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. *As for multiplication, $(R, \cdot)$ is not a group for $R = \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{R}[x]$ or $\mathbb{Z}_n$, since 0^{-1} does not exist in all cases.*
5. *$\mathbb{Z}\backslash\{0\}$ is still not a group, since 3^{-1} does not exist. But $(\mathbb{Q}\backslash\{0\}, \cdot)$ is a group.*
6. *$(GL_n(\mathbb{R}), \cdot)$ is a group under matrix multiplication.*

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 rings* 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* $1_R \cdot r = r \cdot 1_R = r$ *for all* $r \in R$*;*
- **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*

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

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

Example 1.3.2. (1) $\mathbb{Z}$ *is not a field since* $\frac{1}{5} = 5^{-1}$ *does not exist in* $\mathbb{Z}$*.*

(2) $\mathbb{Z}_6$ *is not a field for a slightly different reason -* $[5] \cdot [5] = [25] = [1]$ *in* $\mathbb{Z}_6$*, so* $[5]^{-1} = [5]$ *has an inverse. However,* $[2]^{-1}$ *does not exist.*

(3) $\mathbb{Q}$*,* $\mathbb{R}$*,* $\mathbb{C}$ *are fields.*

(4) *In general,* $\mathbb{Z}_p$ *is a field if and only if* p *is a prime number. Indeed, for any* $[a] \neq [0]$*, one has* $\gcd(a, p) = 1$ *and by Bezout's theorem, there exists integers* $r, s \in \mathbb{Z}$ *such that*

$$1 = ar + ps$$

and hence

$$[1] = [a] \cdot [r] + [p] \cdot [s] = [a] \cdot [r] + [0] \cdot [s] = [a] \cdot [r],$$

that is, $[a]^{-1} = [r]$*. In such a case,* $\mathrm{char}(\mathbb{Z}_p) = p$*.*

In the Homework set, you will construct a field $\mathbb{F}$ with 9 elements. In general, all $\mathbb{F}$ with finitely many elements must have

$$|\mathbb{F}| = p^r$$

for some prime number p, with $\mathrm{char}(\mathbb{F}) = p$.

2. Vector Spaces

2.1. **Definition of Vector Space.** As in MAT2042, here is our 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*

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

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*

$$\textit{for all } \alpha, \beta \in \mathbb{F} \textit{ and all } w_1, w_2 \in W,\ \alpha w_1 + \beta w_2 \in W.$$

Example 2.1.5.

(1) *(MAT2042) In* $V = \mathbb{F}^n$*, all vector subspaces are of the form* $W = \mathrm{Span}_{\mathbb{F}}\{v_1, \dots, v_k\}$ *for some vectors* $v_i \in \mathbb{F}^n$ *(By convention, let* $\mathrm{Span}_{\mathbb{F}}(\emptyset) = \{0\}$*).*

(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 $\alpha_1, \dots, \alpha_k \in \mathbb{F}$*,* $s_1, \dots, s_k \in S$ *and* $k \in \mathbb{N}$*. 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*:*

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

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

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

- 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*

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

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

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

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

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

Example 2.2.3.

(1) *For $V = \mathbb{F}^n$, the* **canonical basis** *$\{e_1, \dots, e_n\}$ is a basis of V.*
(2) *For $V = M_{m\times n}(\mathbb{F})$, $\{E_{ij} \mid 1 \le i \le m,\ 1 \le j \le n\}$ (here E_{ij} is the matrix with 1 on the (i,j)-entry and 0 on the other entries) 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*

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

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

$$(*) \qquad v = \alpha_1 v_1 + \cdots + \alpha_n v_n = \beta_1 v_1' + \cdots + \beta_m v_m'$$

for some *nonzero* $\alpha_i, \beta_j \in \mathbb{F}$ and $\{v_1, \dots v_n\}, \{v_1', \dots, v_m'\} \subseteq \mathcal{B}$ has no repeated elements, then one has

(1) $n = m$,
(2) after some reordering of the indices, $v_i = v_i'$ and $\alpha_i = \beta_i$ for all i.

($\Rightarrow$)If $\mathcal{B}$ is a basis of V, by definition, $v \in \mathrm{Span}(\mathcal{B})$ for all $v \in V$. As for uniqueness, suppose (*) 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

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

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

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

and

$$\alpha_{k+1} = \cdots = \alpha_n = \beta_{k+1} = \cdots = \beta_m = 0.$$

But the latter is impossible since we assume that all $\alpha_i, \beta_j \neq 0$, so one must have $m = n = k$, and

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

($\Leftarrow$)If all $v \in V$ can be expressed by elements in $\mathrm{Span}(\mathcal{B})$, then $\mathcal{B}$ 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 $\mathrm{Span}(\mathcal{B})$, a contradiction. Then $\mathcal{B}$ is both a spanning set and linearly independent, and hence forms 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. *(Basis Extension Theorem) 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.*

In this course, the **cardinality** of any set is equal to one of the following three possibilities:

- a finite number $n \in \mathbb{N}$;
- infinity ∞ (and countable);
- infinity ∞ (and **un**countable)

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).*

(5) $\dim(C^\infty(\mathbb{R})) = \infty$ *(and uncountable). Note that* $\mathcal{L} = \{e^{rx} \mid r \in \mathbb{R}\}$ *is a linearly independent set in* $C^\infty(\mathbb{R})$.

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:*

$$\sum_{i \in I} W_i := \{w_{i_1} + \cdots + w_{i_k} \mid w_{i_l} \in W_{i_l},\ \{i_1, \ldots, i_k\} \subseteq I,\ k \in \mathbb{N}\}$$

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 $V = \mathbb{R}^3$, $W_1 = \mathrm{Span}\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\right\}$ and $W_2 = \mathrm{Span}\left\{\begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$. 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.

More generally, for **two** vector subspaces $W_1, W_2 \leq V$, one can check from that above definition that

$$W_1 + W_2 = W_1 \oplus W_2$$

iff $W_1 \cap W_2 = \{\mathbf{0}\}$.

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. ☐

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 $f(i) \in V_i$ of f. For example, if $I = \{1, 2, \dots, n\}$ is a finite set, we can write:*

$$f \longleftrightarrow (\overbrace{f(1)}^{\in V_1}, \overbrace{f(2)}^{\in V_2}, \dots, \overbrace{f(n)}^{\in V_n})$$

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

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

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

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

- *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*

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

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 $(i_1, j_1), \ldots, (i_k, j_k)$ be such that $j_l \in J_{i_l}$ for all l, consider

$$\alpha_1 f_{i_1}^{j_1} + \cdots + \alpha_k f_{i_k}^{j_k} = \mathbf{0}.$$

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

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

Then one has

$$\begin{aligned} (\alpha_1 f_{i_1}^{j_1} + \cdots + \alpha_k f_{i_k}^{j_k})(\iota) &= \mathbf{0}(\iota) \\ \alpha_l f_\iota^{j_l}(\iota) + \alpha_{x_1} f_\iota^{j_{x_1}}(\iota) + \cdots + \alpha_{x_m} f_\iota^{j_{x_m}}(\iota) &= \mathbf{0}_{V_\iota} \\ \alpha_l b_\iota^{j_l} + \alpha_{x_1} b_\iota^{j_{x_1}} + \cdots + \alpha_{x_m} b_\iota^{j_{x_m}} &= \mathbf{0}_{V_\iota} \end{aligned}$$

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*

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

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

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

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

$$v \sim v' \iff v - v' \in W.$$

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

$$u + W = u' + W \quad \text{and} \quad v + W = v' + W$$

(for possibly $u \neq u'$ *and* $v \neq v'$*), one needs to show that*

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

To see so, note that

$$\begin{aligned} &(u + W) + (v + W) = (u' + W) + (v' + W) \\ \Leftrightarrow\ &(u + v) + W = (u' + v') + W \\ \Leftrightarrow\ &(u + v) - (u' + v') \in W \\ \Leftrightarrow\ &(u - u') + (v - v') \in W \end{aligned}$$

where the second $\Leftrightarrow$ *follow from Proposition 2.6.1. On the other hand, , since* $u+W = u' + W$ *and* $v + W = v' + W$ *one can apply Proposition 2.6.1 again to conclude that*

$$u - u',\ v - v' \in W$$

and hence $(u - u') + (v - v') \in W$ *since* $W \leq V$ *is a vector subspace.*

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*

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

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 the last statement of 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)**.** *Let V and W be vector spaces over $\mathbb{F}$. A* **linear transformation** *from V to W is a map $T : V \to W$ satisfying:*

$$T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2)$$

for all $\alpha, \beta \in \mathbb{F}$ and $v_1, v_2 \in V$.

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*

$$tr(M) := M_{11} + \cdots + M_{nn}$$

is a linear transformation.

(4) *(determinant is* **not** *linear) Let* $\mathrm{char}(\mathbb{F}) = 0$ *and* $\det : M_{n\times n}(\mathbb{F}) \to \mathbb{F}$ *given by*

$$\det(M) := \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} M_{i,\sigma(i)}$$

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. Then the following holds:*

(1) $T(\mathbf{0}_V) = \mathbf{0}_W$.

(2) *The composition $S \circ T : V \to U$ is also a linear transformation.*

(3) *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.*

(4) *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)**.** *Let $T : V \to W$ be a linear transformation.*
- *The* **kernel** *of T is defined by* $\ker(T) := \{v \in V \mid T(v) = \mathbf{0}_W\}$
- *The* **image** *of T is defined by* $im(T) := \{T(v) \in W \mid v \in V\}$

The following result should be well-known in MAT2042:

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

(1) $\ker(T) \le V$, $\operatorname{im}(T) \le W$.

(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.*

Conversely, if $T : V \to W$ *is an isomorphism, then* $\dim(V) = \dim(W)$*. More precisely, if* $\{b_i \mid i \in I\}$ *is a basis of* V*, then one can check (in Homework) that* $\{T(b_i) \mid i \in I\}$ *is a basis of* W*.*

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*

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

There is an $\dim(V) = \infty$ version of the above theorem. But for practical applications one usually take $\dim(V) < \infty$.

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 and is called **canonical projection***, 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. For the well-definedness of $\overline{T}$, suppose $v_1 + S = v_2 + S$, then $v_1 - v_2 = s \in S$. By Proposition 2.6.1, hence

$$\begin{gathered} T(v_1 - v_2) = T(s) = \mathbf{0}_W \\ T(v_1) = T(v_2) \\ \overline{T}(v_1 + S) = \overline{T}(v_2 + S). \end{gathered}$$

Therefore, $\overline{T}$ is well-defined.

Then we need to check that $\overline{T}$ is a linear transformation:

$$\begin{aligned} \overline{T}(\alpha(v+S) + \beta(u+S)) &= \overline{T}((\alpha v) + S + (\beta u) + S) \\ &= \overline{T}((\alpha v + \beta u) + S) \\ &= T(\alpha v + \beta u) \\ &= \alpha T(v) + \beta T(u) \\ &= \alpha \overline{T}(v+S) + \beta \overline{T}(u+S) \end{aligned}$$

for all $v, u \in V$ and $\alpha, \beta \in \mathbb{F}$. $\square$

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

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

Proof. By applying Proposition 3.2.2 with $S = \ker(T)$, one has a linear transformation $\overline{T} : V/\ker(T) \to \operatorname{im}(T)$ given by

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

To conclude the proof, we need to check that $\overline{T}$ is bijective:

- $\overline{T}$ is surjective:
  Surjectivity is obviously true.
- $\overline{T}$ is injective:
  In this case, one only needs to check $\ker(\overline{T}) = \{\mathbf{0}\}$. To see so,

  $$\overline{T}(v + \ker(T)) = \mathbf{0} \;\Leftrightarrow\; T(v) = \mathbf{0} \;\Leftrightarrow\; v \in \ker(T) \;\Leftrightarrow\; v + \ker(T) = \mathbf{0}_{V/\ker(T)}$$

$\square$

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

Proof. By Homework, one has

$$\dim(V/\ker(T)) = \dim(V) - \dim(\ker(T)).$$

Also, one has

$$\dim(V/\ker(T)) = \dim(\operatorname{im}(T))$$

by Example 3.1.8. So the result follows. $\square$

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 $\mathcal{B} = \{E_{ij} \mid 1 \le i, j \le n\}$. Then*

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

and hence the trace transformation in Example 3.1.2 satisfies $tr = f_{11} + \cdots + f_{nn}$.

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^*.*

Proof. Let $\{f_{n_i}\}_{1\le i\le n}$ be a finite subset of $\mathcal{B}^*$, and $\alpha_i \in \mathbb{F}$ be such that

$$\sum_{i=1}^{m} \alpha_i f_{n_i} = \mathbf{0} \quad \text{(where } \mathbf{0} \in V^* \text{ is the zero map).}$$

Then for each $1 \leq j \leq n$, we have

$$\sum_{i=1}^{m} \alpha_i f_{n_i}(b_{n_j}) = 0 \quad \text{(where 0 is the number zero)}$$

In other words,

$$\alpha_j = \sum_{i=1}^{m} \alpha_i \delta_{n_i,n_j} = 0.$$

Therefore we conclude that $\alpha_j = 0$ for each j, and the result follows. □

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.

3.4. **Annihilator of a Set.**

Definition 3.4.1. *Let $S \subset V$ be a subset. Then* **annihilator** *of S is*

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

Example 3.4.2.

(1) *Let $V = \mathbb{F}^3$, $\mathcal{B} = \{e_1, e_2, e_3\}$ be the canonical basis and $\mathcal{B}^* = \{f_1, f_2, f_3\}$ be the corresponding dual basis. For $S = \{\begin{pmatrix}1\\2\\0\end{pmatrix}\}$, one has*

$$f_3\begin{pmatrix}1\\2\\0\end{pmatrix} = 0 \qquad \text{and} \qquad (6f_1 - 3f_2)\begin{pmatrix}1\\2\\0\end{pmatrix} = 6\cdot 1 - 3\cdot 2 = 0,$$

hence $f_3, 6f_1 - 3f_2 \in \operatorname{Ann}(S)$.

(2) *Let $V = \mathbb{F}[x]$ and $S = \operatorname{Span}\{1, x\}$. Then $f \in V^*$ given by*

$$f(p) := p''(x)$$

is in $\operatorname{Ann}(S)$.

Proposition 3.4.3. *Let V be a vector space over $\mathbb{F}$.*

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

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

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

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

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

Lemma 3.4.4 (Complementation)**.** *Let V be a vector space over $\mathbb{F}$, and $W \le V$. Then there exists a complementary subspace $U \le V$ 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. □

Proposition 3.4.5. *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. □

Theorem 3.4.6. *Let V be a vector space over $\mathbb{F}$, and $W \leq V$. Then one has a natural 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 $g = T(f) \in \mathrm{im}(T)$.

$\square$

3.5. **Universal Properties.** In this section, we will put our constructions of new vector spaces (quotient space, direct sum, direct product) and new linear transformations from the old ones **under a panoramic perspective**.

The following proposition is just re-statement of Proposition 3.2.2:

Proposition 3.5.1. *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.*

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

Then

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

$$(V/W, \pi_W) \in \mathcal{C}_{qs}.$$

(b) *For* $T : V \to U$ *such that* $(U, T) \in \mathcal{C}_{qs}$*, there is a uniquely defined* $\beta : V/W \to U$ *satisfying*

$$\beta(v + W) := T(v) \quad \Rightarrow \quad T = \beta \circ \pi_W.$$

In other words, there is a unique β *such that following diagram commutes:*

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

Proof. Statement (a) is easy - since $W \leq \ker(\pi_W) = W$. As for (b), 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 with $S = W$, and one can take $\beta = \overline{T} : V/W \to U$ in the proposition with $\beta(v + W) = \beta \circ \pi_W(v) = T(v)$ to obtain the result. $\square$

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

Proposition 3.5.2. *Let* $\{V_i \mid i \in I\}$ *be collection of vector spaces over* $\mathbb{F}$*. Consider the collection of all linear transformations* $\phi_j : V_j \to X$*, i.e.*

$$\mathcal{C}_{ds} := \{(X, \{\phi_j\}_{j \in I}) \mid \phi_j : V_j \to X\}.$$

Then

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

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

(b) *For* $T_j : V_j \to U$ *such that* $(U, \{T_j\}_{j \in I}) \in \mathcal{C}_{ds}$*, there is a unique* $\beta : \bigoplus_{i \in I} V_i \to U$ *such that*

$$\beta((\cdots, v_j, \cdots)) := \textstyle\sum_{j \in I} T_j(v_j) \quad \Rightarrow \quad T_j = \beta \circ \iota_j \text{ for all } j \in I$$

(note that the sum $\sum_{j \in I} T_j(v_j)$ *is finite by the definition of external direct sum).*

In other words, there is a unique β *making the following diagram commute:*

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

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

Proposition 3.5.3. *Let $\{V_i \mid i \in I\}$ be collection of vector spaces over $\mathbb{F}$. Consider the collection of all linear transformations $\psi_j : X \to V_j$, i.e.*

$$\mathcal{C}_{dp} := \{(X, \{\psi_j\}_{j\in I}) \mid \psi_j : X \to V_j\}.$$

Then

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

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

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

$$\gamma(u) := (\cdots, T_i(u), \cdots) \quad \Rightarrow \quad T_j = \tau_j \circ \gamma \text{ for all } j \in I$$

In other words, there is a unique γ making the following diagram commute:

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

3.6. **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 3.6.1. *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 3.6.2. *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* $\mathbb{R}^n$ $|n \in \mathbb{N}\}$, $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) = \{$*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 3.6.3. *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*
$$i_x : \mathcal{I} \to X \qquad \text{in } \mathrm{Hom}(\mathcal{I}, X).$$
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*
$$t_x : X \to \mathcal{T} \qquad \text{in } \mathrm{Hom}(X, \mathcal{T}).$$

Example 3.6.4. *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:
$$i_W(w) := \mathbf{0} \qquad \text{for all } w \in W.$$

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

Remark 3.6.5. (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* $\gamma \in \mathrm{Hom}(V, W)$*,* $\iota_W \circ \gamma \in \mathrm{Hom}(V, \mathcal{T})$*. 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 3.6.6. *Let* V *be a vector space over* $\mathbb{F}$*, and* $W \leq V$ *be a fixed vector subspace of* V*. Consider the category*

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

with

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

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

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

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

Proof. By Proposition 3.5.1, for each $(U,T)\in\mathcal{C}_{qs}$, there exists a unique $\beta=\overline{T}:V/W\to U$ such that $T=\beta\circ\pi_W$. So

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

has exactly one element. $\square$

Similarly, one has

Proposition 3.6.7. *Let* $\{V_i\mid i\in I\}$ *be collection of vector spaces over* $\mathbb{F}$. *Consider the category*

$$\mathcal{C}_{ds}:=\{(X,\{\phi_j\}_{j\in I})\mid \phi_j:V_j\to X\}.$$

with

$$\mathrm{Hom}((X,\{\phi_j\}_{j\in I}),(Y,\{\phi'_j\}_{j\in I})):=\{\beta:X\to Y\mid \phi_j=\beta\circ\phi'_j \text{ for all } j\}$$

for $(X,\{\phi_j\}_{j\in I}),(Y,\{\phi'_j\}_{j\in I})\in\mathrm{Obj}(\mathcal{C}_{ds})$. *Then*

$$\mathcal{I}_{ds}=\left(\bigoplus_{i\in I}T_i(v_i),\{\iota_j\}_{j\in I}\right)$$

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

As for external direct product, one has

Proposition 3.6.8. *Let* $\{V_i\mid i\in I\}$ *be collection of vector spaces over* $\mathbb{F}$. *Consider the category*

$$\mathcal{C}_{dp}:=\{(X,\{\psi_j\}_{j\in I})\mid \psi_j:X\to V_j\}.$$

with

$$\mathrm{Hom}((X,\{\psi_j\}_{j\in I}),(Y,\{\psi'_j\}_{j\in I})):=\{\gamma:X\to Y\mid \psi_j=\psi'_j\circ\gamma \text{ for all } j\}$$

for $(X,\{\psi_j\}_{j\in I}),(Y,\{\psi'_j\}_{j\in I})\in\mathrm{Obj}(\mathcal{C}_{ds})$. *Then*

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

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

These observations seems very complicated, but they will become very useful in understand tensor products and linear transformations on tensor products in the next chapter.

4. Tensor Product and the Determinant

4.1. **Motivation.** We begin by giving some motivations on studying tensor products:

Definition 4.1.1. *Let V_1, V_2, W be vector spaces over $\mathbb{F}$. A* **bilinear map** *is a function*

$$f : V_1 \times V_2 \to W$$

satisfying

$$f(\alpha v_1 + \beta v_1', v_2) = \alpha f(v_1, v_2) + \beta f(v_1', v_2), \quad f(v_1, \alpha v_2 + \beta v_2') = \alpha f(v_1, v_2) + \beta f(v_1, v_2').$$

More generally, let $V_1, \ldots, V_k, W$ be vector spaces over $\mathbb{F}$. A **k-linear map** *is a function*

$$f : V_1 \times \cdots \times V_k \to W$$

satisfying

$$f(\cdots, v_{i-1}, \alpha v_i + \beta v_i', v_{i+1}, \cdots) = \alpha f(\cdots, v_{i-1}, v_i, v_{i+1}, \cdots) + \beta f(\cdots, v_{i-1}, v_i', v_{i+1}, \cdots).$$

for all $1 \le i \le k$.

Example 4.1.2.

(1) *The inner product of real vector spaces (e.g. dot product)*

$$\langle\ ,\ \rangle : V \times V \to \mathbb{R}$$

is bilinear.

(2) *The cross product of $\mathbb{R}^3$*

$$\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$$

is bilinear.

(3) *The determinant function:*

$$\det : \overbrace{\mathbb{F}^n \times \cdots \times \mathbb{F}^n}^{n \text{ copies}} \to \mathbb{F}$$

$$\det(v_1, \ldots, v_n) := \det((v_1 \mid \ \ldots \ \mid v_n))$$

is n-linear.

One would like to use linear algebra to study k-linear maps. However, they are almost always not a linear transformation if we treat $V_1 \times \cdots \times V_k = \prod_{i=1}^k V_i$ as a external direct product space: For instance, if $f : V_1 \times V_2 \to W$ is bilinear, then

$$f(3(v_1, v_2)) = f(3v_1, 3v_2) = 3f(v_1, 3v_2) = 3 \cdot 3f(v_1, v_2) \neq 3f(v_1, v_2).$$

Inspired by Section 3.6, one would like to make the following:

Definition 4.1.3 (Universal Property of Tensor Product)**.** *Let $V_1, \dots, V_k$ be vector spaces over $\mathbb{F}$. Consider the category*

$$\mathcal{C}_{tp} := \{(W, f) \mid f : V_1 \times \cdots \times V_k \to W \text{ is } k\text{-linear}\}$$

with

$$\mathrm{Hom}((W, f), (U, g)) := \{\beta : W \to U \mid g = \beta \circ f\}.$$

Then the **tensor product**

$$\mathcal{I}_{tp} := (V_1 \otimes \cdots \otimes V_k,\ \iota : V_1 \times \cdots \times V_k \to V_1 \otimes \cdots \otimes V_k)$$

(both terms to be determined later) is defined to be the initial object in $\mathcal{C}_{tp}$.

Remark 4.1.4. *Note that there is a* **"unique" initial object** *in any category $\mathcal{C}$ - more precisely, in the Homework set, you will prove that if $\mathcal{I}$ and $\mathcal{I}'$ are both initial objects, then there exists a unique isomorphism*

$$i_{\mathcal{I}'} : \mathcal{I} \xrightarrow{\cong} \mathcal{I}'$$

between them. In other words, the initial object is **unique up to unique isomorphism.**

By the property of initial object, we can study all k-linear maps

$$g : V_1 \times \cdots \times V_k \to U$$

as follows:

- $\mathrm{Hom}((V_1 \otimes \cdots \otimes V_k,\ \iota), (U, g)) = \{\beta\}$ has only one element; so
- $g = \beta \circ \iota$, where
$$\beta : V_1 \otimes \cdots \otimes V_k \to U$$
is a **linear transformation**

In other words,

$$\boxed{\text{understand } k\textbf{-linear map } g \leftrightarrow \text{understand } \textbf{linear transformation } \beta}$$

Question: how to construct $V_1 \otimes \cdots \otimes V_k$ and $\iota : V_1 \times \cdots \times V_k \to V_1 \otimes \cdots \otimes V_k$?

4.2. **Construction of Tensor Product Space.** For simplicity, we will only construct $V \otimes W$ in this section. The general case with k vector spaces $V_1 \otimes \cdots \otimes V_k$ can be constructed similarly.

Definition 4.2.1. *Let V, W be vector spaces over $\mathbb{F}$. Consider the set*

$$\mathcal{S} = \{(v, w) \mid v \in V,\ w \in W\},$$

then we define the vector space

$$\mathcal{X} = \mathrm{Span}(\mathcal{S}) := \{\sum_{i=1}^{k} \alpha_i (v_i, w_i) \mid \alpha_i \in \mathbb{F},\ (v_i, w_i) \in \mathcal{S},\ k \in \mathbb{N}\},$$

with the usual addition and scalar multiplication rule, so that $\mathcal{S}$ is a basis of $\mathcal{X}$.

Remark 4.2.2. *Note that we only consider $\mathcal{S}$ as a* **set**, *but not a vector space $V \times W$. In particular, there are no relations on the elements $(v, w) \in \mathcal{X}$. For instance $s := (v, w)$ and $s' := (2v, 2w)$ are two different elements in $\mathcal{S}$, so*

$$(2v, 2w) \neq 2(v, w)$$

in $\mathcal{X}$, since $s' \neq 2s$ for two linearly independent vectors in $\mathcal{X}$.

Similarly, $s_1 = (\mathbf{0}, w)$, $s_2 = (v, \mathbf{0})$ and $s_3 = (v, w)$ are three different elements in $\mathcal{S}$, so

$$(\mathbf{0}, w) + (v, \mathbf{0}) \neq (v, w)$$

in $\mathcal{X}$, since $s_1 + s_2 \neq s_3$ for three linearly independent vectors in $\mathcal{X}$.

The only legitimate rule in $\mathcal{X}$ is

$$3(v, w) + 2(v, w) = 5(v, w)$$

(and is **not** *equal to $(5v, 5w)$!).*

Definition 4.2.3. *Let $\mathcal{Y} \leq \mathcal{X}$ be a vector subspace spanned by vectors of the form*

$$\{1 \cdot (v_1 + v_2, w) - 1 \cdot (v_1, w) - 1 \cdot (v_2, w)\} \qquad \{1 \cdot (v, w_1 + w_2) - 1 \cdot (v, w_1) - 1 \cdot (v, w_2)\}$$

and

$$\{(\alpha v, w) - \alpha(v, w)\} \qquad \{(v, \alpha w) - \alpha(v, w)\}$$

for all $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$ and $\alpha \in \mathbb{F}$.

Then the **tensor product** *$V \times W$ is defined by the quotient space*

$$V \otimes W := \mathcal{X}/\mathcal{Y},$$

For $v \in V$, $w \in W$, we define $v \otimes w \in V \otimes W = \mathcal{X}/\mathcal{Y}$ by:

$$v \otimes w := (v, w) + \mathcal{Y}.$$

We now see some arithmetic in $V \otimes W$:

Example 4.2.4. *In $V \otimes W$, one has:*

$$\begin{aligned}
(v_1 + v_2) \otimes w &= (v_1 + v_2, w) + \mathcal{Y} \\
&= ((v_1 + v_2, w) - ((v_1 + v_2, w) - (v_1, w) - (v_2, w))) + \mathcal{Y} \\
&= ((v_1, w) + (v_2, w)) + \mathcal{Y} \\
&= [(v_1, w) + \mathcal{Y}] + [(v_2, w) + \mathcal{Y}] \\
&= v_1 \otimes w + v_2 \otimes w
\end{aligned}$$

Similarly, one has

$$v \otimes (w_1 + w_2) = (v \otimes w_1) + (v \otimes w_2)$$
$$(\alpha v) \otimes w = \alpha(v \otimes w)$$
$$v \otimes (\alpha w) = \alpha(v \otimes w)$$

Example 4.2.5. *Let $V = W = \mathbb{R}^2$. Then $\begin{pmatrix} 3 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} -4 \\ 2 \end{pmatrix}$ can be rewritten in term of e_1 and e_2:*

$$\begin{aligned}
\begin{pmatrix} 3 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} -4 \\ 2 \end{pmatrix} &= (3e_1 + 2e_2) \otimes (-4e_1 + 2e_2) \\
&= (3e_1) \otimes (-4e_1 + 2e_2) + (e_2) \otimes (-4e_1 + 2e_2) \\
&= (3e_1) \otimes (-4e_1) + (3e_1) \otimes (2e_2) + (e_2) \otimes (-4e_1) + e_2 \otimes (2e_2) \\
&= -12(e_1 \otimes e_1) + 6(e_1 \otimes e_2) - 4(e_2 \otimes e_1) + 2(e_2 \otimes e_2)
\end{aligned}$$

As an exercise, check that $e_1 \otimes e_2 + e_2 \otimes e_1$ cannot be re-written as $(ae_1 + be_2) \otimes (ce_1 + de_2)$ for any $a, b, c, d \in \mathbb{R}$.

Remark 4.2.6. *We mention some fundamental differences between product space $V \times W$ and the tensor product space $V \otimes W$:*

(1) $(v, \mathbf{0}) \neq \mathbf{0}_{V \times W}$ *in $V \times W$, but $v \otimes \mathbf{0} = \mathbf{0}_{V \otimes W}$:*
$$v \otimes \mathbf{0} = v \otimes (0 \cdot w) = 0(v \otimes w) = \mathbf{0}_{V \otimes W}$$

(2) $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$ *in $V \times W$, but $(v_1 \otimes w_1) + (v_2 \otimes w_2)$ cannot be simplified further in general for $V \otimes W$. Therefore, a general element in $V \otimes W = \mathcal{X}/\mathcal{Y}$ is of the form:*
$$(\alpha_1(v^{(1)}, w^{(1)}) + \cdots + \alpha_l(v^{(l)}, w^{(l)})) + \mathcal{Y} \quad \in \mathcal{X}/\mathcal{Y},$$
which is equal to
$$[(\alpha_1 v^{(1)}, w^{(1)}) + \mathcal{Y}] + \cdots + [(\alpha_l v^{(l)}, w^{(l)}) + \mathcal{Y}]$$
by the addition rule of quotient space and the calculations in Example 4.2.4 above. **So, a general element in $V \otimes W$ is:**
$$\widetilde{v}^{(1)} \otimes w^{(1)} + \cdots + \widetilde{v}^{(l)} \otimes w^{(l)}$$
where $\widetilde{v}^{(i)} = \alpha_i v^{(i)} \in V$ is an arbitrary vector.

Now we prove that our definition of $V \otimes W = \mathcal{X}/\mathcal{Y}$ satisfies the universal property:

Theorem 4.2.7. *Let $V \otimes W = \mathcal{X}/\mathcal{Y}$ be as defined above. Consider the map*
$$\iota : V \times W \to V \otimes W$$

defined by

$$\iota(v, w) := v \otimes w$$

Then ι is a bilinear map, and $(V \otimes W, \iota)$ is the initial object for $\mathcal{C}_{tp}$.

Proof. We firstly check that if ι is a bilinear map. By Example 4.2.4, one has

$$\begin{aligned}
\iota(v_1 + v_2, w) = (v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w = \iota(v_1, w) + \iota(v_2, w) \\
\iota(v, w_1 + w_2) = v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2 = \iota(v, w_1) + \iota(v, w_2) \\
\iota(\alpha v, w) = (\alpha v) \otimes w = \alpha(v \otimes w) = \alpha\iota(v, w) \\
\iota(v, \alpha w) = v \otimes (\alpha w) = \alpha(v \otimes w) = \alpha\iota(v, w).
\end{aligned}$$

Therefore, ι is a bilinear map.

To show that $(V \otimes W, \iota)$ is the initial object for $\mathcal{C}_{tp}$, one needs to show that for any **bilinear map**

$$f : V \times W \to U,$$

there is a unique **linear transformation**

$$\beta : V \otimes W \to U$$

such that $f = \beta \circ \iota$, that is

$$\beta(v \otimes w) = f(v, w).$$

Let $f : V \times W \to U$ be bilinear. Define a linear transformation

$$\Phi : \mathcal{X}(:= \mathrm{Span}(\mathcal{S})) \to U$$

given by

$$\Phi(v, w) := f(v, w)$$

for all $(v, w) \in \mathcal{S}$ (recall $\mathcal{S}$ is a basis of $\mathcal{X}$, and by Remark 3.1.5, we describe the linear transformation Φ by specifying its image under the basis of $\mathcal{X}$). Then

$$\begin{aligned}
\Phi((v_1 + v_2, w) - (v_1, w) - (v_2, w)) &= \Phi(v_1 + v_2, w) - \Phi(v_1, w) - \Phi(v_2, w) \\
&= f(v_1 + v_2, w) - f(v_1, w) - f(v_2, w) \\
&= 0
\end{aligned}$$

hence $(v_1 + v_2, w) - (v_1, w) - (v_2, w) \in \ker(\Phi)$. Similarly,

$$\begin{aligned}
(v, w_1 + w_2) - (v, w_1) - (v, w_2) &\in \ker(\Phi) \\
(\alpha v, w) - \alpha(v, w) &\in \ker(\Phi) \\
(v, \alpha w) - \alpha(v, w) &\in \ker(\Phi).
\end{aligned}$$

Therefore $\mathcal{Y} \leq \ker(\Phi)$. By Proposition 3.5.1, there is a unique linear transformation

$$\overline{\Phi} : \mathcal{X}/\mathcal{Y} \to U$$

such that

$$\overline{\Phi}((v,w)+\mathcal{Y}) = f(v,w).$$

Then one can take $\beta := \overline{\Phi} : V \otimes W \to U$ such that

$$\beta(v \otimes w) = \overline{\Phi}(v \otimes w) = f(v,w).$$

□

Corollary 4.2.8 (Universal Property of Tensor Product)**.** *Let $f : V \times W \to U$ be a bilinear map. Then there exists a unique linear transformation $\beta : V \otimes W \to U$ such that*

$$\beta(v \otimes w) = f(v,w)$$

for all $v \in V$ and $w \in W$.

In other words, β contains all information about f.

4.3. **Basis of Tensor Product.** Let V, W be finite-dimensional vector spaces over $\mathbb{F}$, with $\{v_1, \ldots, v_n\}$, $\{w_1, \ldots, w_m\}$ being bases of V and W. In this section, we will prove that

$$\mathcal{B} := \{v_i \otimes w_j | 1 \leq i \leq n, 1 \leq j \leq m\}$$

is a basis of $V \otimes W$.

Proposition 4.3.1. *$\mathcal{B}$ spans the tensor product space $V \otimes W$.*

Proof. Note that a general element in $V \otimes W$ is of the form $v^{(1)} \otimes w^{(1)} + \cdots + v^{(r)} \otimes w^{(r)}$. So it suffices to express each $v \otimes w := v^{(l)} \otimes w^{(l)}$ in terms of linear combinations of $v_i \otimes w_j$.

Suppose that $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ and $w = \beta_1 w_1 + \cdots + \beta_m w_m$. So

$$\begin{aligned} v \otimes w &= (\alpha_1 v_1 + \cdots + \alpha_n v_n) \otimes w \\ &= (\alpha_1 v_1) \otimes w + \cdots + (\alpha_n v_n) \otimes w \\ &= \alpha_1 (v_1 \otimes w) + \cdots + \alpha_n (v_n \otimes w) \end{aligned}$$

For each $v_i \otimes w$, $v_i \otimes w = \beta_1 (v_i \otimes w_1) + \cdots + \beta_m (v_i \otimes w_m)$. Therefore,

$$v \otimes w = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \beta_j (v_i \otimes w_j)$$

and the result follows. □

Theorem 4.3.2. *$\mathcal{B}$ is linearly independent in $V \otimes W$. So it is a basis of $V \otimes W$.*

Proof. Suppose

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\alpha_{ij}(v_i \otimes w_j) = \mathbf{0}$$

for some $\alpha_{ij} \in \mathbb{F}$, then we will use universal property to show that all α_{ij} are zeros.

Let $\{\phi_1, \ldots, \phi_n\}$ be the dual basis of V^*, and $\{\psi_1, \ldots, \psi_m\}$ be the dual basis of W^*. For $1 \leq p \leq n$ and $1 \leq q \leq m$, define

$$\pi_{p,q} : V \times W \to \mathbb{F}$$

by $\pi_{p,q}(v, w) := \phi_p(v)\psi_q(w)$. Then it is easy to check that $\pi_{p,q}$ is bilinear: for instance,

$$\begin{aligned}
\pi_{p,q}(\alpha v_1 + \beta v_2, w) &= \phi_p(\alpha v_1 + \beta v_2)\psi_q(w) \\
&= (\alpha\phi_p(v_1) + \beta\phi_p(v_2))\psi_q(w) \\
&= \alpha\phi_p(v_1)\psi_q(w) + \beta\phi_p(v_2)\psi_q(w) \\
&= \alpha\pi_{p,q}(v_1, w) + \beta\pi_{p,q}(v_2, w).
\end{aligned}$$

Therefore, $(\mathbb{F}, \pi_{p,q}) \in \mathcal{C}_{tp}$. By the universal property of the tensor product, there is a unique linear transformation

$$\Pi_{p,q} : V \otimes W \to \mathbb{F}$$

with $\pi_{p,q}(v, w) = \Pi_{p,q} \circ \iota(v, w)$. In other words,

$$\Pi_{p,q}(v \otimes w) = \pi_{p,q}(v, w) = \phi_p(v)\psi_q(w).$$

Applying the mapping $\Pi_{p,q}$ to

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\alpha_{ij}(v_i \otimes w_j) = \mathbf{0}$$

one has:

$$\begin{aligned}
\Pi_{p,q}\left(\sum_{i=1}^{n}\sum_{j=1}^{m}\alpha_{ij}(v_i\otimes w_j)\right) &= \Pi_{p,q}(\mathbf{0})\\
\sum_{i=1}^{n}\sum_{j=1}^{m}\alpha_{ij}\Pi_{p,q}\left((v_i\otimes w_j)\right) &= 0\\
\sum_{i=1}^{n}\sum_{j=1}^{m}\alpha_{ij}\pi_{p,q}(v_i,w_j) &= 0\\
\sum_{i=1}^{n}\sum_{j=1}^{m}\alpha_{ij}\phi_p(v_i)\psi_q(w_j) &= 0\\
\alpha_{pq} &= 0
\end{aligned}$$

for all p, q. So the result follows. □

Remark 4.3.3. *Note that the above proofs work perfectly well for infinite-dimensional vector spaces. Namely: Suppose $\{v_i \mid i \in I\}$ is a basis of V and $\{w_j \mid j \in J\}$ is a basis of W. Then*

$$\mathcal{B} := \{v_i \otimes w_j \mid i \in I,\ j \in J\}$$

is a basis of $V \otimes W$.

Corollary 4.3.4. *If $\dim(V), \dim(W) < \infty$, then $\dim(V \otimes W) = \dim(V)\dim(W)$.*

In the proof of the above theorem, one finds it very helpful to construct a linear transformation on $V \otimes W$ by a bilinear map on $V \times W$. We do one more example of such trick:

Theorem 4.3.5. *Let V, W be vector spaces over $\mathbb{F}$ (not necessarily finite dimensional). Then*

$$V \otimes W \cong W \otimes V.$$

Proof. Define

$$\phi : V \times W \to W \otimes V$$

by $\phi(v, w) := w \otimes v$. Then one can check that ϕ is bilinear map. So universal property implies that there exists a linear transformation

$$\Phi : V \otimes W \to W \otimes V$$

satisfying $\phi(v, w) = \Phi \circ \iota(v, w)$, i.e.

$$\Phi(v \otimes w) = w \otimes v.$$

Similarly, one can also construct a linear transformation

$$\Psi : W \otimes V \to V \otimes W$$

satisfying

$$\Psi(w \otimes v) = v \otimes w.$$

Then for a general element $v^{(1)} \otimes w^{(1)} + \cdots + v^{(r)} \otimes w^{(r)} \in W \otimes V$,

$$\begin{aligned}
\Psi \circ \Phi(v^{(1)} \otimes w^{(1)} + \cdots + v^{(r)} \otimes w^{(r)}) &= \Psi(\Phi(v^{(1)} \otimes w^{(1)}) + \cdots + \Phi(v^{(r)} \otimes w^{(r)})) \\
&= \Psi(w^{(1)} \otimes v^{(1)} + \cdots + w^{(r)} \otimes v^{(r)}) \\
&= \Psi(w^{(1)} \otimes v^{(1)}) + \cdots + \Psi(w^{(r)} \otimes v^{(r)}) \\
&= v^{(1)} \otimes w^{(1)} + \cdots + v^{(r)} \otimes w^{(r)}
\end{aligned}$$

So $\Psi \circ \Phi$ is the identity map. And similarly, $\Phi \circ \Psi$ is also an identity map. So they are inverses to each other, and the result follows. □

4.4. **Linear Transformation on Tensor Product Spaces.** As in the discussions on vector spaces in the last chapter, once we have constructed new vector spaces, we will construct new linear transformations.

Proposition 4.4.1. *Let V, V', W, W' be vector spaces over $\mathbb{F}$. Suppose that $T : V \to V'$ and $S : W \to W'$ are linear transformations, then there exists a unique linear transformation*

$$T \otimes S : V \otimes W \to V' \otimes W'$$

satisfying $(T \otimes S)(v \otimes w) = T(v) \otimes S(w)$.

Proof. Once again, we use universal property to help. Define the map

$$T \times S : V \times W \to V' \otimes W'$$

by $(T \times S)(v, w) := T(v) \otimes S(w)$. Then the map is bilinear, and as before there exists a unique linear transformation

$$T \otimes S : V \otimes W \to V' \otimes W'$$

satisfying $(T \otimes S)(v \otimes w) = T(v) \otimes S(w)$. □

Proposition 4.4.2. *The operation $T \otimes S$ satisfies all the properties of tensor product, i.e.*

$$(\alpha T_1 + \beta T_2) \otimes S = \alpha(T_1 \otimes S) + \beta(T_2 \otimes S) \quad T \otimes (\alpha S_1 + \beta S_2) = \alpha(T \otimes S_1) + \beta(T \otimes S_2)$$

Therefore, the notation "$\otimes$" in the definition of $T \otimes S$ is justified.

Proof. By Remark 4.3.3, $v_i \otimes w_j$ forms a basis of $V \otimes W$. Therefore, one only needs to check that both sides of the equation gives the same image upon applying $v_i \otimes w_j$. For instance,

$$\begin{aligned}((\alpha T_1 + \beta T_2) \otimes S)(v_i \otimes w_j) &= (\alpha T_1 + \beta T_2)(v_i) \otimes S(w_j) \\ &= (\alpha T_1(v_i) + \beta T_2(v_i)) \otimes S(w_j) \\ &= \alpha(T_1(v_i) \otimes S(w_j)) + \beta(T_2(v_i) \otimes S(w_j)) \\ &= \alpha(T_1 \otimes S)(v_i \otimes w_j) + \beta(T_2 \otimes S)(v_i \otimes w_j).\end{aligned}$$

□

Remark 4.4.3. *Let V, V', W, W' be vector spaces over $\mathbb{F}$, with*

- $\mathcal{A} := \{v_i \mid 1 \leq i \leq n\}$ *is a basis of* V.
- $\mathcal{A}' := \{v'_j \mid 1 \leq j \leq n'\}$ *is a basis of* V'.
- $\mathcal{B} := \{w_k \mid 1 \leq k \leq m\}$ *is a basis of* W.
- $\mathcal{B}' := \{w'_l \mid 1 \leq l \leq m'\}$ *is a basis of* W'.

Suppose $T : V \to V'$ has a matrix representation $T_{\mathcal{A}'\mathcal{A}} = A = (a_{ij})$, and $S : W \to W'$ has a matrix representation $T_{\mathcal{B}'\mathcal{B}} = B = (b_{kl})$.

Take

$$\begin{aligned}\mathcal{D} &:= \{v_1 \otimes w_1, \ldots, v_1 \otimes w_m, \quad \cdots, \quad v_n \otimes w_1, \ldots, v_n \otimes w_m\} \\ \mathcal{D}' &:= \{v'_1 \otimes w'_1, \ldots, v'_1 \otimes w'_{m'}, \quad \cdots, \quad v'_{n'} \otimes w'_1, \ldots, v'_{n'} \otimes w'_{m'}\}\end{aligned}$$

as **ordered** *bases of $V \otimes W$ and $V' \otimes W'$ respectively. Then the matrix representation of $T \otimes S : V \otimes W \to V' \otimes W'$ is the* **Kronecker product***:*

$$T_{\mathcal{D}'\mathcal{D}} = \begin{pmatrix} a_{11}B & a_{12}B & \ldots & a_{1n}B \\ a_{21}B & a_{22}B & \ldots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{n'1}B & a_{n'2}B & \ldots & a_{n'n}B \end{pmatrix}$$

We omit the proof here.

4.5. **Exterior Tensor Product.** Before going to the definition of exterior product, let us briefly mention some results on multilinear tensor product $V_1 \otimes \cdots \otimes V_k$ without proving them:

- **General vector:**

$$v_1^{(1)} \otimes \cdots \otimes v_k^{(1)} + \cdots + v_1^{(r)} \otimes \cdots \otimes v_k^{(r)}$$

- **k-linearity:**

$$\begin{aligned}& v_1 \otimes \cdots \otimes (\alpha v_i + \beta v'_i) \otimes \cdots \otimes v_k \\ = {}& \alpha(v_1 \otimes \cdots \otimes v_i \otimes \cdots \otimes v_k) + \beta(v_1 \otimes \cdots \otimes v'_i \otimes \cdots \otimes v_k)\end{aligned}$$

- **Dimension:**
$$\dim(V_1 \otimes \cdots \otimes V_k) = \dim(V_1) \times \cdots \times \dim(V_k)$$
- **Universal property:** Let
$$\iota : V_1 \times \cdots \times V_k \to V_1 \otimes \cdots \otimes V_k$$
$$\iota(v_1, \ldots, v_k) := v_1 \otimes \cdots \otimes v_k,$$
then for any *k-linear map* $f : V_1 \times \cdots \times V_k \to U$, there exists a unique *linear transformation*
$$\beta : V_1 \otimes \cdots \otimes V_k \to U$$
satisfying
$$\beta(v_1 \otimes \cdots \otimes v_k) = f(v_1, \ldots, v_k).$$
- **Linear transformation:** Let $T_i : V_i \to W_i$ be linear transformations. Then there exists a unique linear transformation
$$T_1 \otimes \cdots \otimes T_k : V_1 \otimes \cdots \otimes V_k \to W_1 \otimes \cdots \otimes W_k$$
satisfying
$$T_1 \otimes \cdots \otimes T_k(v_1 \otimes \cdots \otimes v_k) = T_1(v_1) \otimes \cdots \otimes T_k(v_k).$$

In the case when $V_1 = \cdots = V_k = V$, $W_1 = \cdots = W_k = W$ and $T_1 = \cdots = T_k = T$, we will use the shortand e
$$T^{\otimes k} : V^{\otimes k} \to W^{\otimes k}$$

Definition 4.5.1. *A k-linear map $f : V \times \cdots \times V \to W$ is called* **alternating** *if*
$$f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = \mathbf{0}_W$$
whenever $v_i = v_j$ for some $i \neq j$.

Example 4.5.2. *The cross product and determinant function in Example 4.1.2 are alternating, since $v \times v = \mathbf{0}$ and the determinant of a matrix with 2 identical columns is equal to* 0.

Lemma 4.5.3. *If $f : V \times \cdots \times V \to W$ is an alternating map, then*
$$f(v_1, \ldots, v, \ldots, w, \ldots, v_k) = -f(v_1, \ldots, w, \ldots, v, \ldots, v_k)$$
for all $v_i, v, w \in V$

Proof. Note that $f(v_1, \ldots, v + w, \ldots, v + w, \ldots, v_k) = 0$, hence one has
$$\begin{aligned} &f(v_1, \ldots, v, \ldots, v, \ldots, v_k) + f(v_1, \ldots, v, \ldots, w, \ldots, v_k) \\ &+ f(v_1, \ldots, w, \ldots, v, \ldots, v_k) + f(v_1, \ldots, w, \ldots, w, \ldots, v_k) = \mathbf{0} \end{aligned}$$

by k-linearlity of f. Note that the first and the last term are zero since f is alternating, so

$$f(v_1, \ldots, v, \ldots, w, \ldots, v_k) + f(v_1, \ldots, w, \ldots, v, \ldots, v_k) = 0$$

and the result follows. $\square$

Definition 4.5.4. *Let V be a vector space over $\mathbb{F}$. Consider the vector subspace $\mathcal{U} \leq V^{\otimes k}$ spanned by vectors of the form*

$$\{v_1 \otimes \cdots \otimes v \otimes \cdots \otimes v \otimes \cdots \otimes v_k\}$$

*for all $v, v_1, \ldots, v_k \in V$. Then the k-**exterior product** of V is the quotient space*

$$\wedge^k V := V^{\otimes k}/\mathcal{U}.$$

We write $v_1 \wedge \cdots \wedge v_k := v_1 \otimes \cdots \otimes v_k + \mathcal{U} \in \wedge^k V$, and

$$\iota : V \times \cdots \times V \to \wedge^k V$$

is defined by $\iota(v_1, \ldots, v_k) := v_1 \wedge \cdots \wedge v_k$.

By definition, one can check that

- $\wedge^k V$ is k-linear:

$$\begin{aligned} &v_1 \wedge \cdots \wedge (\alpha v_i + \beta v_i') \wedge \cdots \wedge v_k \\ =\ &\alpha(v_1 \wedge \cdots \wedge v_i \wedge \cdots \wedge v_k) + \beta(v_1 \wedge \cdots \wedge v_i' \wedge \cdots \wedge v_k) \end{aligned}$$

  for $i = 1, \ldots, k$.
- $\wedge^k V$ is alternating:

$$\begin{aligned} v_1 \wedge \cdots \wedge v \wedge \cdots \wedge v \wedge \cdots \wedge v_k &:= v_1 \otimes \cdots \otimes v \otimes \cdots \otimes v \otimes \cdots \otimes v_k + \mathcal{U} \\ &= \mathbf{0} + \mathcal{U} \\ &= \mathbf{0}_{\wedge^k V} \end{aligned}$$

- $\wedge^k V$ changes sign by swapping two entries:

$$v_1 \wedge \cdots \wedge v \wedge \cdots \wedge w \wedge \cdots \wedge v_k = -(v_1 \wedge \cdots \wedge w \wedge \cdots \wedge v \wedge \cdots \wedge v_k)$$

Theorem 4.5.5 (Universal Property of Exterior Product)**.** *Let V be a vector space over $\mathbb{F}$. Consider the category*

$$\mathcal{C}_{ep} := \{(W, f) \mid f : V \times \cdots \times V \to W \text{ is } k\text{-alternating}\}$$

with

$$\mathrm{Hom}((W, f), (U, g)) := \{\beta : W \to U \text{ linear transformation} \mid g = \beta \circ f\}.$$

Then the k-exterior product $\mathcal{I}_{ep} := (\wedge^k V,\ \iota : V \times \cdots \times V \to \wedge^k V)$ is the initial object in $\mathcal{C}_{ep}$.

*In other words, for all k-**alternating maps** $f : V \times \cdots \times V \to U$, there exists a* **linear transformation**

$$\beta : \wedge^k V \to U$$

such that $f = \beta \circ \iota$, i.e. $\beta(v_1 \wedge \cdots \wedge v_k) = f(v_1, \ldots, v_k)$.

Example 4.5.6. *Let $T : V \to W$ be a linear transformation, one can define*

$$f : V \times \cdots \times V \to \wedge^k W$$

by $f(v_1, \ldots, v_k) := T(v_1) \wedge \cdots \wedge T(v_k)$. Then one can easily check that f is k-alternating, and the universal property of exterior product implies that there exists

$$T^{\wedge k} := \beta : \wedge^k V \to \wedge^k W$$

satisfying $T^{\wedge k}(v_1 \wedge \cdots \wedge w_k) = T(v_1) \wedge \cdots \wedge T(v_k)$.

4.6. **The Determinant.** We are now in the position to define the determinant of a linear operator $T : V \to V$ for a vector space V with $\dim(V) = n < \infty$.

Lemma 4.6.1. *Let V be a vector space over $\mathbb{F}$ with $\dim(V) = n$. For $0 \leq k \leq n$, one has*

$$\dim(\wedge^k V) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Proof. Suppose $\{v_1, \ldots, v_n\}$ is a basis of V, we claim that

$$\mathcal{E} := \{v_{i_1} \wedge \cdots \wedge v_{i_k} \mid 1 \leq i_1 < \cdots < i_k \leq n\}$$

is a basis of $\wedge^k V$.

It is easy to see $\mathcal{E}$ spans $\wedge^k V$: Namely, recall the canonical projection

$$\pi_{\mathcal{U}} : V^{\otimes k} \to \wedge^k V = V^{\otimes k}/\mathcal{U}$$

given by

$$\pi_{\mathcal{U}}(v^{(1)} \otimes \cdots \otimes v^{(k)}) = (v^{(1)} \otimes \cdots \otimes v^{(k)}) + \mathcal{U} = v^{(1)} \wedge \cdots \wedge v^{(k)}$$

is surjective, so it maps a spanning set of $V^{\otimes k}$ to a spanning set of $\wedge^k V$. Since

$$\{v_{j_1} \otimes \cdots \otimes v_{j_k} \mid 1 \leq j_l \leq n \text{ for all } l\}$$

is a basis (and hence a spanning set) of $V^{\otimes k}$,

$$\{v_{j_1} \wedge \cdots \wedge v_{j_k} \mid 1 \leq j_l \leq n \text{ for all } l\}$$

is a spanning set of $\wedge^k V$. But it is easy to see that each element in the above set is either zero (if there are repeated terms), or (if there are no repeated terms) it is equal to an element in $\mathcal{E}$ up to a sign.

The proof of linear independence is left as an exercise - for instance, one can follow the proof of linear independence in the tensor product section. $\square$

Therefore, in the particular case when $k = n$, $\dim(\wedge^n V) = 1$, and is spanned by

$$\zeta := v_1 \wedge \cdots \wedge v_n$$

for any choice of basis $\{v_1, \ldots, v_n\}$ of V.

Definition 4.6.2. *Let V be a finite dimensional vector space over $\mathbb{F}$ with $\dim(V) = n$, and $T : V \to V$ be a linear operator.*

Then $\wedge^n V = \mathrm{Span}(\zeta)$ is one-dimensional, and the linear operator

$$T^{\wedge n} : \wedge^n V \to \wedge^n V$$

defined by

$$T^{\wedge n}(u_1 \wedge \cdots \wedge u_n) := T(u_1) \wedge \cdots \wedge T(u_n)$$

satisfies

$$T^{\wedge n}(\zeta) = \Delta_T \cdot \zeta$$

for some scalar Δ_T. We define the **determinant** *of T by*

$$\det(T) := \Delta_T.$$

Example 4.6.3. *Let $T : \mathbb{F}^2 \to \mathbb{F}^2$ be given by $T\begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$. Then $\wedge^2 \mathbb{F}^2 = \mathrm{Span}(e_1 \wedge e_2)$, and*

$$T^2 : \wedge^2 \mathbb{F}^2 \to \wedge^2 \mathbb{F}^2$$

satisfies

$$\begin{aligned}
T^{\wedge 2}(e_1 \wedge e_2) &= T(e_1) \wedge T(e_2) \\
&= \begin{pmatrix} a \\ c \end{pmatrix} \wedge \begin{pmatrix} b \\ d \end{pmatrix} \\
&= (ae_1 + ce_2) \wedge (be_1 + de_2) \\
&= (ae_1) \wedge (be_1) + (ae_1) \wedge (de_2) + (ce_2) \wedge (de_1) + (ce_2) \wedge (de_2) \\
&= 0 + (ad)e_1 \wedge e_2 + (bc)e_2 \wedge e_1 + 0 \\
&= (ad - bc)e_1 \wedge e_2
\end{aligned}$$

Therefore, $\det(T) = ad - bc$ as expected.

More generally, suppose Let $T : \mathbb{F}^n \to \mathbb{F}^n$ be given by

$$T(v) := Av, \quad \textit{where} \quad A = \big(a_1 | a_2 | \ldots | a_n\big)$$

By writing $\det(A) := \det(T)$ as in MAT2042, one has

$$T^{\wedge n}(e_1 \wedge e_2 \wedge \cdots \wedge e_n) = a_1 \wedge a_2 \wedge \cdots \wedge a_n = \det(A)(e_1 \wedge e_2 \wedge \cdots \wedge e_n)$$

So if A have identical columns $a_i = a_j$ for $i \neq j$, one has $a_1 \wedge a_2 \wedge \cdots \wedge a_n = \mathbf{0}$ and hence $\det(A) = 0$.

Also, if $A' = \big(a_2|a_1|\ldots|a_n\big)$ and $T'(v) := A'v$. Then

$$\begin{aligned}\det(A')(e_1 \wedge e_2 \wedge \cdots \wedge e_n) &= (T')^{\wedge n}(e_1 \wedge e_2 \wedge \cdots \wedge e_n)\\ &= a_2 \wedge a_1 \wedge \cdots \wedge a_n\\ &= -(a_1 \wedge a_2 \wedge \cdots \wedge a_n)\\ &= -\det(A)(e_1 \wedge e_2 \wedge \cdots \wedge e_n)\end{aligned}$$

and hence $\det(A') = -\det(A)$.

By similar arguments, one can conclude the following defining properties of the determinant:

Corollary 4.6.4. *Let $A \in M_{n\times n}(\mathbb{F})$. Then the following holds:*

- $\det(I_{n\times n}) = 1$*;*
- $\det\big(a_1|\ldots|a_i|\ldots|a_j|\ldots|a_n\big) = -\det\big(a_1|\ldots|a_j|\ldots|a_i|\ldots|a_n\big)$*;*
- *For any $\alpha \in \mathbb{F}$,* $\det\big(a_1|\ldots|\alpha a_i|\ldots|a_n\big) = \alpha\det\big(a_1|\ldots|a_i|\ldots|a_n\big)$*;*

As another advantage of understanding determinant in such a way, we have

Theorem 4.6.5. *Let $S, T : V \to V$ be linear operators. Then*

$$\det(S \circ T) = \det(S)\det(T)$$

Proof. Pick any basis $\{v_1 \wedge \cdots \wedge v_n\}$ of $\wedge^n V$, then

$$\begin{aligned}\det(T \circ S)(v_1 \wedge \cdots \wedge v_n) &= (T \circ S)^{\wedge n}(v_1 \wedge \cdots \wedge v_n)\\ &= (T \circ S)(v_1) \wedge \cdots \wedge (T \circ S)(v_n)\\ &= T(S(v_1)) \wedge \cdots \wedge T(S(v_n))\\ &= (T^{\wedge n})((S(v_1) \wedge \cdots \wedge S(v_n))\\ &= (T^{\wedge n}) \circ (S^{\wedge n})(v_1 \wedge \cdots \wedge v_n)\\ &= (T^{\wedge n})(\det(S)(v_1 \wedge \cdots \wedge v_n))\\ &= \det(S)T^{\wedge n}(v_1 \wedge \cdots \wedge v_n)\\ &= \det(S)\det(T)(v_1 \wedge \cdots \wedge v_n)\end{aligned}$$

and hence $\det(T \circ S) = \det(T)\det(S)$. □

Corollary 4.6.6. *Let $A, B \in M_{n\times n}(\mathbb{F})$, then*

$$\det(AB) = \det(A)\det(B).$$

5. Modules

5.1. **Basic Definitions.** Modules can be seen as a generalization of "vector spaces over a ring R". The theory can get very general for different kinds of R's, so we would like to focus on a specific kind of rings:

Definition 5.1.1. *Let R be a unital commutative ring. A* **zerodivisor** *of R is a nonzero element $a \in R \backslash \{0\}$ such that there exists $b \in R \backslash \{0\}$ such that*

$$a \cdot b = 0.$$

If R has no zerodivisors, we call R an **integral domain (ID)**.

Example 5.1.2. *Here are some (non)-examples of integral domains:*

(1) *Let $R = \mathbb{Z}_6$, then $[2] \cdot [3] = [0]$, so $[2]$ and $[3]$ are zero-divisors. In other words, R is* **not** *an ID.*
(2) *$\mathbb{Z}$ is an ID.*
(3) *All fields $\mathbb{F}$ are IDs - namely, for all nonzero $a \in \mathbb{F} \backslash \{0\}$, one always has a^{-1} such that $a^{-1} \cdot a = 1$. Therefore,*

$$a \cdot b = 0 \quad \Rightarrow \quad a^{-1} \cdot a \cdot b = a^{-1} \cdot 0 \quad \Rightarrow \quad 1 \cdot b = 0 \quad \Rightarrow \quad b = 0$$

which implies there are no $b \neq 0$ such that $a \cdot b = 0$.
(4) *If R is an ID, then the polynomial ring $R[x]$ is also an ID.*

Unless specified otherwise, we will focus on the case when R is an integral domain.

Definition 5.1.3. *Let $(R, +, \cdot)$ be an integral domain.* **A (left) R-module** *is a set M with operations*

$$+ : M \times M \to M \quad \text{and} \quad \cdot : R \times M \to M$$

such that the following holds:

- *$(M, +)$ is an abelian group, i.e.*
  (1) *$(m_1 + m_2) + m_3 = m_1 + (m_2 + m_3)$.*
  (2) *There exists $\mathbf{0}_M$ such that $\mathbf{0}_M + m = m = m + \mathbf{0}_M$.*
  (3) *For every $m \in M$, there exists $(-m) \in M$ such that $m + (-m) = (-m) + m = \mathbf{0}_M$.*
  (4) *$m_1 + m_2 = m_2 + m_1$ for any $m_1, m_2 \in M$.*
- *For every $r, s \in R$, $u, v \in M$, we have*
  (1) *$r \cdot (u + v) = r \cdot u + r \cdot v$.*
  (2) *$(r + s) \cdot u = r \cdot u + s \cdot u$.*
  (3) *$(r \cdot s) \cdot u = r \cdot (s \cdot u)$.*
  (4) *$1 \cdot u = u$.*

Example 5.1.4. (1) *Let $R = \mathbb{F}$, then all vector spaces V over $\mathbb{F}$ is an R-module.*

(2) *For any R, $R^n = \{(r_1, \cdots, r_n) \mid r_i \in R\}$ with*
- $(r_1, \cdots, r_n) + (s_1, \cdots, s_n) := (r_1 + s_1, \cdots, r_n + s_n)$;
- $r \cdot (r_1, \cdots, r_n) := (r \cdot r_1, \cdots, r \cdot r_n)$

is an R-module.

(3) *Let $R = \mathbb{Z}$, then $M = \mathbb{Z}_n$ with*
- $[a] + [b] := [a + b]$;
- $r \cdot [a] := [r \cdot a]$

is an R-module.

(4) *(Important Example) Let V be a vector space over $\mathbb{F}$ and $T : V \to V$ is a linear operator. Then V is an $\mathbb{F}[x] = R$-module by*
- $v + w := v + w$ *(usual vector addition);*
- $(a_0 + a_1 x + \cdots + a_n x^n) \cdot v := (a_0 I + a_1 T + \cdots + a_n T^n) \cdot v = a_0 v + a_1 T(v) + \cdots + a_n T^n(v)$

for any $v, w \in V$ and $p(x) = a_0 + \cdots + a_n x^n \in R$.

By Cayley-Hamilton Theorem in MAT2042, the characteristic polynomial $\chi_T(x) \in \mathbb{F}[x]$ satisfies
$$\chi_T(T)v = \mathbf{0}$$
for all $v \in V$. In other words, one has $\chi_T(x) \cdot v = \mathbf{0}$ for any $v \in V$.

Remark 5.1.5. *In linear algebra, a single nonzero vector $\{v\}$ is always linear independent. This is not true anymore for modules! For instance, in Example (3) above,*
$$n \cdot [a] = [0]$$
so $\{[a]\}$ is not linearly independent. Also, in Example (4) above,
$$\chi_T(x) \cdot v = \mathbf{0},$$
so $\{v\}$ is not linearly independent.

5.2. **Submodules and Ideals.**

Definition 5.2.1. *Let $(M, +, \cdot)$ be an R-module. A subset $I \subseteq M$ is called a* **submodule** *of M if*
$$+|_{I \times I} : I \times I \to I \quad \text{and} \quad \cdot|_{R \times I} : R \times I \to I.$$
In other words, $(I, +|_{I\times I}, \cdot|_{R\times I})$ is a submodule of M if
$$i + j \in I \quad \text{and} \quad r \cdot I \in I$$
for all $i, j \in I$ and $r \in R$.

Example 5.2.2. *For any R, let $M = R^n$, then*
$$I := \{(r_1, r_2, \cdots, r_k, 0, \cdots, 0) \mid r_i \in R\}$$
is a submodule of M.

In the special case when one take $M = R$ as an R-module, we have the following definition for submodules of R, which is of *utmost importance* in both module theory and abstract algebra.

Definition 5.2.3 (Ideals)**.** *Let R be an integral domain, and $M = R^1 = R$ is an R-module. The submodules $I \lhd M(= R)$ are called* **ideals** *of R. In other words, $I \lhd R$ if and only if*

$$i + j \in I \quad \text{and} \quad r \cdot i \in I$$

for any $i, j \in I$ and $r \in R$.

Example 5.2.4. *Let $R = \mathbb{Z} = M$, then*

$$\begin{aligned} I_1 &= \langle\langle 2 \rangle\rangle = \{2 \cdot k \mid k \in \mathbb{Z}\} = 2\mathbb{Z} \lhd \mathbb{Z}, \\ I_2 &= \langle\langle 6, 9 \rangle\rangle = \{6 \cdot a + 9 \cdot b \mid a, b \in \mathbb{Z}\} = 3\mathbb{Z} = \langle\langle 3 \rangle\rangle \lhd \mathbb{Z}. \end{aligned}$$

5.3. **Spanning Set and Linear Independence.**

Definition 5.3.1 (Spanning Set)**.** *Let M be an R-module and $B \subseteq M$ is a subset. Then the* **submodule generated by** *B is*

$$\langle\langle B \rangle\rangle (= Span_R(B)) := \{r_1 \cdot b_1 + \cdots + r_n \cdot b_n \mid b_i \in B, r_i \in R, n \in \mathbb{N}\}.$$

More precisely,

- *If $|B| = n < \infty$, we say $\langle\langle B \rangle\rangle$ is* **n-generated***, or* **finitely generated***.*
- *If $B = \{b\}$, then $\langle\langle b \rangle\rangle$ is* **1-generated***, or is the* **cyclic submodule** *generated by b.*

Example 5.3.2. (1) *Let $R = \mathbb{Z}$, $M = \mathbb{Z}_n$. Then $M = \langle\langle 1 \rangle\rangle$ is 1-generated (cyclic).*

(2) *For any R, $M = R^n$ is generated by $M = \langle\langle e_1, \cdots, e_n \rangle\rangle$, where*

$$e_i := (0, \ldots, 0, \overbrace{1}^{i\text{-th coordinate}}, 0, \ldots, 0).$$

(3) *Let $R = \mathbb{Z}[x]$, then $M = R$ is 1-generated since $M = \langle\langle 1 \rangle\rangle$. However,*

$$\begin{aligned} M' &:= \{p(x_1, x_2, \cdots) \mid p(x_1, x_2, \cdots) \text{ has an even constant term}\} \\ &= \langle\langle 2, x \rangle\rangle \end{aligned}$$

is a submodule of M with 2 generators which cannot be reduced to

$$\langle\langle 2, x \rangle\rangle = \langle\langle f \rangle\rangle.$$

So even if M is n-generated, $M' \leq M$ may have more than n generators!

(4) *Let* $R = \mathbb{Z}[x_1, x_2, \cdots]$, *then* $M = R$ *is 1-generated. Consider*

$$\begin{aligned} M' &:= \{p(x_1, x_2, \cdots) \mid p(0, 0, \cdots) = 0 \text{ has zero constant term}\} \\ &= \langle\langle \{x_1, x_2, \cdots\} \rangle\rangle \end{aligned}$$

which is **NOT** *finitely generated. So even if* M *is finitely generated,* $M' \leq M$ *may not be finitely generated!*

Definition 5.3.3 (Principal Ideal Domain)**.**

(1) *An ideal* $I \lhd R$ *is called* **principal** *if* $I = \langle\langle a \rangle\rangle$ *is 1-generated.*
(2) *An integral domain* R *is called* **Principal Ideal Domain (PID)** *if all ideals* $I \lhd R$ *are 1-generated.*

Example 5.3.4. (1) $R = \mathbb{Z}$ *is a PID, and* $I \lhd \mathbb{Z}$ *are of the form* $I = n\mathbb{Z}$.
(2) $R = \mathbb{F}[x]$ *is a PID.*
(3) $R = \mathbb{Z}[x]$ *is* **NOT** *a PID by Example 5.3.2(3).*

For the later part of this course, we will study the special case when R is a Principal Ideal Domain in full detail. In particular, we will get (generalizations of) Primary Decomposition Theorem and Jordan Normal Form Theorem in the setting of R-modules with R being a PID.

Definition 5.3.5 (Linear Independence)**.** *Let* M *be an* R*-module. A subset* $S \subseteq M$ *is* **linearly independent** *if for any subset* $\{s_1, \cdots, s_k\} \subseteq S$, *one has*

$$r_1 s_1 + \cdots + r_k s_k = 0 \quad \Leftrightarrow \quad r_1 = \cdots = r_k = 0.$$

Otherwise, we say S *is* **linearly dependent**.

Example 5.3.6. *Linear dependence for modules and vector spaces are* **different**.

(1) *(A nonzero vector* $\{v\}$ *can be linearly dependent.)*
Let $R = \mathbb{Z}$ *and* $M = \mathbb{Z}_n$, *then* $[1] \in M$ *is linearly dependent since*

$$n \cdot [1] = [0].$$

(2) *(*$\{v_1, v_2\}$ *is linearly dependent does not implies* v_1 *is a scalar multiple of* v_2*.)*
Let $R = \mathbb{Z}$ *and* $M = R = \mathbb{Z}$, *then* $\{2, 3\}$ *is linearly dependent since*

$$6 \cdot 2 + (-4) \cdot 3 = 0,$$

but $2 \neq r \cdot 3$ *for any* $r \in \mathbb{Z}$.

5.4. **Torsion and Annihilators.** As we saw in Example 5.3.6(1), "bad" behavior happens if there exists $r \neq 0$ such that

$$r \cdot m = 0$$

for some $m \neq 0$. So we make the following:

Definition 5.4.1. *Let M be an R-module. The set*

$$M_{tor} := \{m \in M \mid r \cdot m = 0 \text{ for some } r \in R \setminus \{0\}\}$$

is called the **torsion elements** *in M.*

- *If $M_{tor} = \{0\}$, we say M is* **torsion-free**.
- *If $M_{tor} = M$, we say M is a* **torsion module**.

Example 5.4.2.

(1) *Let $R = Z$ and $M = \mathbb{Z} \oplus \mathbb{Z}_3 = \{(a, [b]) \mid a \in \mathbb{Z}, [b] \in \mathbb{Z}_3\}$. Then*

$$\begin{aligned} M_{tor} &= \{(a, [b]) \mid (r \cdot a, r \cdot [b]) = (0, [0]) \text{ for some nonzero } r \in \mathbb{Z}\} \\ &= \{(0, [b]) \mid [b] \in \mathbb{Z}_3\}. \end{aligned}$$

(2) *Recall the important example in Example 5.1.4(4), where V is an $R = \mathbb{F}[x]$-module given by*

$$p(x) \cdot v := p(T) \cdot v$$

Since $\chi_T(x) \cdot v = \chi_T(T) \cdot v = 0$ for all $v \in V$, so one can take $r = \chi_T(x)$ in the definition of torsion elements to conclude that

$$V_{tor} = V$$

is a torsion module.

Proposition 5.4.3. *Let R be an ID, and M be an R-module. Then $M_{tor} \leq M$ is a submodule.*

Proof. Let $t_1, t_2 \in M_{tor}$, in other words, there exists $r_1, r_2 \neq 0$ such that

$$r_1 \cdot t_1 = r_2 \cdot t_2 = 0.$$

Since R is ID, $r_1, r_2 \neq 0$ implies $r_1 \cdot r_2 \neq 0$. Then

$$\begin{aligned} r_1 r_2 \cdot (p \cdot t_1 + q \cdot t_2) &= r_1 r_2 p \cdot t_1 + r_1 r_2 q \cdot t_2 \\ &= r_2 p \cdot (r_1 \cdot t_1) + r_1 q \cdot (r_2 \cdot t_2) \\ &= r_2 p \cdot 0 + r_1 q \cdot 0 \\ &= 0 \end{aligned}$$

for any $p, q \in R$, and we have $p \cdot t_1 + q \cdot t_2 \in M_{tor}$. Hence $M_{tor} \leq M$. □

In the definition of torsion elements, we look at which $m \in M$ gets 'killed' by some $r \in R$. We now change our perspective by looking at which $r \in R$ 'kills' an element $m \in M$:

Definition 5.4.4 (Annihilator)**.** *Let M be an R-module.*

- *The* **annihilator** *of $m \in M$ is*

$$\mathrm{Ann}(m) := \{r \in R \mid r \cdot m = 0\}.$$

- *The* **annihilator** *of M is*

$$\mathrm{Ann}(M) := \{r \in R \mid r \cdot m = 0 \textit{ for all } m \in M\}.$$

Example 5.4.5. (1) *If M is torsion-free, then $\mathrm{Ann}(M) = \{0_R\}$ for all $m \in M$.*
(2) *On the other extreme, if $M = M_{tor}$ is a torsion module, then $Ann(m) \neq \{0_R\}$ for all $m \in M$. However, this* **does not** *necessarily imply $Ann(M) \neq \{0_R\}$ in general!*

Example 5.4.6 (Minimal Polynomials for Linear Operators)**.** *As a special case of torsion module, consider our important Example 5.1.4(2) with $R = \mathbb{F}[x]$ and $M = V$. We have already seen from there that $\chi_T(x) \in Ann(v)$ for all $v \in V$. So one has*

$$\chi_T(x) \in Ann(V).$$

More generally, for any linear operator $T : V \to V$, define the **minimal polynomial** $m_T(x)$ **of** T *such that*

(1) *$m_T(x)$ is monic, i.e. the leading power coefficient of $m_T(x)$ is 1.*
(2) *$m_T(T)v = \mathbf{0}$ for all $v \in V$.*
(3) *$m_T(x)$ is the polynomial of smallest positive such that (1) and (2) holds.*

A possible candidate satisfying (1) and (2) is the characteristic polynomial $\chi_T(x)$. However, there may be polynomials with smaller degree such that both (1) and (2) holds. In general, one has:

$$\boxed{\textit{If } f(x) \in \mathbb{F}[x] \textit{ satisfies } f(T)v = \mathbf{0}\ \forall\ v \in V\ \textit{(e.g. } f(x) = \chi_T(x)\textit{), then } m_T(x) | f(x)}$$

Under this perspective, one has

$$\begin{aligned} Ann(V) &:= \{f(x) \in \mathbb{F}[x] \mid f(x) \cdot v = \mathbf{0} \textit{ for all } v \in V\} \\ &= \{f(x) \mid m_T(x) | f(x)\} \\ &= \{f(x) = m_T(x)p(x) \mid p(x) \in \mathbb{F}[x]\} \\ &= \langle\langle m_T(x) \rangle\rangle \end{aligned}$$

Proposition 5.4.7. *Let M be a R-module, and $m \in M$. Then $\mathrm{Ann}(m) \lhd R$ is an ideal.*

Proof. Let $a, a' \in \mathrm{Ann}(m)$, one has $a \cdot m = a' \cdot m = \mathbf{0}$. Then

$$(a + a') \cdot m = a \cdot m + a' \cdot m = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

$$(ra) \cdot m = r \cdot (a \cdot m) = r \cdot \mathbf{0} = \mathbf{0}$$

for all $r \in R$, hence $a + a' \in \mathrm{Ann}(m)$ and $r \cdot a \in \mathrm{Ann}(m)$. □

The same proposition and the same argument hold for $\mathrm{Ann}(M)$

5.5. **Basis and Free Modules.**

Definition 5.5.1. *Let M be an R-module, We say M is a* **free module** *if there exists a linearly independent spanning set $\mathcal{B}$ of M.*

Example 5.5.2.

(1) *(Not all M has a basis) Let $R = \mathbb{Z}$, then $M = \mathbb{Z}_n$ is not free since every element in M is a torsion element.*
(2) *For any ring R, $M = R^n$ is free with $\mathcal{B} = \{e_1, e_2, ..., e_n\}$.*
(3) *(M is free module does* **NOT** *imply $N \leq M$ is a free module) Let $R = \mathbb{Z} \times \mathbb{Z}$ (with addition and multiplication defined in the usual way), then $M = R$ is a free module by (2). However, the submodule*
$$N := \{(a, 0) | a \in \mathbb{Z}\} \leq M$$
is no longer free. (Since $(0, 1) \cdot (a, 0) = (0, 0)$, no single element in N is linearly independent.)

Free modules (over R) behave as nicely as vector space (over $\mathbb{F}$). For instance, we have:

Theorem 5.5.3. *Let M be a free R-module with basis $\mathcal{B}$, then*

(1) *Every $m \in M$ can be* **uniquely** *expressed as*
$$m = r_1 b_1 + r_2 b_2 + \cdots + r_k b_k$$
where $r_i \in R$, $b_i \in \mathcal{B}$, $k \in \mathbb{N}$.
(2) *$\mathcal{B}$ is a minimal spanning set of M (i.e, if you remove an element $b \in B$, $\mathcal{B} - \{b\}$ is* **NOT** *a spanning set).*
(3) *$\mathcal{B}$ is a maximal linearly independent set of M (i.e. If we add $m \in M$ to $\mathcal{B}$, then $\mathcal{B} \cup \{m\}$ is linear dependence).*

A natural question to ask is that whether all bases of a free module M has the same **cardinality**. The answer of this question is yes, as we will see below. The proof requires the use of quotient modules, which is given by:

Definition 5.5.4 (Quotient Module)**.** *Let M be an R-module, and $N \leq M$ is a submodule. The* **quotient module** *M/N is an R-module*
$$M/N := \{m + M \mid m \in M\},$$
with $+$ and $\cdot$ defined by
$$(m_1 + N) + (m_2 + N) := (m_1 + m_2) + N$$
$$r \cdot (m + N) := (r \cdot m) + N$$
for all $r \in R$.

Example 5.5.5. *Let* $R = \mathbb{Z}$, $M = \mathbb{Z}$ *and* $N := \langle\langle n\rangle\rangle = n\mathbb{Z}$. *Then*

$$\begin{aligned} M/N &= \mathbb{Z}/n\mathbb{Z} \\ &= \{a + n\mathbb{Z} \mid a \in \mathbb{Z}\} \\ &= \{0 + n\mathbb{Z}, 1 + n\mathbb{Z}, \cdots, (n-1) + n\mathbb{Z}\} \\ &= \{\overline{0}, \overline{1}, \cdots, \overline{n-1}\} \end{aligned}$$

with

$$\overline{2} + \overline{3} = (2 + n\mathbb{Z}) + (3 + n\mathbb{Z}) = (2+3) + n\mathbb{Z} = \overline{2+3}$$
$$4 \cdot \overline{3} = \overline{4 \cdot 3} = \overline{12}.$$

So $M/N = \{\overline{0}, \overline{1}, \cdots, \overline{n-1}\}$ *behaves like* $\mathbb{Z}_n = \{[0], [1], \cdots, [n-1]\}$ *(more precisely, we will say* $\mathbb{Z}/n\mathbb{Z}$ *and* $\mathbb{Z}_n$ *are* **isomorphic** *as* $\mathbb{Z}$*-modules later). Therefore,*

$$n \cdot \overline{a} = \overline{n \cdot a} = \overline{0}$$

for all $\overline{a} \in M/N$, *and hence* $\{\overline{a}\}$ *is linearly independent. So* M/N *has no basis, i.e.* M/N *is* **NOT** *free.*

In conclusion,

$$M = \langle\langle 1\rangle\rangle \text{ is free with } \mathcal{B}_M = \{1\}$$
$$N = \langle\langle n\rangle\rangle \text{ is free with } \mathcal{B}_N = \{n\}$$

while M/N *is not free.*

Theorem 5.5.6. *Let* M *be a free* R*-module. Then any bases of* M *have the same cardinality.*

To prove the theorem, we need to some facts from abstract algebra: Let R be an integral domain, then

(1) For any ideal $I \lhd R$, the quotient $(R/I, +, \cdot)$ is also a ring (called **quoitent ring**) with

$$(r + I) + (s + I) := (r + s) + I$$
$$(r + I) \cdot (s + I) := (r \cdot s) + I$$

(2) In the special case when $I \lhd R$ is a **maximal** ideal (here maximal means if

$$I \lhd N \lhd R,$$

then $N = I$ or $N = R$). Then the quotient ring $\mathbb{F} = R/I$ is a field.

Proof. Let $I \lhd R$ be a maximal ideal, and consider

$$IM := \{i_1m_1 + i_2m_2 + \cdots + i_km_k | i_j \in I, m_k \in M, k \in \mathbb{N}\}$$

It is easy to check that IM is a submodule of M. So we have a quotient R-module

$$M/IM := \{m + IM | m \in M\}$$

In fact, M/IM is an $\mathbb{F} = R/I$-module with scalar multiplication defined by

$$(r+I)\dot{(}m+IM) := rm + IM$$

(check this is well-defined). In other words, M/IM **is a $\mathbb{F}$-vector space**.

Now we make the following claim:

Claim. *If*

$$\mathcal{B} = \{b_s \mid s \in S\} \qquad \mathcal{C} = \{c_j \mid j \in J\}$$

are two bases of M (as R-modules), then

$$\mathcal{B}' = \{b_s + IM \mid s \in S\} \qquad \mathcal{C}' = \{c_j + IM \mid j \in J\}$$

are two bases of M/IM (as $\mathbb{F} = R/I$-vector spaces).

If the claim holds, then $|\mathcal{B}'| = |\mathcal{C}'|$ have the same cardinality by our knowledge of bases over vector spaces. Then the result follows since $|\mathcal{B}| = |\mathcal{B}'|$ and $|\mathcal{C}| = |\mathcal{C}'|$.

Proof of claim.

- $\mathcal{B}'$ is a **spanning set**.
  Take any $m + IM \in M/IM$, since $\langle\langle \mathcal{B} \rangle\rangle = M$

$$\begin{aligned} m + IM &= (r_1 b_1 + \cdots + r_n b_n) + IM \\ &= (r_1 + I)(b_1 + IM) + \cdots + (r_n + I)(b_n + IM) \end{aligned}$$

- $\mathcal{B}'$ is **linearly independent**.
  Suppose $(r_1 + I)(b_1 + IM) + \cdots + (r_n + I)(b_n + IM) = \mathbf{0}$

$$\Rightarrow r_1 b_1 + \cdots + r_n b_n + IM = \mathbf{0}$$

$$\Rightarrow r_1 b_1 + \cdots + r_n b_n \in IM$$

$$\Rightarrow r_1 b_1 + \cdots + r_n b_n = i_1 m_1 + \cdots + i_k m_k \in IM$$

$$\Rightarrow r_1 b_1 + \cdots + r_n b_n = i_1 (r_1^{(1)} b_1 + \cdots + r_1^{(n)} b_n) + \cdots + i_k (r_k^{(1)} b_1 + \cdots + r_k^{(n)} b_n)$$

$$\Rightarrow r_1 b_1 + \cdots + r_n b_n = (i_1 r_1^{(1)} + \cdots + i_k r_k^{(1)}) b_1 + \cdots + (i_1 r_1^{(n)} + \cdots i_k r_k^{(n)}) b_n$$

By the fact that $\mathcal{B}$ is linear independent, one has

$$\begin{aligned} r_1 &= r_1^{(1)} i_1 + \cdots + r_k^{(1)} i_k \\ r_2 &= r_1^{(2)} i_1 + \cdots + r_k^{(2)} i_k \\ &\vdots \\ r_n &= r_1^{(1)} i_1 + \cdots + r_k^{(n)} i_k \end{aligned}$$

But $i_1, \ldots, i_k \in I$ and $I \lhd R$ is an R-module, so each summand $r_\star^{(\bullet)} i_* \in I$ in the above equations. Hence we conclude that $r_1, r_2, \cdots, r_n \in I$, that is

$$r_1 + I = r_2 + I = \cdots = r_n + I = 0.$$

$\square$

Now we are safe to make the following:

Definition 5.5.7. *Let M be a free R-module, The* **rank** *of M is the cardinality of any choice of basis $\mathcal{B}$ of M.*

5.6. **Homomorphisms.** Homomorphism can be seen as the analog of linear transformation for R-modules:

Definition 5.6.1 (Homomorphism)**.** *Let M, N be R-modules. A map $\phi : M \to N$ is an R-***homomorphism** *if*

$$\phi(r_1 \cdot m_1 + r_2 \cdot m_2) = r_1 \cdot \phi(m_1) + r_2 \cdot \phi(m_2)$$

for all $r_1, r_2 \in R$ and $m_1, m_2 \in M$.

The set of all R-homomorphism is denoted as

$$\mathrm{Hom}_R(M, N) := \{\phi : M \to N \mid \phi \textit{ is an } R\textit{-homomorphism}\},$$

which is an analog of $\mathcal{L}(V, W)$ for modules. And as in the case of $\mathcal{L}(V, W)$ for vector spaces, $\mathrm{Hom}_R(M, N)$ is an R-module given by:

$$(\phi + \psi)(m) := \phi(m) + \psi(m) \qquad \textit{and} \qquad (r \cdot \phi)(m) := r \cdot \phi(m).$$

Theorem 5.6.2. *Let M be a free R-module with basis $\mathcal{B} = \{b_i | i \in I\}$, then for any R-homomorphism $\phi : M \to N$, ϕ is uniquely determined by*

$$\{\phi(b_i) \in N \mid b_i \in \mathcal{B}\}.$$

Conversely, let $\{n_i \in N \mid i \in I\}$ be a collection of elements in N. Then there is a (unique) R-homomorphism $\phi : M \to N$ satisfying

$$\phi(b_i) = n_1.$$

More precisely, ϕ is defined by

$$\phi(r_1 \cdot b_{i_1} + \cdots + r_k \cdot b_{i_k}) := r_1 \cdot n_{i_1} + \cdots + r_k \cdot n_{i_k},$$

for all $r_l \in R$ and $b_{i_l} \in \mathcal{B}$.

Definition 5.6.3 (Kernel and Image)**.** *Let $\phi : M \to N$ be an R-homomorphism, then the* **kernel** *and* **image** *of ϕ are defined by:*

(1) $\ker(\phi) := \{m \in M \mid \phi(m) = 0_N\} \leq M$,
(2) $\mathrm{im}(\phi) := \{\phi(m) \in N \mid m \in M\} \leq N$.

Theorem 5.6.4. *Let $\phi \in \mathrm{Hom}_R(M, N)$,*

(1) ϕ *is* **injective** $\Leftrightarrow \ker(\phi) = \{0_M\}$,
(2) ϕ *is* **surjective** $\Leftrightarrow \mathrm{im}(\phi) = N$.

Definition 5.6.5. *If $\phi \in \mathrm{Hom}_R(M, N)$ is bijective, we say ϕ is an* **isomorphism** *between M and N.*

As in the case of vector spaces, we have the **correspondence** theorem and **isomorphism** theorems for R-modules:

Theorem 5.6.6 (Correspondence Theorem)**.** *Let $N \leq M$ be R-modules. Then there is a 1-1 correspondence between*

$$\begin{array}{rcl} \{S \mid N \leq S \leq M\} & \leftrightarrow & \{X \mid 0 \leq X \leq M/N\} \\ S & \mapsto & S/N \\ \displaystyle\bigcup_{x \in X} (x + N) & \leftarrow\!\shortmid & X \end{array}$$

Theorem 5.6.7 (Isomorphism Theorem)**.** *Let $\phi : M \to N$ be an R-homomorphism. Then the map $\overline{\phi} : M/\ker(\phi) \to \mathrm{im}(\phi)$ defined by*

$$\overline{\phi}(m + \ker(\phi)) := \phi(m)$$

is an R-isomorphism between $M/\ker(\phi)$ and $\mathrm{im}(\phi)$.

Similarly one can construct new R-modules from old ones using direct sum, direct product, tensor product as in the case of vector spaces.

6. Noetherian Rings and Noetherian Modules

6.1. **Basic Definitions.** In this section, we will specialize our attention to a certain kind of integral domains R. In particular, we want to **avoid** the following situation:

Example 6.1.1. *Let $R = \mathbb{F}[x_1, x_2, \cdots]$. Consider the R-module*

$$M = R = \langle\langle 1 \rangle\rangle$$

which is **1-generated**. *But the submodule $N \leq M$ with zero constant term*

$$N = \langle\langle \{x_1, x_2, \cdots\} \rangle\rangle$$

is **not finitely generated**.

In other words, we want the following holds: If M is finitely generated, then any $N \leq M$ is also finitely generated.

Definition 6.1.2 (Noetherian Module)**.** *Let M be an R-module. We say M is a* **noetherian module** *if it satisfies the ascending chain condition (ACC) of submodules: Let*

$$N_1 \leq N_2 \leq N_3 \leq \cdots$$

be an ascending chain of submodules of M. Then the chain must become equal somewhere, i.e. there must be some k such that

$$N_{k-1} < N_k = N_{k+1} = N_{k+2} = \cdots .$$

Example 6.1.3. *Let $R = \mathbb{F}[x_1, x_2, \cdots]$ and $M = R = \langle\langle 1 \rangle\rangle$. Then*

$$\langle\langle x_1 \rangle\rangle < \langle\langle x_1, x_2 \rangle\rangle < \langle\langle x_1, x_2, x_3 \rangle\rangle < \cdots$$

does not satisfy (ACC)! So $M = R$ is **NOT** *noetherian.*