = v_1', \dots, v_k = v_k'\}.$$

and by subtraction, one has

$$(\alpha_1 - \beta_1)v_1 + \dots + (\alpha_k - \beta_k)v_k + \alpha_{k+1}v_{k+1} + \dots \alpha_{k+1}v_n + (-\beta_{k+1})v_{k+1}' + \dots + (-\beta_m)v_m' = \mathbf{0}$$

Now by applying linear independence of $\mathcal{B}$ on the finite subset

$$\{v_1 = v_1', \dots, v_k = v_k', v_{k+1}, \dots, v_n, v_{k+1}', \dots, v_m'\},$$

the above equation only has zero solutions

$$\alpha_1 - \beta_1 = \dots = \alpha_k - \beta_k = 0$$

and $\alpha_{k+1} = \dots \alpha_n = \beta_{k+1} = \dots = \beta_m = 0$. But the latter is impossible since we assume that all $\alpha_i, \beta_j \ne 0$, so one must have $m = n = k$, and

$$\alpha_i = \beta_i \qquad \text{for all } 1 \le i \le k.$$

($\Leftarrow$)If all $v \in V$ can be expressed by elements in $\text{Span}(\mathcal{B})$, then $\mathbb{B}$ is obviously a spanning set. Now suppose on contrary that $\mathcal{B}$ is **not** linearly independent, then there exists $\{b_1, \dots, b_n\} \subseteq \mathcal{B}$ and $\gamma_1, \dots, \gamma_n$ not all zeros such that

$$\gamma_1 b_1 + \cdots + \gamma_n b_n = \mathbf{0}$$

By reordering the indices in the above expression, we assume $\gamma_1 \neq 0$ without loss of generality. Then

$$v = b_1 = (-\frac{\gamma_2}{\gamma_1})b_2 + \cdots + (-\frac{\gamma_n}{\gamma_1})b_n$$

are two different expressions of v in $\text{Span}(\mathcal{B})$, a contradiction. Hence $\mathcal{B}$ is both a spanning set and linearly independent, and form a basis of V. $\square$

The following theorem are given in MAT2042 in the finite dimensional case:

Theorem 2.2.5. *Let V be a vector space over $\mathbb{F}$. Then the following holds:*

(1) *V has a basis.*
(2) *Let $\mathcal{L}$ be a linearly independent set in V, then one can extend $\mathcal{L}$ to $\mathcal{B} = \mathcal{L} \sqcup \mathcal{L}'$ such that $\mathcal{B}$ is a basis of V.*
(3) *Let $\mathcal{S}$ be a linearly independent set in V, then there exists a subset $\mathcal{B} \subseteq \mathcal{S}$ such that $\mathcal{B}$ is a basis of V.*
(4) *All bases of V have the same cardinality.*

Given the last statement of the theorem above, we have:

Definition 2.2.6. *Let V be a vector space over $\mathbb{F}$. Then the* **dimension** $\dim(V)$ *of V is the cardinality of any basis of V.*

Example 2.2.7. (1) $\dim(\mathbb{F}^n) = n$;
(2) $\dim(M_{m\times n}(\mathbb{F}) = mn$;
(3) $\dim(\mathbb{F}[x]) = \infty$ *(and countable).*
(4) *For V_3 in Example 2.1.2,* $\dim(V_3) = \infty$ *(and countable).*

2.3. **Internal Direct Sum.** In the following sections, we will give some constructions of new vector spaces from old ones.

Definition 2.3.1 (Internal Sum). *Let V be a vector space over $\mathbb{F}$, and $\{W_i \mid i \in I\}$ be a collection of vector subspaces of V. The* **internal sum** *is defined by:*

$$\sum_{i\in I} W_i := \{w_{i_1} + \cdots + w_{i_k} \mid w_{i_l} \in W_{i_l},\ \{i_1, \dots, i_k\} \text{ is a finite subset of } I\}$$

Note that $\sum_{i\in I} W_i$ is a vector subspace of V, and if $\mathcal{S}_i$ spans W_i for all $i \in I$, then $\bigcup_{i\in I} \mathcal{S}_i$ spans $\sum_{i\in I} W_i$. However, if $\mathcal{L}_i$ is linearly independent in W_i for all $i \in I$, then $\bigcup_{i\in I} \mathcal{L}_i$ may not be linearly independent.

To see so, let $V = \mathbb{R}^3$, $W_1 = \text{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$ and $W_2 = \text{Span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$. Then

$$W_1 + W_2 = V$$

but $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ is linearly dependent.

Definition 2.3.2. *Let V be a vector space over $\mathbb{F}$, and $\{W_i \mid i \in I\}$ be a collection of vector subspaces of V. We say $\sum_{i\in I} W_i = \bigoplus_{i\in I} W_i$ is a* **direct internal sum** *if for any $\{i_1, \ldots, i_k\} \subseteq I$ and any $w_{i_l} \in W_{i_l}$,*

$$w_{i_1} + \cdots + w_{i_k} = \mathbf{0} \qquad \Leftrightarrow \qquad w_{i_1} = \cdots = w_{i_k} = \mathbf{0}.$$

In the paragraph above the definition, one has

$$w_1 := \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \in W_1, \qquad w_2 = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} = (-1)\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in W_2$$

with $w_1 + w_2 = \mathbf{0}$ yet $w_1, w_2 \neq \mathbf{0}$. So the sum $W_1 + W_2$ is **not** direct.

Theorem 2.3.3. *Let V be a vector space over $\mathbb{F}$, and $\{W_i \mid i \in I\}$ be a collection of vector subspaces of V such that $\bigoplus_{i\in I} W_i$ is a direct sum. Suppose $\mathcal{B}_i$ is a basis of W_i, then*

$$\mathcal{B} = \bigsqcup_{i\in I} \mathcal{B}_i$$

is a basis of $\bigoplus_{i\in I} W_i$.

Proof. To begin with, we show that all $v \in V$ can be uniquely written as elements of the form $w_{i_1} + \cdots + w_{i_k}$ for some $w_{i_l} \in W_{i_l}$. Indeed, suppose

$$v = w_{i_1} + \cdots + w_{i_n} = w'_{j_1} + \cdots + w'_{j_m},$$

for some nonzero vectors $w_\bullet \in W_\bullet$. Then by reordering the indices and subtraction as in the proof of Theorem 2.2.4, one assumes that

$$\{i_1, \ldots, i_n\} \cap \{j_1, \ldots, j_m\} = \{i_1 = j_1, \ldots, i_k = j_k\}$$

and

$$\mathbf{0} = (w_{i_1} - w'_{j_1}) + \cdots + (w_{i_k} - w'_{i_k}) + w_{i_{k+1}} + \cdots + w_{i_n} - w'_{j_{k+1}} - \cdots - w_{j_m},$$

where each term in the summand is in the same vector subspace $W_\bullet \leq V$. By the definition of direct internal sum, all the summand in the above equation must be $\mathbf{0}$, and one must have $m = n = k$ and $w_{i_1} = w'_{j_1}, \cdots, w_{i_k} = w'_{i_k}$.

Now we proceed to proving that $\mathcal{B}$ is a basis. In view of Theorem 2.2.4, it suffices to show that all $v \in V$ can be uniquely written as an element in $\mathrm{Span}(\mathcal{B})$. By definition of direct internal sum, it is obvious that $v \in \mathrm{Span}(\mathcal{B})$. Moreover, in such a case, one has

$$v = b_{i_1} + \cdots + b_{i_k}, \qquad b_{i_l} \in \mathrm{Span}(\mathcal{B}_{i_l}).$$

But $b_{i_l} = \mathrm{Span}(\mathcal{B}_{i_l}) = W_{i_l}$, so the arguments in the beginning of the proof implies that the W_{i_l}'s appearing in the summand of v is unique, and each $b_{i_l} \in W_{i_l}$ is uniquely determined. Furthermore, the expression of each b_{i_l} by the linear combination of $\mathcal{B}_{i_l}$ is also unique since it is a basis of W_{i_l}. So the expression of v in terms of $\mathrm{Span}(\mathcal{B})$ is unique, and the result follows. $\square$

As an immediate consequence of the above theorem, one has:

Corollary 2.3.4. *Let V be a vector space over $\mathbb{F}$, and $\{W_i \mid i \in I\}$ be a collection of vector subspaces of V such that $\bigoplus_{i\in I} W_i$ is a direct sum. Then*

$$\dim\left(\bigoplus_{i\in I} W_i\right) = \sum_{i\in I} \dim(W_i)$$

(here we do not distinguish countable and uncountable ∞).

2.4. **External Direct Sum.**

Definition 2.4.1. *Let $\{V_i \mid i \in I\}$ be a collection of vector spaces over $\mathbb{F}$. The* **external direct sum** *is defined by*

$$\bigoplus_{i\in I} V_i := \{f : I \to \cup_{i\in I} V_i \mid f(i) \in V_i \text{ such that only finitely many } f(i) \text{ is nonzero}\}.$$

The addition and scalar multiplication on $\bigoplus_{i\in I} V_i$ is given by:

- *If $f, g \in \bigoplus_{i\in I} V_i$, then $f + g \in \bigoplus_{i\in I} V_i$ is defined by $(f+g)(i) := f(i) + g(i)$ for all $i \in I$.*
- *If $f \in \bigoplus_{i\in I} V_i$, then $\alpha f \in \bigoplus_{i\in I} V_i$ is defined by $(\alpha f)(i) := \alpha(f(i))$ for all $i \in I$.*

and the zero vector $\mathbf{0}$ is given by $\mathbf{0}(i) := \mathbf{0}_{V_i}$ for all $i \in I$.

Remark 2.4.2.

- *Any $f \in \bigoplus_{i\in I} V_i$ is uniquely determined by the image $f(i)$ of f. For example, if $I = \{1, 2, ..., n\}$ is a finite set, we can write:*

$$f \longleftrightarrow (f(1), f(2), \ldots, f(n))$$

or if $I = \mathbb{N}$, we can write:

$$f \longleftrightarrow (f(1), f(2), \ldots,)$$

Informally, we will see $\bigoplus_{i\in I} V_i$ as:

$$\{(\cdots, \overbrace{v_i}^{\text{position } i}, \cdots) \mid v_i \in V_i,\ \text{only finitely many } v_i \text{ not equal to } \mathbf{0}_{V_i}\}$$

- *If I is a finite set, then there only only finitely many $f(i)$'s, and hence the condition that only finitely many $f(i) \neq \mathbf{0}_{V_i}$ is vacuous.*
- *Under the above interpretation of $\bigoplus_{i\in I} V_i$, if*

  $$u \longleftrightarrow (\cdots, u_i, \cdots, u_j, \cdots), \qquad v \longleftrightarrow (\cdots, v_i, \cdots, v_j, \cdots)$$

  Then the addition and scalar multiplication given by the definition above can be understood as:

  $$u + v \longleftrightarrow (\cdots, u_i + v_i, \cdots, u_j + v_j, \cdots), \quad \alpha u \longleftrightarrow (\cdots, \alpha u_i, \cdots, \alpha u_j, \cdots).$$

Example 2.4.3. *Let $I = \mathbb{N}$ and $V_i = \mathbb{R}$ for all i. Then*

$$\bigoplus_{i\in\mathbb{N}} \mathbb{R} \approx \{(\alpha_1, \alpha_2, \dots) \mid \alpha_i \in \mathbb{R}, \text{only finitely many } \alpha_i \neq 0\}$$

is isomorphic to V_3 given in Example 2.1.2 (we have not defined what isomorphism means in this course, but you should know what it is about in MAT2042).

The condition that $\bigoplus_{i\in I} V_i$ allows only finitely many nonzero element seems artificial, but it is natural in the sense that one only allows a finite sum for linear combination and the definition of internal sum. In particular, it is essential for the following:

Theorem 2.4.4. *Let $\{V_i \mid i \in I\}$ be a collection of vector spaces over $\mathbb{F}$. Suppose $\mathcal{B}_i := \{b_i^j \mid j \in J_i\}$ is a basis of V_i for each i, consider $f_i^j \in \bigoplus_{i\in I} V_i$ given by*

$$f_i^j(k) := \begin{cases} b_i^j & \text{if } k = i \\ \mathbf{0}_{V_k} & \text{if } k \neq i \end{cases}$$

(Informally, one can interpret $f_i^j \longleftrightarrow (\cdots, \mathbf{0}, \mathbf{0}, \overbrace{b_i^j}^{\text{position } i}, \mathbf{0}, \mathbf{0}, \cdots)$). Then

$$\mathcal{B} := \{f_i^j \mid i \in I,\ j \in J_i\}$$

is a basis of $\bigoplus_{i\in I} V_i$.

Proof. For linear independence, suppose $(i_1, j_1), \dots, (i_k, j_k)$ be such that $j_l \in J_{i_l}$ for all l, consider

$$\alpha_i f_{i_1}^{j_1} + \cdots + \alpha_i f_{i_k}^{j_k} = \mathbf{0}.$$

For each $1 \leq l \leq k$, write $i_l := \iota$. Suppose

$$\{l' \mid i_{l'} = \iota\} = \{l, x_1, \dots, x_m\}$$

Then one has

$$
\begin{aligned}
(\alpha_i f_{i_1}^{j_1} + \cdots + \alpha_i f_{i_k}^{j_k})(\iota) &= \mathbf{0}(\iota) \\
\alpha_l f_\iota^{j_l}(\iota) + \alpha_{x_1} f_\iota^{j_{x_1}}(\iota) + \cdots + \alpha_{x_m} f_\iota^{j_{x_m}}(\iota) &= \mathbf{0}_{V_\iota} \\
\alpha_l b_\iota^{j_l} + \alpha_{x_1} b_\iota^{j_{x_1}} + \cdots + \alpha_{x_m} b_\iota^{j_{x_m}} &= \mathbf{0}_{V_\iota}
\end{aligned}
$$

But $\{b_\iota^j \mid j \in I_\iota\}$ is a basis of V_ι, hence it is linearly independent, and hence ($\alpha_{x_1} = \cdots = \alpha_{x_m} =$) $\alpha_l = 0$. Note that we can apply the same argument for all $1 \leq l \leq k$, so one has $\alpha_1 = \cdots = \alpha_k = 0$.

As for spanning set, consider $f \in \bigoplus_{i \in I} V_i$. By definition of direct external sum, the set

$$\{i \in I \mid f(i) \neq \mathbf{0}_{V_i}\} = \{i_1, \ldots, i_k\}$$

is finite, with

$$f(i_l) = \alpha_{l,1} b_{i_l}^{j_{l,1}} + \cdots + \alpha_{l,n_l} b_{i_l}^{j_{l,n_l}} \in V_{i_l} = \operatorname{Span}(\mathcal{B}_{i_l})$$

Then one can check that

$$f = (\alpha_{1,1} f_{i_1}^{j_{1,1}} + \cdots + \alpha_{1,n_1} f_{i_1}^{j_{1,n_1}}) + \cdots + (\alpha_{k,1} f_{i_k}^{j_{k,1}} + \cdots + \alpha_{1,n_1} f_{i_k}^{j_{k,n_k}})$$

is in $\operatorname{Span}(\mathcal{B})$ (note that it is a finite sum). $\square$

Corollary 2.4.5. *Let $\{V_i \mid i \in I\}$ be a collection of vector spaces over $\mathbb{F}$. Then*

$$\dim(\bigoplus_{i \in I} V_i) = \sum_{i \in I} \dim(V_i).$$

2.5. **External Direct Product.** As we discussed before, one may remove the finitely many nonzero condition. In such a case, we have

Definition 2.5.1. *Let $\{V_i \mid i \in I\}$ be a collection of vector spaces over $\mathbb{F}$. The* **external direct product** *is defined by*

$$\prod_{i \in I} V_i := \{f : I \to \cup_{i \in I} V_i \mid f(i) \in V_i\}.$$

The addition, scalar multiplication and $\mathbf{0}$ on $\prod_{i \in I} V_i$ is given exactly as in that of external direct sum.

Example 2.5.2. *The vector*

$$v := (1, 1, \ldots) \in \prod_{i \in \mathbb{N}} \mathbb{R}$$

but is not in $\bigoplus_{i \in \mathbb{N}} \mathbb{R}$. More generally, one has $V_1 \cong \prod_{i \in \mathbb{N}} \mathbb{R}$ for the V_1 defined in Example 2.1.2.

Also, by the discussions in Section 2.4, the set $\mathcal{B} := \{f_i \mid i \in \mathbb{N}\}$ *given by*

$$f_i(k) := \delta_{ik} = \begin{cases} 1 & if\ i = k \\ 0 & if\ i \neq k \end{cases}$$

is a basis of $\bigoplus_{i\in\mathbb{N}} \mathbb{R}$ *(informally,* $f_i = (0, \ldots, 0, \overbrace{1}^{i-position}, 0, \ldots,))$*. However,* $\mathcal{B}$ **is only linearly independent but not a spanning set in** $\prod_{i\in\mathbb{N}} \mathbb{R}$ *- namely*

$$v \text{ “=” } f_1 + f_2 + \ldots \text{ is an infinite sum, so } v \notin \mathrm{Span}(\mathcal{B}).$$

2.6. **Quotient Spaces.** Let V be a vector space over $\mathbb{F}$, and $W \leq V$. Define an equivalence relationship $\sim$ by

$$v \sim v' \quad \Longleftrightarrow v - v' \in W.$$

The equivalence class with representative $v \in V$ is defined by

$$v + W := \{v' \in V \mid v' \sim v\}$$

We call $v + W$ a **coset with representative** v.

Proposition 2.6.1. *Let* $v + W$*,* $u + W$ *be cosets.*

- $v + W = \{v + w \mid w \in W\}$.
- *As subsets of* V*, either* $(v+W) = (u+W)$ *are equal or* $(v+W)\cap(u+W) = \phi$ *are disjoint.*
- $(v + W) = (u + W)$ *iff* $u = v + w$ *for some* $w \in W$*.*

Example 2.6.2. *Let* $V = \mathbb{R}^3$, $W = \{\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \mid x, y \in \mathbb{R}\}$ *be the xy-plane. Then the coset* $\begin{pmatrix} a \\ b \\ c \end{pmatrix} + W$ *is the horizontal plane in* $\mathbb{R}^3$ *elevated/lowered to level* c*. In particular, one has*

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + W = \begin{pmatrix} 5 \\ 10 \\ 1 \end{pmatrix} + W = \begin{pmatrix} \pi \\ e \\ 1 \end{pmatrix} + W.$$

Definition 2.6.3. *Let* V *be a vector space over* $\mathbb{F}$*, and* $W \leq V$*. The* **quotient space** V/W *is defined by the set of cosets:*

$$V/W := \{v + W \mid v \in V\}$$

(i.e. a ‘vector’ in $v + W \in V/W$ *is a coset), with addition and scalar multiplication given by*

- $(v+W)+(u+W):=(v+u)+W$;
- $\alpha\cdot(v+W):=(\alpha v)+W$

Remark 2.6.4. *Although the arithmetic is simple in V/W, one needs to be careful that one may have* **different** *expressions for the* **same** *element in V/W. In such a case, one needs to show that we get the same addition and scalar multiplication even if we use different representatives.*

For instance, suppose $v+W=v'+W$ and $u+W=u'+W$ (for possibly $v\neq v'$ and $u\neq u'$, one needs to show that

$$(u+W)+(v+W)=(u'+W)+(v'+W).$$

One can apply Proposition 2.6.1 to verify this.

Example 2.6.5. (1) *Let $V=\mathbb{R}^3$ and W is the xy-plane as in Example 2.6.2. We have seen that there are a lot of repetitions $v+W=v'+W$. But they are all be reduced to*

$$V/W:=\{\begin{pmatrix}0\\0\\z\end{pmatrix}+W \mid z\in\mathbb{R}\}$$

and the operations are given by

$$\left(\begin{pmatrix}0\\0\\z_1\end{pmatrix}+W\right)+\left(\begin{pmatrix}0\\0\\z_2\end{pmatrix}+W\right)=\begin{pmatrix}0\\0\\z_1+z_2\end{pmatrix}+W,$$

$$\alpha\left(\begin{pmatrix}0\\0\\z\end{pmatrix}+W\right)=\begin{pmatrix}0\\0\\\alpha z\end{pmatrix}+W.$$

So the space is 'isomorphic' (once again we have not defined it in this course yet) to $\mathbb{R}$.

(2) *Let $V=\mathbb{F}[x]$, and $W=\{p(x)\mid p(x)$ is divisible by $(x^2+1)\}=\{(x^2+1)q(x)\mid q(x)\in\mathbb{F}[x]\}$ (check the latter is a vector subspace). Then the elements in V/W are of the form $p(x)+W$. To remove all repetitions, note that division algorithm of polynomials give*

$$p(x)=(x^2+1)q(x)+r(x)$$

for $r(x)=ax+b$ of degree less than $2=\deg(x^2+1)$. Then one has

$$p(x)+W=r(x)+\overbrace{(x^2+1)q(x)}^{\in W}+W=(ax+b)+W,$$

where the last equality comes from Proposition 2.6.1. In other words,

$$V/W=\{(ax+b)+W \mid a,b\in\mathbb{F}\}$$

Note that $\{1+W, x+W\}$ *is a basis of* V/W.

(3) *Let* $V = \prod_{i\in\mathbb{N}} \mathbb{R}$ *and* $W = \{(\alpha_1, \alpha_2, \dots) \in V \mid \alpha_1 = 0\}$. *Then all* $(\alpha_1, \alpha_2, \dots) \in V$ *can be written as*

$$(\alpha_1, \alpha_2, \dots) = (\alpha_1, 0, 0, \dots) + \overbrace{(0, \alpha_2, \alpha_3, \dots)}^{\in W}$$

and hence one has

$$V/W = \{(\alpha, 0, 0, \dots) + W \mid \alpha \in \mathbb{R}\}.$$

3. Linear Transformation

3.1. **Basic Definitions.** The notion of linear transformation is the same as in MAT2042. We will quickly go through them in this section.

Definition 3.1.1 (Linear Transformation)**.** *s*

Example 3.1.2.
(1) *(matrix transformation) Let $A \in M_{m\times n}(\mathbb{F})$, $T : \mathbb{F}^n \to \mathbb{F}^m$ given by*
$$T(x) := Ax$$
is a linear transformation.
(2) *$T : M_{n\times n}(\mathbb{F}) \to \mathbb{F}$ given by*
$$T(M) := M_{ij}$$
is a linear transformation.
(3) *(trace) $tr : M_{n\times n}(\mathbb{F}) \to \mathbb{F}$ given by*
$$T(M) := M_{11} + \cdots + M_{nn}$$
is a linear transformation.
(4) *(determinant is* **not** *linear) $det : M_{n\times n}(\mathbb{F}) \to \mathbb{F}$ given by*
$$det(M) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} M_{i\sigma(i)}$$
is **not** *a linear transformation for $n \geq 2$, since* $\det(2M) = 2^n \det(M) \neq 2\det(M)$ *in general.*
(5) *(differentiation) $D : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ given by*
$$D(f) := f'$$
is a linear transformation.
(6) *(integration) $I : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ given by*
$$I(f) := \int_a^x f(t)dt$$
is a linear transformation.

Definition 3.1.3. *Let V, W be vector spaces over $\mathbb{F}$.*
(1) *The set of all linear transformations*
$$\mathcal{L}(V,W) := \{T : V \to W \mid T \text{ is linear transformation}\}$$
has a vector space structure given by: for $T, S \in \mathcal{L}(V,W)$ and $\alpha \in \mathbb{F}$,
- *$(T + S) \in \mathcal{L}(V,W)$ is defined by $(T + S)(v) := T(v) + S(v)$;*
- *$\alpha T \in \mathcal{L}(V,W)$ is defined by $(\alpha T)(v) := \alpha(T(v))$.*

(2) *If $W = V$, we write $\mathcal{L}(V) := \mathcal{L}(V,V)$, and $T \in \mathcal{L}(V)$ is called a* **linear operator** *of V.*
(3) *If $W = \mathbb{F}$ (treated as a 1-dimensional vector space), we write $V^* := \mathcal{L}(V,\mathbb{F})$ the* **dual vector space** *of V. And $f \in V^*$ is called a* **linear functional** *of V.*

The following theorem is well-known and has the same proof as in MAT2042:

Theorem 3.1.4. *Let V, W, U be vector spaces over $\mathbb{F}$, and $T : V \to W$, $S : W \to U$ are linear transformations.*

(1) *The composition $S \circ T : V \to U$ is also a linear transformation.*
(2) *Suppose T is bijective, then its inverse $T^{-1} : W \to V$ is also a linear transformation. In such a case we call $V \cong W$ are* **isomorphic**, *and T is an* **isomorphism** *between V and W.*
(3) *Let $\mathcal{B}$ be a basis of V, then T is uniquely determined by the values*

$$\{T(b) \mid b \in \mathcal{B}\}.$$

Remark 3.1.5. *As a converse to the last statement of the above theorem, suppose $\mathcal{B} = \{b_i \mid i \in I\}$ is a basis of V, and $\{w_i \mid i \in I\}$ is any subset of W. Then one can define a linear transformation $T : V \to W$ given by*

$$T(\alpha_1 b_{i_1} + \cdots + \alpha_k b_{i_k}) := \alpha_1 w_{i_1} + \cdots + \alpha_k w_{i_k}$$

for all $\alpha_l \in \mathbb{F}$ and all finite subset $\{i_1, \ldots, i_k\} \subseteq I$. In particular, this is the unique linear transformation satisfying

$$T(b_i) = w_i$$

for all $i \in I$.

Definition 3.1.6 (Kernel and Image)**.**

Theorem 3.1.7. *Let $T : V \to W$ be a linear transformation.*

(1) $\ker(T) \leq V$, $\mathrm{im}(T) \leq W$.
(2) *T is injective iff $\ker(T) = \{\mathbf{0}_V\}$.*
(3) *T is surjective iff $\mathrm{im}(T) = W$.*

Example 3.1.8. *Suppose $\dim(V) = \dim(W) = n$ with $\mathcal{B} = \{v_1, \ldots, v_n\}$ and $\mathcal{C} = \{w_1, \ldots, w_n\}$ be bases of V and W respectively. Then $T : V \to W$ defined by*

$$T(v_i) := w_i$$

(c.f. Remark 3.1.5) is an isomorphism of vector space by checking its kernel and image. The same result holds if $\dim(V)$ and $\dim(W)$ are of countable infinite dimension.

Finally, we have an important theorem in MAT2042. We will give an alternative proof of it using quotient spaces.

Theorem 3.1.9 (Rank-Nullity Theorem)**.** *Let V be a finite dimensional vector space over $\mathbb{F}$, and $T : V \to W$ be a linear transformation. Then*

$$\dim(\ker(T)) + \dim(\mathrm{im}(T)) = \dim(V).$$

3.2. **Linear Transformation on Quotient Spaces.** In the previous Chapter, we have constructions of new vector spaces from old ones. In this section and the next section, we will construct new linear transformations from old ones.

Proposition 3.2.1. *Let V be a vector space over $\mathbb{F}$, and $V' \leq V$. The map $\pi_{V'} : V \to V/V'$ given by*

$$\pi_{V'}(v) := v + V'$$

is a linear transformation with $\ker(\pi_{V'}) = V'$ and $\mathrm{im}(\pi_{V'}) = V/V'$.

Proposition 3.2.2. *Let $T : V \to U$ be a linear transformation. Suppose $S \leq \ker(T)$, then one can define a linear transformation*

$$\overline{T} : V/S \to U$$

given by

$$\overline{T}(v + S) := T(v).$$

In other words, $T = \overline{T} \circ \pi$, i.e. the following diagram commutes:

$$\begin{tikzcd} V \arrow[r, "\pi"] \arrow[dr, "T"'] & V/S \arrow[d, "\overline{T}"] \\ & U \end{tikzcd}$$

Proof. □

Theorem 3.2.3. *Let $T : V \to W$ be a linear transformation. Then there is an isomorphism*

$$V/\ker(T) \cong \mathrm{im}(T).$$

Proof. □

Corollary 3.2.4. *The rank-nullity theorem (Theorem 3.1.9) holds.*

Proof. □

3.3. **Dual Spaces.** In this section, we will study some properties and linear transformations related to dual vector spaces $V^* = \mathcal{L}(V, \mathbb{F})$.

Definition 3.3.1. *Let V be a vector space over $\mathbb{F}$, and $\mathcal{B} = \{b_i \mid i \in I\}$ be a basis of V. For each $i \in I$, let $f_i \in V^*$ be given by*

$$f_i(b_j) := \delta_{ij}$$

(c.f. Remark 3.1.5). Define

$$\mathcal{B}^* := \{f_i \mid i \in I\}$$

Note that $\mathcal{B}^*$ and $\mathcal{B}$ have the same cardinality.

Example 3.3.2. (1) *Let $V = \mathbb{F}^n$ and $\mathcal{B} = \{e_1, \ldots, e_n\}$ be the canonical basis of V. Then $f_i \in V^*$ is defined by*

$$f_i \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_i.$$

Check that f_i satisfies the condition given by the above definition.

(2) *Let $V = M_{n\times n}(\mathbb{F})$ and $\mathcal{B} = \{E_{ij} \mid 1 \leq i, j \leq n\}$, where E_ij is the matrix with 1 on the (i,j)-entry and zero on the other entries. Then*

$$f_{ij}(M) := M_{ij}$$

and hence $tr = f_{11} + \cdots + f_{nn}$.

We hope that $\mathcal{B}^*$ is a basis of V^*. To check if it is the case, we begin by proving the following:

Proposition 3.3.3. *$\mathcal{B}^*$ is linearly independent in V^*.*

Corollary 3.3.4. *Let V be a finite dimensional vector space over $\mathbb{F}$ with basis $\mathcal{B}$. Then $\mathcal{B}^*$ is a basis of V^*, and $V^* \cong V$ are isomorphic.*

Proof. By Homework, one check that

$$\dim(V^*) = \dim(\mathcal{L}(V, \mathbb{F}) = \dim(V) \cdot \dim(\mathbb{F}) = \dim(V) \cdot 1 = \dim(V) = n.$$

Now $\mathcal{B}^*$ is a linearly independent set with $\dim(V^*) = n$ elements. By basis extension (Theorem 2.2.5(2)), one can extend $\mathcal{B}^*$ to a basis $\mathcal{B}^* \sqcup \mathcal{E}$ of V^*. But $|\mathcal{B}^*| = |\mathcal{B}| = n$, so $\mathcal{E} = \phi$ is empty, otherwise we will have a basis having more than n elements, violating Theorem 2.2.5(4).

The last statement of the corollary is given in Example 3.1.8. □

How about the case when $\dim(V) = \infty$? We still have $\mathcal{B}^*$ linear independent, but it is no longer a spanning set in general:

Example 3.3.5. *Let* $V = \mathbb{F}[x]$ *and* $\mathcal{B} = \{x^i \mid i \in \mathbb{N} \cup \{0\}\}$*. Then* $\mathcal{B}^* = \{f_i \mid i \in \mathbb{N} \cup \{0\}\}$ *with*

$$f_i(\alpha_n x^n + \cdots + \alpha_1 x + \alpha_0) := \alpha_i.$$

Now consider $\phi \in V^*$ *given by* $\phi(p(x)) := p(1)$*. Then we claim that* $\phi \notin \mathrm{Span}(\mathcal{B}^*)$ *- suppose on contrary*

$$\phi = \gamma_0 f_0 + \cdots + \gamma_k f_k,$$

then applying x^{k+1} *on both sides yield:*

$$\begin{aligned} \phi(x^{k+1}) &= \gamma_0 f_0(x^{k+1}) + \cdots + \gamma_k f_k(x^{k+1}) \\ 1^{k+1} &= \gamma_0 \cdot 0 + \cdots + \gamma_k \cdot 0 \\ 1 &= 0 \end{aligned}$$

which gives a contradiction.

Indeed, one can check that if $\dim(V) = \infty$ is countable , then $\dim(V^*) = \infty$ is uncountable. We omit the details here.

Definition 3.3.6. *Let* $S \subset V$ *be a subset. Then* **annihilator** *of* S *is*

$$\mathrm{Ann}(S) := \{f \in V^* \mid f(s) = 0 \textit{ for all } s \in S\}.$$

Example 3.3.7.

Proposition 3.3.8. *Let* V *be a vector space over* $\mathbb{F}$*.*

(1) *If* $S \subset S'$ *then* $\mathrm{Ann}(S) \supset \mathrm{Ann}(S')$*.*
(2) *If* $W_1, W_2 \leq V$*, then*
$$\mathrm{Ann}(W_1 \cap W_2) = \mathrm{Ann}(W_1) + \mathrm{Ann}(W_2) \textit{ and}$$
$$\mathrm{Ann}(W_1 + W_2) = \mathrm{Ann}(W_1) \cap \mathrm{Ann}(W_2).$$
(3) *If* $\dim(V) < \infty$ *and* $W \leq V$*, then* $\dim(\mathrm{Ann}(W)) = \dim(V) - \dim(W)$*.*

By Example 3.1.8, if V is finite dimensional, then $\mathrm{Ann}(W)$ and V/W have the same dimension and hence

$$\mathrm{Ann}(W) \cong V/W.$$

However, the isomorphism given by example 3.1.8 is not **natural** (i.e. the isomorphism involves a choice of basis). Moreover, it cannot be generalized to infinite dimensional cases. We wish to construct a natural isomorphism instead, but it will be on the dual space

$$\mathrm{Ann}(W) \cong (V/W)^*.$$

Lemma 3.3.9 (Complementation)**.** *Let* V *be a vector space over* $\mathbb{F}$*, and* $W \leq V$*. Then there exists a complementary subspace* $U \leq V$ *such that*

$$V = W \oplus U.$$

Proof. Consider the set

$$\mathcal{S} := \{U' \mid U' \leq V, \quad U' \cap W = \{\mathbf{0}\}\}.$$

Then

- $\mathcal{S}$ is an ordered set by inclusion as sets; and
- For all $U' \in \mathcal{S}$, the sum $W \oplus U'$ is direct

By Zorn's Lemma, $\mathcal{S}$ has a maximal element U. We claim that

$$V = W \oplus U.$$

Suppose on contrary that there exists $v \in V \backslash (W \oplus U)$. Then by letting $\widetilde{U} := \mathrm{Span}(U \cup \{v\})$, one checks that $\widetilde{U} > U$, and $\widetilde{U} \cap W = \{\mathbf{0}\}$. In other words, $\widetilde{U} \in \mathcal{S}$, violating the maximality of U. $\square$

Proposition 3.3.10. *Suppose $V = W \oplus U$ as in the Lemma above, then*

$$V/W \cong U.$$

Proof. Let $p : V \to U$ be defined as follows: for all $v \in V$, $v = w + u$ for some unique choices of $w \in W$ and $u \in U$. Then

$$p(v) = p(w + u) := u.$$

One checks that p is a linear transformation, with $\ker(p) = W$ and $\mathrm{im}(p) = U$. Then the result follows from the first isomorphism theorem. $\square$

Theorem 3.3.11. *Let V be a vector space over $\mathbb{F}$, and $W \leq V$. Then one has an isomorphism*

$$(V/W)^* \cong \mathrm{Ann}(W)$$

Proof. Take $V = W \oplus U$ as in the complementation lemma, then one has $V/W \cong U$ by the above proposition, and hence it suffices to check that

$$U^* \cong \mathrm{Ann}(W)$$

(check that if $V \cong V'$, then $V^* \cong (V')^*$).

We now construct $T : U^* \to \mathrm{Ann}(W)$ - for $f \in U^*$, let $\widetilde{f} : V \to \mathbb{F}$ by

$$\widetilde{f}(v) = \widetilde{f}(w + u) := f(u),$$

where $v = w + u$ is the unique expression of v. Then one defines

$$T : U^* \to \mathrm{Ann}(W)$$

$$T(f) := \widetilde{f}.$$

There are a few things to check:

(1) $\widetilde{f} \in V^*$: let $v = w + u, v' = w' + u'$ be the unique expressions of $v, v' \in V$. Then

$$\begin{aligned}\widetilde{f}(\alpha v + \beta v') &= \widetilde{f}((\alpha w + \beta w') + (\alpha u + \beta u')) \\ &= f(\alpha u + \beta u') \\ &= \alpha f(u) + \beta f(u') \\ &= \alpha \widetilde{f}(v) + \beta \widetilde{f}(v')\end{aligned}$$

(2) $\widetilde{f} \in \mathrm{Ann}(W)$: for all $w \in W$, $w = \mathbf{0} + w$ is its unique expression. So

$$\widetilde{f}(w) = \widetilde{f}(\mathbf{0} + w) = f(\mathbf{0}) = \mathbf{0}$$

(3) $T : U^* \to \mathrm{Ann}(W)$ is a linear transformation: for $f, g \in U^*$,

$$\begin{aligned}T(\alpha f + \beta g)(v) &= (\widetilde{\alpha f + \beta g})(w + u) \\ &= (\alpha f + \beta g)(u) \\ &= \alpha f(u) + \beta g(u) \\ &= \alpha \widetilde{f}(v) + \beta \widetilde{g}(v) \\ &= (\alpha \widetilde{f} + \beta \widetilde{g})(v).\end{aligned}$$

So $T(\alpha f + \beta g) = \alpha \widetilde{f} + \beta \widetilde{g}$.

(4) T is injective: Suppose $T(f) = \widetilde{f}$ is zero, then $\widetilde{f}(w + u) = f(u) = 0$ for all $v = w + u$. Hence $f \in U^*$ is the zero transformation.

(5) T is surjective: for any $g \in \mathrm{Ann}(W)$, then for all $v = w + u \in V$, one has

$$g(v) = g(w + u) = g(w) + g(u) = g(u).$$

Define $f \in U^*$ by $f(u) := g(u)$. Then f is obviously linear, and

$$T(f)(v) = \widetilde{f}(v) = f(u) = g(u) = g(v)$$

for all $v \in V$. So $g = T(f) \in \mathrm{im}(T)$.

□

3.4. **Universal Property of Vector Spaces.** In this section, we will construct some new linear transformations from some known ones which possesses some nice properties. Our prototype is Proposition 3.2.2:

Proposition 3.4.1. *Let V be a vector space over $\mathbb{F}$, and $W \leq V$. Consider the collection of all linear transformations ϕ with $W \leq \ker(\phi)$, i.e.*

$$\mathcal{C}_{qs} := \{(U, \phi) \mid \phi : V \to U \quad satisfying \quad W \leq \ker(\phi)\}$$

Then

(a) *Let* $\pi_W : V \to V/W$ *be the canonical projection. Then*

$$(W, \pi_W) \in \mathcal{C}_{qs};$$

(b) *For* $T : V \to U$ *such that* $(U, T) \in \mathcal{C}_{qs}$*, one has*

$$T = \beta \circ \pi_W \quad \text{with} \quad \beta(v + W) := T(v).$$

In other words, there is a uniquely defined $\beta : V/W \to U$ *such that following diagram is commutative:*

$$\begin{tikzcd} V \arrow[r, "\pi_W"] \arrow[dr, "T"'] & V/W \arrow[d, "\beta"] \\ & U \end{tikzcd}$$

Proof. Since $W \leq \ker(\phi)$ by the fact that $(U, \phi) \in \mathcal{C}_{qs}$, so the hypothesis of Proposition 3.2.2 is satisfied, and one can take $\beta = \overline{T}$ in the proposition to obtain the result. $\square$

As for *external direct sum*, one has the following:

Proposition 3.4.2. *Let* $\{V_i \mid i \in I\}$ *be collection of vector spaces over* $\mathbb{F}$*. Consider the collection of all linear transformations* $\phi_i : V_i \to U$*, i.e.*

$$\mathcal{C}_{ds} := \{(U, \{\phi_i\}_{i \in I}) \mid \phi_i : V_i \to U\}.$$

Then

(a) *For all* $i \in I$*, let* $\alpha_i : V_i \to \bigoplus_{i \in I} V_i$ *be defined by* $\alpha_i(v_i) := (\cdots, \mathbf{0}, \overbrace{v_i}^{\text{position } i}, \mathbf{0}, \cdots)$*. Then*

$$\left(\bigoplus_{i \in I} V_i,\ \{\alpha_i\}_{i \in I}\right) \in \mathcal{C}_{ds};$$

(b) *For* $T_i : V_i \to U$ *such that* $(U, \{T_i\}_{i \in I}) \in \mathcal{C}_{ds}$*, one has*

$$T_i = \beta \circ \alpha_i \quad \text{with} \quad \beta((\cdots, v_i, \cdots)) := \textstyle\sum_{i \in I} T_i(v_i)$$

(note that the sum $\sum_{i \in I} T_i(v_i)$ *is finite by the definition of external direct sum).*

In other words, there is a unique $\beta : \bigoplus_{i \in I} V_i \to U$ *such that the following diagram is commutative:*

$$\begin{tikzcd} V_i \arrow[r, "\alpha_i"] \arrow[dr, "T_i"'] & \bigoplus_{i \in I} V_i \arrow[d, "\beta"] \\ & U \end{tikzcd}$$

While for *external direct product*, things are slightly different:

Proposition 3.4.3. *Let $\{V_i \mid i \in I\}$ be collection of vector spaces over $\mathbb{F}$. Consider the collection of all linear transformations $\psi_i : U \to V_i$, i.e.*

$$\mathcal{C}_{dp} := \{(U, \{\psi_i\}_{i\in I}) \mid \psi_i : U \to V_i\}.$$

Then

(a) *For all $i \in I$, let $\alpha_i : \prod_{i\in I} V_i \to V_i$ be defined by $\alpha_i((\cdots, \overbrace{v_i}^{\text{position } i}, \cdots)) := v_i$. Then*

$$\left(\prod_{i\in I} V_i,\ \{\alpha_i\}_{i\in I}\right) \in \mathcal{C}_{dp};$$

(b) *For $T_i : U \to V_i$ such that $(U, \{T_i\}_{i\in I}) \in \mathcal{C}_{dp}$, one has*

$$T_i = \alpha_i \circ \gamma \quad \text{with} \quad \gamma(u) := (\cdots, T_i(u), \cdots)$$

(note that the sum is finite by the definition of external direct sum).

In other words, there is a unique $\gamma : U \to \prod_{i\in I} V_i$ such that the following diagram is commutative:

$$\begin{tikzcd} U \arrow[d, "\gamma"'] \arrow[dr, "T_i"] & \\ \prod_{i\in I} V_i \arrow[r, "\alpha_i"'] & V_i \end{tikzcd}$$

3.5. **Rough Introduction to Category Theory.** Category theory is used to describe similarities between different branches of mathematics. These branches have something in common - (1) certain ‘objects’ with some specified structures, and (2) certain ‘maps’ between these objects preserving their structures. Examples include:

- (MAT 2040) n-vectors $\{\mathbb{R}^n\}_{n\in\mathbb{N}}$ and matrix transformations $\{A : \mathbb{R}^n \to \mathbb{R}^m\}$;
- (MAT 2042) Vectors spaces $\{V\}$ and linear transformations $\{T : V \to W\}$;
- (MAT 3004) Groups $\{G\}$ and homomorphisms $\{\phi : G \to H\}$;
- (MAT 4002) Topological spaces $\{X\}$ and continuous functions $\{f : X \to Y\}$;

Here is a formal definition:

Definition 3.5.1. *A (small) category $\mathcal{C}$ consists of the following three mathematical entities:*

- *A set of objects* $\mathrm{Obj}(\mathcal{C})$ *(or simply $\mathcal{C}$);*
- *For every $X, Y \in \mathrm{Obj}(\mathcal{C})$, a set of morphisms*

$$\mathrm{Hom}(X, Y) := \{f : X \to Y\}$$

- *For every $X, Y, Z \in \mathrm{Obj}(\mathcal{C})$, one can compose morphisms:*

$$\circ : \mathrm{Hom}(X, Y) \times \mathrm{Hom}(Y, Z) \to \mathrm{Hom}(X, Z)$$
$$(f, g) \mapsto g \circ f$$

satisfying

(1) *For all morphisms f, g and h, one has $(f \circ g) \circ h = f \circ (g \circ h)$.*
(2) *For any $X \in \mathrm{Obj}(\mathcal{C})$, there exists an identity element $1_X \in \mathrm{Hom}(X, X)$ such that*
$$1_X \circ f = f \qquad g \circ 1_X = g$$
for all $f \in \mathrm{Hom}(U, X)$ and $g \in \mathrm{Hom}(X, Y)$.

Example 3.5.2. *Here are some examples of categories:*

(1) $\mathcal{C}_{set} = \{$*all possible sets* $A\}$, $\mathrm{Hom}(A, B) = \{$*all functions* $f : A \to B\}$.
(2) $\mathcal{C}_{vec} = \{$*all vectors* $\mathbb{R}^n \mid n \in \mathbb{N}\}$, $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) = \{$*all matrix transformations* $A : \mathbb{R}^n \to \mathbb{R}^m\} = M_{m\times n}(\mathbb{R})$.
(3) $\mathcal{C}_{vs} = \{$*all vectors spaces* $V\}$, $\mathrm{Hom}(V, W) = \{$*all linear transformations* $T : V \to W\}$.

Definition 3.5.3. *Let $\mathcal{C}$ be a category. An* **initial object** *is an object $\mathcal{I} \in \mathrm{Obj}(\mathcal{C})$ such that for any $X \in \mathrm{Obj}(\mathcal{C})$, there exists exactly one*
$$i_x : \mathcal{I} \to X \qquad \textit{in } \mathrm{Hom}(\mathcal{I}, X).$$
A **terminal object** *is an object $\mathcal{T} \in \mathrm{Obj}(\mathcal{C})$ such that for any $X \in \mathrm{Obj}(\mathcal{C})$, there exists exactly one*
$$t_x : X \to \mathcal{T} \qquad \textit{in } \mathrm{Hom}(X, \mathcal{T}).$$

Example 3.5.4. *Consider $\mathcal{C}_{vs}$. Then $\mathcal{I} = \{\mathbf{0}\}$ is an initial object. Namely, for all $W \in \mathrm{Obj}(\mathcal{C}_{vs})$, there is only one possible linear transformation $i_W : \mathcal{I} \to W$ given by:*
$$i_W(\mathbf{0}) := \mathbf{0_W}$$
Similarly, $\mathcal{T} = \{\mathbf{0}\}$ is also a terminal object, since there is only one possible linear transformation $t_W : W \to \mathcal{T}$ given by:
$$i_W(w) := \mathbf{0} \qquad for\ all\ w \in W.$$

In general, initial objects and terminal objects may not exist in a category. But if it does, it possesses some very nice properties:

Remark 3.5.5. (1) *Suppose $\mathcal{I}$ is an initial object, then $i_{\mathcal{I}} : \mathcal{I} \to \mathcal{I}$ must be the identity map $i_{\mathcal{I}} = 1_{\mathcal{I}}$.*
(2) *Similarly, if $\mathcal{T}$ is a terminal object, then $t_{\mathcal{T}} : \mathcal{T} \to \mathcal{T}$ must be the identity map $t_{\mathcal{T}} = 1_{\mathcal{T}}$.*
(3) *Suppose $\mathcal{I} \in \mathrm{Obj}(\mathcal{C})$ is an initial object. Then for any $V, W \in \mathrm{Obj}(\mathcal{C})$ and $\beta \in \mathrm{Hom}(V, W)$, $\beta \circ i_V \in \mathrm{Hom}(\mathcal{I}, W)$. But there is only one element $i_W \in \mathrm{Hom}(\mathcal{I}, W)$ by the definition of initial object, so*
$$i_W = \beta \circ i_V,$$
i.e. the following diagram commutes for all $V, W, X \in \mathrm{Obj}(\mathcal{C})$:

$$\begin{tikzcd}
\mathcal{I} \arrow[r, "i_V"] \arrow[dr, "i_W"] \arrow[ddr, "i_X"'] & V \arrow[d, "\beta"] \\
& W \arrow[d, "\beta'"] \\
& X \\
& \vdots
\end{tikzcd}$$

(4) *Similarly, if* $\mathcal{T} \in \mathrm{Obj}(\mathcal{C})$ *is an initial object. Then for any* $V, W \in \mathrm{Obj}(\mathcal{C})$ *and* $\beta \in \mathrm{Hom}(V, W)$*,* $\iota_W \circ \gamma \in \mathrm{Hom}(V, \mathcal{T})$*. But there is only one element* $t_V \in \mathrm{Hom}(V, \mathcal{T})$ *by the definition of terminal object, so*

$$i_V = i_W \circ \gamma,$$

i.e. the following diagram commutes for all $U, V, W \in \mathrm{Obj}(\mathcal{C})$*:*

$$\begin{tikzcd}
\vdots \\
U \arrow[d, "\gamma'"'] \arrow[ddr, "t_U"] \\
V \arrow[d, "\gamma"'] \arrow[dr, "t_V"] \\
W \arrow[r, "t_W"'] & \mathcal{T}
\end{tikzcd}$$

Proposition 3.5.6. *Let* V *be a vector space over* $\mathbb{F}$*, and* $W \leq V$ *be a fixed vector subspace of* V*. Consider*

$$\mathcal{C}_{qs} := \{(X, \phi) \mid \phi : V \to X \quad \textit{such that} \quad W \leq \ker(T)\}$$

and for $(X, \phi), (Y, \psi) \in \mathcal{C}_{qs}$*,*

$$\mathrm{Hom}((X, \phi), (Y, \psi)) := \{\beta : X \to Y \mid \psi = \beta \circ \phi\}$$

Then

$$\mathcal{I}_{qs} = (V/W, \pi_W)$$

is an initial object in $\mathcal{C}_{qs}$*.*

Proof. By Proposition 3.4.1, for each $(Y, \psi) \in \mathcal{C}_{qs}$, there exists a unique $\beta : V/W \to Y$ such that $\psi = \beta \circ \pi_W$. So

$$\mathrm{Hom}((V/W, \pi_W), (Y, \psi)) := \{\beta\}$$

has exactly one element. $\square$

Similarly, one has

Proposition 3.5.7. *Let* $\{V_i \mid i \in I\}$ *be collection of vector spaces over* $\mathbb{F}$*. Consider*

$$\mathcal{C}_{ds} := \{(U, \{\phi_i\}_{i\in I}) \mid \phi_i : V_i \to U\}.$$

and for $(U, \{\phi_i\}_{i\in I}), (W, \{\phi'_i\}_{i\in I}) \in \mathrm{Obj}(\mathcal{C}_{ds})$*,*

$$\mathrm{Hom}((U, \{\phi_i\}_{i\in I}), (W, \{\phi'_i\}_{i\in I})) := \{\beta : U \to W \mid \phi_i = \beta \circ \phi'_i \textit{ for all } i\}$$

Then

$$\mathcal{I}_{ds} = \left(\sum_{i\in I} T_i(v_i), \{\alpha_i\}_{i\in I} \right)$$

is an initial object in $\mathcal{C}_{ds}$*.*

As for external direct product, one has

Proposition 3.5.8. *Let* $\{V_i \mid i \in I\}$ *be collection of vector spaces over* $\mathbb{F}$*. Consider*

$$\mathcal{C}_{dp} := \{(U, \{\psi_i\}_{i\in I}) \mid \psi_i : U \to V_i\}.$$

and for $(U, \{\psi_i\}_{i\in I}), (W, \{\psi'_i\}_{i\in I}) \in \mathrm{Obj}(\mathcal{C}_{ds})$*,*

$$\mathrm{Hom}((U, \{psi_i\}_{i\in I}), (W, \{\psi'_i\}_{i\in I})) := \{\gamma : U \to W \mid \psi_i = \psi'_i \circ \beta \textit{ for all } i\}$$

Then

$$\mathcal{I}_{dp} = \left(\prod_{i\in I} T_i(v_i), \{\alpha_i\}_{i\in I} \right)$$

is an terminal object in $\mathcal{C}_{dp}$*.*