MT242P ABSTRACT ALGEBRA

1. Fields

Definition 1.1. *A* **field** *is a non-empty set* $\mathbb{F}$ *together with two operations* $+ : \mathbb{F}\times\mathbb{F} \to \mathbb{F} : (a, b) \to a+b$, *called* **addition**, *and* $\cdot : \mathbb{F}\times\mathbb{F} \to \mathbb{F} : (a, b) \to a\cdot b$, *called* **multiplication**, *such that for all* $a, b, c \in \mathbb{F}$:

- *(F1) (commutativity)* $a + b = b + a$ *and* $a \cdot b = b \cdot a$
- *(F2) (associativity)* $a + (b + c) = (a + b) + c$ *and* $a \cdot (b \cdot c) = (a \cdot b) \cdot c$
- *(F3) (existence of additive and multiplicative identity elements) there are two distinct elements* 0, *called* **zero element**, *and* 1, *called* **one element**, *such that* $a + 0 = a = 0 + a$ *and* $a \cdot 1 = a = 1 \cdot a$.
- *(F4) (existence of additive and multiplicative inverses) there is* $x \in \mathbb{F}$ *such that* $a + x = 0 = x + a$. *We write* $-a$ *for this* x. *If* $a \neq 0$ *there is* $y \in \mathbb{F}$ *such that* $a \cdot y = 1 = y \cdot a$. *We write* a^{-1} *for this* y.
- *(F5) (distributive laws)* $a \cdot (b + c) = a \cdot b + a \cdot c$ *and* $(a + b) \cdot c = a \cdot c + b \cdot c$

Remark/Example 1.2. *(1) For convenience we may write* ab *instead of* $a\cdot b$. *Furthermore we set* $\mathbb{F}^* := \mathbb{F}\backslash\{0\}$.

(2) The zero and one element in a field are unique. Assume for instance that there is $z \in \mathbb{F}$ *such that* $a + z = a$, *for all* $a \in \mathbb{F}$. *Then in particular* $0 + z = 0$. *As furthermore* $0 + z = z$, *by (F3), we get* $z = 0$.

(3) The additive and multiplicative inverses are unique. Given $a \in \mathbb{F}$, *assume for instance besides* $-a$ *there is another additive inverses* $b \in \mathbb{F}$. *Then*

$$b \overset{(F3)}{=} b + 0 = b + (a + (-a)) \overset{(F2)}{=} (b + a) + (-a) = 0 + (-a) \overset{(F3)}{=} -a.$$

(4) Let $a, b, c \in \mathbb{F}$. *Then*

- *(i) If* $a + b = a + c$, *then* $b = c$.
- *(ii)* $a \cdot 0 = 0$
- *(iii)* $-(a \cdot b) = (-a) \cdot b = a \cdot (-b)$
- *(iv) If* $a \cdot b = 0$, *then* $a = 0$ *or* $b = 0$.

Proof. (i) $b \overset{(F3)}{=} 0+b \overset{(F4)}{=} ((-a)+a)+b \overset{(F2)}{=} (-a)+(a+b) = (-a)+(a+c) \overset{(F2)}{=} ((-a) + a) + c \overset{(F4)}{=} 0 + c \overset{(F3)}{=} c$

(ii) It follows from (i), as $(a\cdot 0)+0 \overset{(F3)}{=} a\cdot 0 \overset{(F3)}{=} a\cdot(0+0) \overset{(F5)}{=} (a\cdot 0)+(a\cdot 0)$

(iii) It follows, as $a \cdot b + a \cdot (-b) \overset{(F5)}{=} a \cdot (b + (-b)) \overset{(F4)}{=} a \cdot 0 \overset{(ii)}{=} 0$

(iv) If $b \neq 0$, then b^{-1} exists by (F4) and

$$0 \overset{(ii)}{=} 0 \cdot b^{-1} = (a \cdot b) \cdot b^{-1} \overset{(F2)}{=} a \cdot (b \cdot b^{-1}) \overset{(F4)}{=} a \cdot 1 \overset{(F3)}{=} a.$$

$\square$

(5) *Note that every field is a ring (see Definition 1.1 in Finite Mathematics). However for instance the ring of integers is not a field as there is no multiplicative inverse.*

(6) *Recall the rational numbers $\mathbb{Q} = \{(a,b) : a, b \in \mathbb{Z}, b \neq 0\}$, subject to the identity $(a,b) = (c,d)$ if and only if $ad = bc$, together with the operations*

$$(a,b) + (c,d) = (ad + bc, bd), \qquad (a,b) \cdot (c,d) = (ac, bd),$$

for all $a, b, c, d \in \mathbb{Q}$. Then $(\mathbb{Q}, +, \cdot)$ is a field with additive identity $(0,1)$ and the multiplicative identity is $(1,1)$. Furthermore for all $(a,b) \in \mathbb{Q}$ we have $-(a,b) = (-a,b)$ and, provided $a \neq 0$, $(a,b)^{-1} = (b,a)$.

(7) *The real numbers $\mathbb{R}$ form a field together with standard addition and multiplication.*

(8) *Recall the complex numbers $\mathbb{C} = \{a + bi : \ a, b \in \mathbb{R}\}$, where $i^2 = -1$, together with the two operations*

$$z + w := (a + c) + (b + d)i, \quad z \cdot w := (ac - bd) + (ad + bc)i,$$

where $z = a + bi$ and $w = c + di$. Then $(\mathbb{C}, +, \cdot)$ is a field with additive identity $0 + 0i$ and the multiplicative identity is $1 + 0i$. Furthermore for all $a + bi \in \mathbb{C}$ we have $-(a + bi) = (-a) + (-b)i$ and, provided $a + bi \neq 0$, $(a + bi)^{-1} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2} \cdot i$.

(9) *Let $n \in \mathbb{N}\backslash\{0\}$ and set $\mathbb{Z}_n := \{0, 1, \ldots, n-1\}$. Recall that by Lemma 6.3 of Finite Mathematics, for every $a \in \mathbb{Z}$ there is a unique $r_a \in \mathbb{Z}_n$ such that $a \equiv r_a \mod n$, where r_a is the residue of a modulo n. Hence for all $a, b \in \mathbb{Z}_n$ we can define*

$$a + b := r_{a+b} \qquad \textit{and} \qquad a \cdot b := r_{a \cdot b}$$

Then $(\mathbb{Z}_n, +, \cdot)$ is a ring, called **ring of integers modulo** *n, with 0 and 1 as the respective identity elements. Generally $\mathbb{Z}_n$ is not a field. In $\mathbb{Z}_4$, for instance, we have $2 \cdot 2 = 4 = 0$, contradicting property (5(iv)) above. Alternatively, one can check that 2 has no multiplicative inverse.*

Next assume that $n = p$ is a prime number. Then for every $a \in \{1, \ldots, p-1\}$ we have that $\gcd(a,p) = 1$ and so there are $s, t \in \mathbb{Z}$ such that $as + pt = 1$. Since $s \equiv r_s \mod p$, we have $ar_s \equiv 1 \mod p$ and so $a \cdot r_s = 1$ in $\mathbb{Z}_n$, that is, a has a multiplicative inverse. In fact one can show that $(\mathbb{Z}_n, +, \cdot)$ is a field if and only if n is prime.

For any prime number p one defines $\mathbb{F}_p := \mathbb{Z}_p$. Those $\mathbb{F}_p$ are examples of finite fields. In particular, $\mathbb{F}_2$ is the smallest possible field. We have

$$\begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \qquad \begin{array}{c|cc} \cdot & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

One can show that a finite field with n elements exists if and only if $n = p^r$, for some prime number p and integer $r \geq 1$. Take for instance the quadratic polynomial $f(x) = x^2 + x + 1$ over $\mathbb{F}_2$. As $f(0) = 1 = f(1)$, it follows that f has no roots in $\mathbb{F}_2$. If we define a new element α as a root of f, that is, $\alpha^2 + \alpha + 1 = 0$, then $\alpha + 1$ is also a root of f. Now $\mathbb{F}_4 := \{0, 1, \alpha, \alpha + 1\}$ is a field with

$$\begin{array}{c|cccc} + & 0 & 1 & \alpha & \alpha+1 \\ \hline 0 & 0 & 1 & \alpha & \alpha+1 \\ 1 & 1 & 0 & \alpha+1 & \alpha \\ \alpha & \alpha & \alpha+1 & 0 & 1 \\ \alpha+1 & \alpha+1 & \alpha & 1 & 0 \end{array} \qquad \begin{array}{c|cccc} \cdot & 0 & 1 & \alpha & \alpha+1 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & \alpha & \alpha+1 \\ \alpha & 0 & \alpha & \alpha+1 & 1 \\ \alpha+1 & 0 & \alpha+1 & 1 & \alpha \end{array}$$

Note that $a + a = 0$, for all $a \in \mathbb{F}_4$. Also, obviously $\mathbb{Z}_4 \neq \mathbb{F}_4$.

(10) *Let $\mathbb{F}$ be a field. Then the set of all polynomials over $\mathbb{F}$ is given by*

$$\mathbb{F}[X] := \left\{ \sum_{i=0}^{n} a_i X^i : \ a_i \in \mathbb{F}, n \geq 0 \right\}.$$

If convenient we may write X for X^1 and omit X^0. Next we define

$$\sum_{i=0}^{n} a_i X^i + \sum_{i=0}^{m} b_i X^i := \sum_{i=0}^{\max\{n,m\}} (a_i + b_i) X^i$$

$$\sum_{i=0}^{n} a_i X^i \cdot \sum_{i=0}^{m} b_i X^i := \sum_{i=0}^{n+m} \left(\sum_{j,k:j+k=i} a_j b_k \right) X^i$$

Then $(\mathbb{F}[X], +, \cdot)$ is a ring, but not a field, with $0 = 0 \cdot X^0$ the zero element and $1 = 1 \cdot X^0$ the one element. We can extent $\mathbb{F}[X]$ to a field, by setting

$$\mathbb{F}(X) := \left\{ \frac{f}{g} : \ f, g \in \mathbb{F}[X], g \neq 0 \right\}.$$

We identify two elements $\dfrac{f}{g}$ and $\dfrac{h}{k}$ precisely if $f \cdot k = h \cdot g$. Next we define

$$\frac{f}{g} + \frac{h}{k} := \frac{f \cdot k + h \cdot g}{g \cdot k} \qquad \text{and} \qquad \frac{f}{g} \cdot \frac{h}{k} := \frac{f \cdot h}{g \cdot k}$$

Then $(\mathbb{F}(X), +, \cdot)$ is a field, called **field of rational functions**.

Definition 1.3. *Let* $(\mathbb{F}, +, \cdot)$ *be a field and* E *a subset* $\mathbb{F}$. *We call* E *a* **subfield** *of* $\mathbb{F}$, *if* $(\mathbb{E}, +, \cdot)$ *is a field in its own right.*

Remark 1.4. *(1) A subset* E *of a field* $\mathbb{F}$ *is a subfield, if and only if* $E \cap \mathbb{F}^* \neq \emptyset$ *and for all* $a, b \in E$, *one has* $a+(-b) \in E$ *and, provided* $b \neq 0$, $a \cdot b^{-1} \in E$.

(2) $\mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ *and* $\mathbb{Q}(X) \subseteq \mathbb{R}(X) \subseteq \mathbb{C}(X)$

(3) Given a field $\mathbb{F}$, *the set* $\{aX^0 : a \in \mathbb{F}\}$ *is a subfield of* $\mathbb{F}(X)$.

(4) For a field $\mathbb{F}$ *with subfield* $\mathbb{E}$ *and* $r \in \mathbb{F}$, *a root of some* $g \in \mathbb{E}[X]$, *we set*

$$\mathbb{E}[r] := \{f(r) : f \in \mathbb{E}[X]\} = \left\{ \sum_{i=0}^{n} a_i \cdot r^i : \ a_i \in \mathbb{E}, n \geq 0 \right\}$$

Then $\mathbb{E}[r]$ *is a subfield of* $\mathbb{F}$. *For instance let* $f(X) = X^2 + X + 1 \in \mathbb{Q}[X]$ *with root* $\omega \in \mathbb{C}$. *Then* $\omega^2 = -\omega - 1$. *Hence* $\mathbb{Q}[\omega] = \{a + b\omega : a, b \in \mathbb{Q}\}$ *is a subfield of* $\mathbb{C}$.

(5) Let $\mathbb{F}$ *be a finite field with one element* $1_{\mathbb{F}}$. *Then there is a minimal* n *such that* $\sum_{i=1}^{n} 1_{\mathbb{F}} = 0$. *One can show that* n *has to be a prime number* p. *If we identify* $a \in \mathbb{F}_p$ *with* $\sum_{i=1}^{a} 1_{\mathbb{F}}$ *in* $\mathbb{F}$, *then* $\mathbb{F}_p$ *is a subfield of* $\mathbb{F}$. *It follows that* $\sum_{i=1}^{p} a = 0$, *for all* $a \in \mathbb{F}$.

2. Vector Spaces

Definition 2.1. *Let* $\mathbb{F}$ *be a field. A non-empty set* V *is called* $\mathbb{F}$**-vector space** *if there is a* **vector addition** $V \times V \to V$, $(v, w) \to v + w$ *and a* **scalar multiplication** $\mathbb{F} \times V \to V$, $(\lambda, v) \to \lambda v$, *such that for all* $v, w \in V$ *and* $\lambda, \mu \in \mathbb{F}$:

(V1) $+$ *is commutative and associative*
(V2) $+$ *has an identity element* 0_V, *called* **the zero vector**, *that is,* $v + 0_V = v = 0_V + v$
(V3) v *has an additive inverse* $-v$, *that is,* $v + (-v) = 0_V = (-v) + v$
(V4) $1v = v$
(V5) $\lambda(v + w) = \lambda v + \lambda w, \quad (\lambda\mu)v = \lambda(\mu v), \quad (\lambda + \mu)v = \lambda v + \mu v$

Remark/Example 2.2. *(1) Set* $V = \{0\}$ *and define* $0 + 0 := 0$ *and* $\lambda 0 := 0$, *for all* $\lambda \in \mathbb{F}$. *Then* V *is an* $\mathbb{F}$*-vector space, called the* **zero vector space** *and usually denoted by* 0. *Furthermore* $\mathbb{F}$ *is an* $\mathbb{F}$*-vector space.*

(2) The zero vector is unique. Also, given $v \in V$, *its inverse* $-v$ *is unique.*

(3) For all $v \in V$ *and* $\lambda \in \mathbb{F}$ *we have*

(i) $0v = 0_V$, *(ii)* $\lambda 0_V = 0_V$, *(iii)* $-v = (-1)v$

(iv) $\lambda v = 0_V$ *if and only if* $\lambda = 0$ *or* $v = 0_V$

(v) $(-\lambda)v = -(\lambda v) = \lambda(-v)$

(4) For an integer $n \geq 1$ let $V_1, \ldots, V_n$ be $\mathbb{F}$-vector spaces. We define

$$\bigoplus_{i=1}^{n} V_i := \{(v_1, \ldots, v_n) : \ v_i \in V_i\},$$

and for all $(v_1, \ldots, v_n), (w_1, \ldots, w_n) \in \bigoplus_{i=1}^n V_i$ and $\lambda \in \mathbb{F}$:

$$(v_1, \ldots, v_n) + (w_1, \ldots, w_n) := (v_1 + w_1, \ldots, v_n + w_n)$$
$$\lambda(v_1, \ldots, v_n) := (\lambda v_1, \ldots, \lambda v_n)$$

Then $\bigoplus_{i=1}^n V_i$ is an $\mathbb{F}$-vector space. If we set $V_i = \mathbb{F}$, for all $i = 1, \ldots, n$, then $\mathbb{F}^n := \bigoplus_{i=1}^n V_i$ is an $\mathbb{F}$-vector space. In this way we get the $\mathbb{Q}$-, $\mathbb{R}$- and $\mathbb{C}$-vector spaces $\mathbb{Q}^n$, $\mathbb{R}^n$ and $\mathbb{C}^n$, respectively.

(5) Just as in (5), the set $\mathbb{F}^\infty := \{(v_1, v_2, v_3, \ldots) : v_i \in \mathbb{F}\}$ of infinite sequences over $\mathbb{F}$ becomes an $\mathbb{F}$-vector space.

(6) If $\mathbb{E}$ is a subfield of $\mathbb{F}$, then every $\mathbb{F}$-vector space is also an $\mathbb{E}$-vector space. In particular, $\mathbb{F}$ is an $\mathbb{E}$-vector space. For instance $\mathbb{C}^n$ is both an $\mathbb{R}$- and $\mathbb{Q}$-vector space, and $\mathbb{R}^n$ is a $\mathbb{Q}$-vector space.

(7) The set $\mathbb{F}[X]$ of all polynomials over $\mathbb{F}$ together with their standard addition and scalar multiplication given by

$$\lambda\left(\sum_{i=0}^{n} \alpha_i X^i\right) := \sum_{i=0}^{n} (\lambda\alpha_i) X^i,$$

for $\lambda \in \mathbb{F}$ and $\sum_{i=0}^n \alpha_i X^i \in \mathbb{F}[X]$, is an $\mathbb{F}$-vector space. Similarly, $\mathbb{F}(X)$ is an $\mathbb{F}$-vector space.

(8) Let $m, n \geq 1$ be integers. Let $\mathcal{M}_{m\times n}(\mathbb{F})$ denote the set of all $m\times n$-matrices over $\mathbb{F}$. Then $\mathcal{M}_{m\times n}(\mathbb{F})$ is an $\mathbb{F}$-vector space, where

$$(\alpha_{ij}) + (\beta_{ij}) := (\alpha_{ij} + \beta_{ij}) \qquad \textit{and} \qquad \lambda(\alpha_{ij}) := (\lambda\alpha_{ij}),$$

for all $(\alpha_{ij}), (\beta_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{F})$ and $\lambda \in \mathbb{F}$.

(9) For an $\mathbb{F}$-vector space V and a set S, let $\mathcal{F}(S, V)$ denote the set of all functions $S \to V$. For $f, g \in \mathcal{F}(S, V)$ and $\lambda \in \mathbb{F}$ we define

$$f + g : s \mapsto f(s) + g(s) \qquad \textit{and} \qquad \lambda f : s \mapsto \lambda f(s)$$

Then $\mathcal{F}(S, V)$ is an $\mathbb{F}$-vector space. In particular, $\mathcal{F}(S, \mathbb{F})$ is an $\mathbb{F}$-vector space. In the case where $S = V$, we write $\mathcal{F}(V)$ for $\mathcal{F}(V, V)$.

Definition 2.3. *Let V be an $\mathbb{F}$-vector space. A non-empty subset U of V is called a* **subspace** *of V if U is closed under (i) vector addition and (ii) scalar multiplication, that is, for all $u, v \in U$ and $\lambda \in \mathbb{F}$ we have (i) $u + v \in U$ and (ii) $\lambda u \in U$.*

Remark/Example 2.4. *(1) A subspace U of V is an $\mathbb{F}$-vector space in its own right together with the same operations that come with V.*

(2) $0_V \in U$ and so 0_V is the zero vector of U

(3) $\{0_V\}$ and V are subspaces of V

(4) $\mathbb{F}v := \{\lambda v : \lambda \in \mathbb{F}\}$ is a subspace of V, for all $v \in V$

(5) Let $\lambda \in \mathbb{F}$ and $\lambda_i \in \mathbb{F}^*$, for $i = 1, \dots, n$. Then the solution set to the linear equation

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n = \lambda$$

is a subspace of $\mathbb{F}^n$ if and only if $\lambda = 0$, that is, the equation is **homogeneous**.

(6) Let $a, b, c, d \in \mathbb{R}$. A line $ax + by = c$ in $\mathbb{R}^2$ is a subspace of $\mathbb{R}^2$ and a plane $ax + by + cz = d$ in $\mathbb{R}^3$ is a subspace of $\mathbb{R}^3$ precisely if they pass through the origin.

(7) For real numbers $a < b$, let $\mathcal{C}([a,b],\mathbb{R})$ be the set of all continuous functions $f : [a,b] \to \mathbb{R}$. Then $\mathcal{C}([a,b],\mathbb{R})$ is a subspace of $\mathcal{F}([a,b],\mathbb{R})$.

(8) For $\mathbb{F} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$, the set of convergent sequences in $\mathbb{F}^\infty$ are a subspace of $\mathbb{F}^\infty$.

(9) The vector space $\mathbb{F}_2^2 = \{(0,0),(1,0),(0,1),(1,1)\}$ contains 4 vectors. The subspaces are 0, $\mathbb{F}_2^2$ and the 3 lines

$$\{(0,0),(1,0)\}, \quad \{(0,0),(0,1)\}, \quad \{(0,0),(1,1)\}.$$

Lemma 2.5. *Let $\mathcal{S}$ be a non-empty collection of subspaces of the $\mathbb{F}$-vector space V. Then $\bigcap_{U \in \mathcal{S}} U$ is a subspace of V.*

Proof. Homework. □

Corollary 2.6. *Let $m, n \geq 1$ be integers and $A \in \mathcal{M}_{m \times n}(\mathbb{F})$. Then the set of solutions to the homogeneous system of linear equations $Ax = 0$ is a subspace of $\mathbb{F}^n$.*

Example 2.7. *In $\mathbb{R}^3$ consider a homogeneous system of linear equations*

$$\begin{aligned} x + y - 2z &= 0 \\ x - 2y + z &= 0 \end{aligned}$$

The solution set to each equation is a plane passing through the origin. As they differ, their intersection is a line. One calculates that the solution set is $\mathbb{R}(1,1,1) = \{(x,x,x) : x \in \mathbb{R}\}$.

Remark 2.8. *Generally, the union of subspaces U_1 and U_2 of an $\mathbb{F}$-vector space V is not a subspace of V. Take for instance $U_1 := \mathbb{R}(1,0) = \{(a,0) : \ a \in \mathbb{R}\}$ and $U_2 := \mathbb{R}(0,1) = \{(0,b) : \ b \in \mathbb{R}\}$ in $\mathbb{R}^2$.*

Definition 2.9. *Let U_1 and U_2 be subspaces in the $\mathbb{F}$-vector space V. We call*

$$U_1 + U_2 := \{u_1 + u_2 : u_1 \in U_1, u_2 \in U_2\}$$

the **sum** *of U_1 and U_2. If furthermore $U_1 \cap U_2 = \{0_V\}$, then we call*

$$U_1 \oplus U_2 := U_1 + U_2$$

the **direct sum** *of U_1 and U_2.*

Lemma 2.10. *Let U_1 and U_2 be subspaces of the $\mathbb{F}$-vector space V. Then $U_1 + U_2$ is the smallest subspace of V containing $U_1 \cup U_2$.*

Proof. As $0_V \in U_1 \cap U_2$ we have that $0_V = 0_V + 0_V \in U_1 + U_2$. Next let $u, v \in U_1 + U_2$, that is, $u = u_1 + u_2$ and $v = v_1 + v_2$, where $u_i, v_i \in U_i$, for $i = 1, 2$, and $\lambda \in \mathbb{F}$. Then

$$u + v = (u_1 + u_2) + (v_1 + v_2) \overset{(V1)}{=} (u_1 + v_1) + (u_2 + v_2) \in U_1 + U_2$$

$$\lambda u = \lambda(u_1 + u_2) \overset{(V5)}{=} \lambda u_1 + \lambda u_2 \in U_1 + U_2$$

Hence, by Definition 2.3, $U_1 + U_2$ is a subspace of V. Finally, as $0_V \in U_1 \cap U_2$, we get that $U_1, U_2 \subseteq U_1 + U_2$. Finally, by additivity, any subspace U of V containing $U_1 \cup U_2$, must contain the elements in $U_1 + U_2$. □

Lemma 2.11. *Let U_1 and U_2 be subspaces in the $\mathbb{F}$-vector space V such that $U_1 \cap U_2 = \{0_V\}$. Then for every $u \in U_1 \oplus U_2$ there are unique $u_1 \in U_1$ and $u_2 \in U_2$ such that $u = u_1 + u_2$*

Proof. Assume that $u = u_1 + u_2 = v_1 + v_2$, where $u_i, v_i \in U_i$, for $i = 1, 2$. Then $(-v_1) + u_1 = v_2 + (-u_2)$. As the LHS lies in U_1 and the RHS lies in U_2, both sides must equal 0_V. Thus uniqueness follows. □

Example 2.12. *In $\mathbb{R}^3$ consider the subspaces $U_1 = \{(x, y, 0) : x, y \in \mathbb{R}\}$, $U_2 = \{(0, y, z) : y, z \in \mathbb{R}\}$ and $U_3 = \{(x, x, x) : x \in \mathbb{R}\}$. Then $U_1 \cap U_2 = \{(0, y, 0) : y \in \mathbb{R}\}$ and for all $(x, y, z) \in \mathbb{R}^3$,*

$$(x, y, z) = (x, y, 0) + (0, 0, z) = (x, 0, 0) + (0, y, z).$$

Hence $\mathbb{R}^3$ is the sum, but not the direct sum, of U_1 and U_2.

Next, $U_1 \cap U_3 = \{(0, 0, 0)\}$. Then $(x, y, z) = (x - z, y - z, 0) + (z, z, z)$ is unique, for all $(x, y, z) \in \mathbb{R}^3$. Hence $\mathbb{R}^3 = U_1 \oplus U_3$.

Remark 2.13. *Given subspaces $U_1, \ldots, U_n$ of the $\mathbb{F}$-vector space V, we define their* **sum** *as*

$$U_1 + \ldots + U_n := \{u_1 + \ldots + u_n : u_i \in U_i, i = 1, \ldots, n\}$$

Then $U_1 + \ldots + U_n$ is the smallest subspace of V containing $U_1 \cup \ldots \cup U_n$. Furthermore this sum is a **direct sum***, that is, the expression for each element in $U_1 + \ldots + U_n$ is unique, if and only if*

$$U_i \cap (U_1 + \ldots + U_{i-1} + U_{i+1} + \ldots + U_n) = \{0_V\},$$

for all $i = 1, \ldots, n$. In this case we write $U_1 \oplus \ldots \oplus U_n$.

Definition 2.14. *Let V be an $\mathbb{F}$-vector space and $M \neq \emptyset$ be a subset of V.*

(1) *A vector $v \in V$ is called a* **linear combination** *of vectors $v_1, \ldots, v_n$ in M, if there are scalars $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ such that*

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n.$$

(2) *The* **span** *of M, denoted by $\operatorname{span}(M)$, is the set of all linear combinations of vectors in M, that is,*

$$\operatorname{span}(M) := \{\lambda_1 v_1 + \ldots + \lambda_n v_n : \; n \geq 1, \lambda_i \in \mathbb{F}, v_i \in M, 1 \leq i \leq n\}.$$

Moreover $\operatorname{span}(\emptyset) := \{0_V\}$. Furthermore, if $U = \operatorname{span}(M)$, then we say that U is **spanned** *by M and m is a* **spanning set** *of U.*

(3) *The set M called* **linearly dependent** *if there is a non-trivial way to express 0_V as a linear combination of distinct vectors in M, that is, there are $\lambda_1, \ldots, \lambda_n \in \mathbb{F}^*$ and distinct $v_1, \ldots, v_n \in M$ such that*

$$0_V = \lambda_1 v_1 + \ldots + \lambda_n v_n.$$

Otherwise we call M **linearly independent***. The empty set is defined as linearly independent.*

Lemma 2.15. *Let V be an $\mathbb{F}$-vector space and $M \subseteq V$. Then $\operatorname{span}(M)$ is the smallest subspace of V which contains M.*

Proof. By definition the span is non-empty and one checks quickly that it is closed under vector addition and scalar multiplication. Hence $\operatorname{span}(M)$ is a subspace of V, by Definition 2.3. Also it is evident that any subspace of V containing M must contain all elements in $\operatorname{span}(M)$. □

Lemma 2.16. *Let V be an $\mathbb{F}$-vector space and M a subset of V. Then the following are equivalent*

(1) *M is linearly independent*

(2) *$\lambda_1 v_1 + \ldots + \lambda_n v_n = 0_V$ implies that $\lambda_1 = \ldots = \lambda_n = 0$, for all $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ and all distinct $v_1, \ldots, v_n \in M$*

(3) *No $v \in M$ is a linear combination of elements in $M \backslash \{v\}$*

Proof. (1) $\Leftrightarrow$ (2) follows from the definition. Next assume (1) and not (3), that is, $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$, for some $v \in M$, $v_i \in M \backslash \{v\}$ and $\lambda_i \in \mathbb{F}$. Then $0_V = (-1)v + \lambda_1 v_1 + \ldots + \lambda_n v_n$ is a non-trivial linear combination of 0_V, contradicting the linear independence of M. Vice versa, assume (3) and not (1). Then $0_V = \lambda_1 v_1 + \ldots + \lambda_n v_n$, for $\lambda_i \in \mathbb{F}^*$ and distinct $v_i \in M$. W.l.o.g, $\lambda_1 = -1$, by multiplying the scalar $(-\lambda_1)^{-1}$ onto the equation. Then $v_1 = \lambda_2 v_2 + \ldots + \lambda_n v_n$, contradicting (3). $\square$

Remark/Example 2.17. *(1) If there is ambiguity over which field $\mathbb{F}$ the span is taken, we write $\operatorname{span}_{\mathbb{F}}(M)$ instead of $\operatorname{span}(M)$. Consider for instance $1, i \in \mathbb{C}$. Here we have $\operatorname{span}_{\mathbb{Q}}(1, i) = \{a + bi : a, b \in \mathbb{Q}\} \subsetneq \mathbb{C}$, while $\operatorname{span}_{\mathbb{R}}(1, i) = \{a + bi : a, b \in \mathbb{R}\} = \mathbb{C}$.*

(2) If $0_V \in M$, then M is linearly dependent.

(3) Let $v \in V$. Then $\operatorname{span}(v) = \{\lambda v : \lambda \in \mathbb{F}\} = \mathbb{F}v$.

(4) For $v_1, \ldots, v_n \in V$ we have

$$\operatorname{span}(v_1, \ldots, v_n) = \{\lambda_1 v_1 + \ldots + \lambda_n v_n : \ \lambda_i \in \mathbb{F}\} = \mathbb{F}v_1 + \ldots + \mathbb{F}v_n$$

If $v_1, \ldots, v_n$ are linearly independent, then $v_i \notin \operatorname{span}(\{v_1, \ldots, v_n\} \backslash \{v_i\})$, for all $i = 1, \ldots, n$, by Lemma 2.16, and so

$$\operatorname{span}(v_1, \ldots, v_n) = \mathbb{F}v_1 \oplus \ldots \oplus \mathbb{F}v_n$$

In particular, for all $v \in \operatorname{span}(v_1, \ldots, v_n)$, there are unique $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ such that $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$.

(5) In $\mathbb{F}^n$, set $e_i := (e^i_1, \ldots, e^i_n)$, for $i = 1, \ldots, n$, where $e^i_j = 0_{\mathbb{F}}$, whenever $i \neq j$, and $e^i_j = 1_{\mathbb{F}}$, if $i = j$. Then for every $v = (v_1, \ldots, v_n) \in \mathbb{F}^n$ we have $v = v_1 e_1 + \ldots + v_n e_n$, and so $\operatorname{span}(e_1, \ldots, e_n) = \mathbb{F}^n$. Furthermore the set $\{e_1, \ldots, e_n\}$ is linearly independent, by Lemma 2.16, because for all $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ we have

$$0_V = \lambda_1 e_1 + \ldots + \lambda_n e_n \Rightarrow 0_V = (\lambda_1, \ldots, \lambda_n) \Rightarrow \lambda_1 = \ldots = \lambda_n = 0.$$

Overall, $\mathbb{F}^n = \operatorname{span}(e_1, \ldots, e_n) = \mathbb{F}e_1 \oplus \ldots \oplus \mathbb{F}e_n$.

(6) For instance, in $\mathbb{R}^3$ we have $e_1 := (1, 0, 0)$, $e_2 := (0, 1, 0)$ and $e_3 := (0, 0, 1)$. Then

$$\operatorname{span}_{\mathbb{R}}(e_1, e_2, e_3) = \mathbb{R}^3 = \mathbb{R}e_1 \oplus \mathbb{R}e_2 \oplus \mathbb{R}e_3$$

If furthermore $v := e_1 + e_2 + e_3 = (1, 1, 1)$, then $\operatorname{span}_{\mathbb{R}}(e_1, e_2, e_3, v) = \mathbb{R}^3$, but $\{e_1, e_2, e_3, v\}$ are not linearly independent.

(7) In $\mathbb{F}^\infty$, for every integer $i \geq 1$, let e_i be the sequence that is zero everywhere, except in position i, which is one. Then $\{e_i : \ i \geq 1\}$ is linearly independent, but does not span $\mathbb{F}^\infty$. Describe $\operatorname{span}(e_i : \ i \geq 1)$.

(8) Let $\mathbb{F} = \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. In $\mathbb{F}[X]$, the set $\{X^i : i \geq 0\} = \{1, X, X^2, X^3, \ldots\}$ is linearly independent and spans $\mathbb{F}[X]$.

(9) Let $m, n \geq 1$ be integers. In $\mathcal{M}_{m\times n}(\mathbb{F})$, let $E_{i,j}$ denote the matrix, with a one in entry (i, j) and zeros elsewhere. Then $\{E_{i,j} : 1 \leq i \leq m,\ 1 \leq j \leq n\}$ is linearly independent and spans $\mathcal{M}_{m\times n}(\mathbb{F})$.

(10) The plane $P:\ x - 2y - z = 0$ is a subspace of $\mathbb{R}^3$. Then $v = (x, y, z) \in \mathbb{R}^3$ lies in P if and only if $z = x - 2y$, that is, $v = x(1, 0, 1) + y(0, 1, -2)$. Hence $P = \operatorname{span}_{\mathbb{R}}((1, 0, 1), (0, 1, -2))$. In fact, $P = \mathbb{R}(1, 0, 1) \oplus \mathbb{R}(0, 1, -2)$. Equally, $v = (x, y, z) \in \mathbb{R}^3$ lies in P if and only if $v = y(2, 1, 0) + z(1, 0, 1)$. So, $P = \operatorname{span}_{\mathbb{R}}((2, 1, 0), (1, 0, 1))$, and again $P = \mathbb{R}(2, 1, 0) \oplus \mathbb{R}(1, 0, 1)$.

(11) In $\mathcal{F}(\mathbb{R}, \mathbb{R})$ let $f(x) = \sin(x)$, $g(x) = \cos(x)$ and $h(x) = \exp(x)$, for all $x \in \mathbb{R}$. Furthermore let $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ such that

$$0 = \lambda_1 f + \lambda_2 g + \lambda_3 h$$

Then for $x = 0$, we have $0 = \lambda_1 \sin(0) + \lambda_2 \cos(0) + \lambda_3 \exp(0) = \lambda_2 + \lambda_3$, that is, $-\lambda_2 = \lambda_3$. Next for $x = \pi$ we get $0 = -\lambda_2 + \lambda_3 \exp(\pi)$, and so $0 = \lambda_3(1 + \exp(\pi))$. This forces $\lambda_3 = 0$ and so $\lambda_2 = 0$. Now for $x = \frac{\pi}{2}$, we have $0 = \lambda_1$. Hence all $\lambda_i = 0$ and so $\{f, g, h\}$ is linearly independent.

Theorem 2.18. *Let $m, n \geq 1$ be integers and $v_1, \ldots, v_n$ be (column) vectors in $\mathbb{F}^m$. Furthermore let $A \in \mathcal{M}_{m\times n}(\mathbb{F})$, where column i is given by vector v_i.*

(a) The following are equivalent:

(1) The set $\{v_1, \ldots, v_n\}$ is linearly independent.

(2) the homogeneous system $Ax = 0$ only has the trivial solution $x = 0$.

(3) the reduced row echelon form of A has n leading ones.

(b) The following are equivalent:

(1) The set $\{v_1, \ldots, v_n\}$ spans $\mathbb{F}^m$.

(2) the system $Ax = v$ has a solution for every column vector $v \in \mathbb{F}^m$.

(3) the reduced row echelon form of A has m leading ones.

Proof. (a) (1) $\Leftrightarrow$ (2): This follows from Lemma 2.16.

(2) $\Leftrightarrow$ (3): Let R be the REF of A. Then $Ax = 0$ and $Rx = 0$ have the same solution set. But $Rx = 0$ has a non-trivial solution precisely if there is at least one column without a leading one.

(b) (1) $\Leftrightarrow$ (2): Obvious

(2) $\Leftrightarrow$ (3): Precisely when R has a row without a leading one (i.e. a zero row), there is some $v' \in \mathbb{F}^m$ so that $Rx = v'$ has no solution, which is equivalent to there being some $v \in \mathbb{F}^m$ so that $Ax = v$ has no solution. $\square$

Example 2.19. *Consider the set $M = \{(1,-1,0),(0,1,-2),(1,0,3)\}$ over $\mathbb{R}$ and $\mathbb{F}_5$, respectively. Here $n = m = 3$. In either case*

$$(A|v) = \left(\begin{array}{ccc|c} 1 & 0 & 1 & a \\ -1 & 1 & 0 & b \\ 0 & -2 & 3 & c \end{array}\right) \overset{R2+R1}{\to} \left(\begin{array}{ccc|c} 1 & 0 & 1 & a \\ 0 & 1 & 1 & a+b \\ 0 & -2 & 3 & c \end{array}\right) \overset{R3+2R2}{\to} \left(\begin{array}{ccc|c} 1 & 0 & 1 & a \\ 0 & 1 & 1 & a+b \\ 0 & 0 & 5 & c+2(a+b) \end{array}\right)$$

Over $\mathbb{R}$, we divide the last row by 5 and thus obtain three leading ones. Hence the set M is linearly independent and spans $\mathbb{R}^3$. However, $5 = 0$ in $\mathbb{F}_3$ and so M is linearly dependent in $(\mathbb{F}_5)^3$ and does not span $(\mathbb{F}_5)^3$. For instance $(1,0,3) = (1,-1,0) + (0,1,-2)$. However, $M' := \{(1,-1,0),(0,1,-2\}$ is linearly independent in $(\mathbb{F}_5)^3$. Furthermore, $Ax = v$ has a solution if and only if $c+2a+2b = 0$. In particular, M (and M') only span the plane $2a+2b+c = 0$ in $(\mathbb{F}_5)^3$.

Corollary 2.20. *Let $m, n \geq 1$ be integers and $M := \{v_1, \dots, v_n\}$ a set of vectors in $\mathbb{F}^m$.*

(1) If M is linearly independent, then $n \leq m$.

(2) If M spans $\mathbb{F}^m$, then $m \leq n$.

Theorem 2.21. *Let M be a subset of an $\mathbb{F}$-vector space V and let $v \in V$.*

(1) (Plus Theorem) If M is linearly independent and $v \notin \operatorname{span}(M)$, then $M \cup \{v\}$ is linearly independent.

(2) (Minus Theorem) If $v \in \operatorname{span}(M \backslash \{v\})$, for some $v \in M$, (i.e. M is linearly dependent), then $\operatorname{span}(M) = \operatorname{span}(M \backslash \{v\})$.

Proof. Homework. $\square$

Example 2.22. *In $\mathbb{F}^\infty$, the set $M = \{e_i : \ i \geq 1\}$ is linearly independent, but does not span $\mathbb{F}^\infty$. Let p be the sequence of all entries equal to $1_{\mathbb{F}}$. Then $p \notin \operatorname{span}(M)$ and so $M \cup \{p\}$ is linearly independent.*

In $\mathbb{R}^3$, the set $\{e_1, e_2, e_3, v\}$, for $v = (1,1,1)$, is linearly dependent and spans $\mathbb{R}^3$. As for instance, $e_1 = v - e_2 - e_3$, we get that $\{e_2, e_3, v\}$ still spans $\mathbb{R}^3$.

Definition 2.23. *Let V be an $\mathbb{F}$-vector space.*

(1) We call V **finite-dimensional** *if V has a finite spanning set, that is, there are vectors $v_1, \dots, v_n$ in V such that $V = \operatorname{span}(v_1, \dots, v_n)$.*

(2) A set $\mathcal{B}$ is called **basis** *of V if*

(a) $\mathcal{B}$ is linearly independent and
(b) $\mathcal{B}$ spans V

(3) A basis $\mathcal{B}$ of V is called **finite** *if the set $\mathcal{B}$ is finite, and* **infinite** *otherwise.*

Lemma 2.24. *(Steinitz Exchange Lemma) Let $n, m \geq 0$ be integers, V an $\mathbb{F}$-vector space and $M = \{w_1, \dots, w_m\}$ and $N = \{u_1, \dots, u_n\}$ subsets of V such that M spans V and N is linearly independent. Then $n \leq m$ and there are n vectors in M, say $\{w_1, \dots, w_n\}$, so that $\{u_1, \dots, u_n, w_{n+1}, \dots, w_m\}$ spans V.*

Proof. We prove the statement by induction on $n \geq 0$. If $n = 0$ there is nothing to show. Now assume that the statement holds for fewer than n elements. Then $n - 1 \leq m$ and there are $n - 1$ vectors in M, say $\{w_1, \dots, w_{n-1}\}$, such that $\{u_1, \dots, u_{n-1}, w_n, \dots, w_m\}$ spans V. Hence there are $\lambda_1, \dots, \lambda_m \in \mathbb{F}$ such that

$$u_n = \sum_{i=1}^{n-1} \lambda_i u_i + \sum_{i=n}^{m} \lambda_i w_i.$$

If $m = n - 1$ or $\lambda_i = 0$, for all $i = n, \dots, m$, then u_n is a linear combination of $u_1, \dots, u_{n-1}$ contradicting the linear independence of N. Hence $n \leq m$ and, say $\lambda_n \neq 0$. It follows that

$$w_n = (\lambda_n)^{-1} u_n + (-\lambda_n)^{-1} \left(\sum_{i=1}^{n-1} \lambda_i u_i + \sum_{i=n+1}^{m} \lambda_i w_i \right)$$

Hence, by the Minus Theorem,

$$V = \operatorname{span}(u_1, \dots, u_{n-1}, u_n, w_n, \dots, w_m) = \operatorname{span}(u_1, \dots, u_n, w_{n+1}, \dots, w_m).$$

□

Corollary 2.25. *Every subspace U of a finite-dimensional vector space V has a finite basis. In particular, U is finite-dimensional.*

Proof. Let $V = \operatorname{span}(v_1, \dots, v_m)$ and U a subspace of V. If $U = \{0_V\}$, we are done (see Remark 2.28). Otherwise pick a finite linearly independent subset N of U. If N spans U, we have a finite basis. Otherwise there is some $u \in U$ such that $u \notin \operatorname{span}(N)$, and so $N \cup \{u\}$ is linearly independent, by the Plus Theorem. This process terminates eventually with a finite basis for U, as any linearly independent subset of U has at most m vectors, by Steinitz Exchange Lemma. □

Corollary 2.26. *Any two bases in a finite-dimensional vector space have the same finite size.*

Proof. Let $V = \operatorname{span}(v_1, \dots, v_m)$ and $\mathcal{B}$ a basis of V. Note that $\mathcal{B}$ is finite as otherwise we could pick $m+1$ linearly independent vectors from $\mathcal{B}$ in contradiction to Steinitz. Next let $\mathcal{B}'$ be a second basis. Then by Steinitz, $\mathcal{B}$ cannot have more elements that $\mathcal{B}'$, while $\mathcal{B}'$ cannot have more elements than $\mathcal{B}$. $\square$

Definition 2.27. *Let V be a finite-dimensional $\mathbb{F}$-vector space. Then the* **dimension** *of V, denoted by $\dim(V)$, is the size of any basis of V. If V is not finite-dimensional, we say V has infinite dimension and write $\dim(V) = \infty$.*

Remark/Example 2.28. *(1) Using a more advanced argument, known as Zorn's Lemma, one can show that every vector space has a basis, even those for which there exist no finite spanning set.*

(2) If there is ambiguity about the field $\mathbb{F}$, we write $\dim_{\mathbb{F}}(V)$ instead of $\dim(V)$ and talk about an $\mathbb{F}$-basis. For instance, the set $\{1, i\}$ is an $\mathbb{R}$-basis of $\mathbb{C}$ and $\dim_{\mathbb{R}}(\mathbb{C}) = 2$. However, though still linearly independent, $\{1, i\}$ does not span $\mathbb{C}$ as a $\mathbb{Q}$-vector space. In fact $\dim_{\mathbb{Q}}(\mathbb{C}) = \infty$. Also $\dim_{\mathbb{Q}}(\mathbb{R}) = \infty$.

(3) The zero vector space has empty basis and dimension zero. For $\mathbb{F}$, as an $\mathbb{F}$-vector space, the set $\{\lambda\}$, for any $\lambda \in \mathbb{F}^*$, is a basis. Hence $\dim_{\mathbb{F}}(\mathbb{F}) = 1$.

(4) Let $v \in V$. Then $\mathbb{F}v$ has basis $\{v\}$ and dimension one. In fact, $\{\lambda v\}$ is a basis of $\mathbb{F}v$, for all $\lambda \in \mathbb{F}^*$.

(5) A set $\mathcal{B} = \{v_1, \dots, v_n\}$ is a basis of V if and only if $V = \mathbb{F}v_1 \oplus \ldots \oplus \mathbb{F}v_n$ if and only if for every $v \in V$ there are unique $\lambda_1, \dots, \lambda_n \in \mathbb{F}$ such that $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$.

(6) The set $\{e_1, \dots, e_n\}$ is the standard basis of $\mathbb{F}^n$. Hence $\dim(\mathbb{F}^n) = n$. One can show that $\{e_1 + (-e_2), e_2 + (-e_3), \dots, e_{n-1} + (-e_n), e_n\}$ is another basis of $\mathbb{F}^n$, while $\{e_1 + (-e_2), e_2 + (-e_3), \dots, e_{n-1} + (-e_n), e_n + (-e_1)\}$ is not.

(7) We have $\dim(\mathbb{F}^\infty) = \infty$, as $\{e_i : i \geq 1\}$ is linearly independent.

(8) Let $\mathbb{F} = \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. We have $\dim(\mathbb{F}[X]) = \infty$, where $\{X^i : i \geq 0\}$ is the standard basis.

(9) Let $m, n \geq 1$ be integers. Then $\dim(\mathcal{M}_{m \times n}(\mathbb{F})) = mn$, where the set $\{E_{i,j} : 1 \leq i \leq m,\ 1 \leq j \leq n\}$ is the standard basis.

(10) Every spanning set of a vector space V contains a basis of V and every linearly independent subset of V, can be extended to a basis of V. In the finite case this is due to the Minus Theorem and Plus Theorem, respectively.

Lemma 2.29. *Let V be an $\mathbb{F}$-vector space with subspace U. Then*

(1) $\dim(U) \leq \dim(V)$

(2) there is a subspace W of V such that $V = U \oplus W$.

Proof. (1) This follows as a basis of U is a linearly independent in V, and hence cannot contain more elements than a basis of V.

(2) Let $\mathcal{B}'$ be a basis of U. Then $\mathcal{B}'$ can be extended to a basis $\mathcal{B} = \mathcal{B}' \cup \mathcal{B}''$ of V. Set $W := \mathrm{span}(\mathcal{B}'')$ and check that $V = U \oplus W$. $\square$

Example 2.30. *(1) The subspaces in $\mathbb{F}^3$ are thus: (i) the zero subspace, (ii) the lines $\mathbb{F}v$, for $v \in \mathbb{F}^3 \backslash \{0_V\}$, (iii) the planes $ax + by + cz = 0$, for $(a, b, c) \in \mathbb{F}^3 \backslash \{0_V\}$ and (iv) $\mathbb{F}^3$.*

(2) For $v := (a, b, c) \in \mathbb{F}^3 \backslash \{0_V\}$ we have

$$\mathbb{F}^3 = \mathbb{F}v \oplus (ax + by + cz = 0)$$

Lemma 2.31. *Let V be an n-dimensional $\mathbb{F}$-vector space and let $\mathcal{B}$ be a set of m vectors in V. Then the following are equivalent:*

(1) $\mathcal{B}$ is basis of V.

(2) $\mathcal{B}$ is linearly independent and $m = n$.

(3) $\mathcal{B}$ spans V and $m = n$.

Proof. Clearly (1) implies (2) and (3). That both (2) and (3) imply (1), follows from Steinitz and the Plus Theorem and Minus Theorem, respectively. $\square$

Lemma 2.32. *Let V be a finite-dimensional $\mathbb{F}$-vector space with subspaces U and W. Then*

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W).$$

In particular, if $U \cap W = \{0_V\}$, then $\dim(U \oplus W) = \dim(U) + \dim(W)$.

Proof. Let $\{a_1, \ldots, a_r\}$ be a basis for $U \cap W$. So $\dim(U \cap W) = r$. We extend it to a basis $\{a_1, \ldots, a_r, b_1, \ldots, b_s\}$ of U and to a basis $\{a_1, \ldots, a_r, c_1, \ldots, c_t\}$ of W. So $\dim(U) = r + s$ and $\dim(W) = r + t$.

We claim that $\mathcal{B} := \{a_1, \ldots, a_r, b_1, \ldots, b_s, c_1, \ldots, c_t\}$ is a basis for $U + W$. It is straight forward to show that $\mathcal{B}$ spans $U + W$. Next consider scalars $\alpha_i, \beta_j, \gamma_k \in \mathbb{F}$ such that $\sum \alpha_i a_i + \sum \beta_j b_j + \sum \gamma_k c_k = 0$, that is,

$$\sum \alpha_i a_i + \sum \beta_j b_j = -\sum \gamma_k c_k. \tag{1}$$

As the LHS lies in U and the RHS lies in W, both sides lie in $U \cap W$. In particular we can write the LHS as $\sum \mu_\ell a_\ell$, for some scalars $\mu_\ell \in \mathbb{F}$. But now $\sum \mu_\ell a_\ell + \sum \gamma_k c_k = 0$. As $\{a_1, \ldots, a_r, c_1, \ldots, c_t\}$ is linearly independent, this implies that each γ_k is zero. Plug this back into (1) gives

$$\sum \alpha_i a_i + \sum \beta_j b_j = 0.$$

Then all α_i and β_j are zero, as $\{a_1, \ldots, a_r, b_1, \ldots, b_s\}$ is linearly independent. Hence $\mathcal{B}$ is linearly independent. Overall $\mathcal{B}$ is a basis for $U + W$. So

$$\dim(U + W) = r + s + t = \dim(U) + \dim(W) - \dim(U \cap W).$$

$\square$

3. Linear Maps / Homomorphisms

Definition 3.1. *Let V, W be $\mathbb{F}$-vector spaces. A function $T : V \to W$ is called* **linear map** *or* **homomorphism** *if for all $v, u \in V$ and $\lambda \in \mathbb{F}$:*

(i) $T(v + u) = T(v) + T(u)$ and
(ii) $T(\lambda v) = \lambda T(v)$

We write $\operatorname{Hom}(V, W)$ *for the set of all homomorphisms $V \to W$.*

Remark/Example 3.2. *(1) The function $T : V \to W : v \mapsto 0_W$, is a linear map, called* **zero map**, *since $T(v + u) = 0_W = 0_W + 0_W = T(v) + T(u)$ and $T(\lambda v) = 0_W = \lambda 0_W = \lambda T(v)$, for all $v, u \in V$ and $\lambda \in \mathbb{F}$.*

(2) If V is a subspace of W, then $I : V \to W : v \mapsto v$, is a linear map as $I(v + u) = v + u = I(v) + I(u)$ and $I(\lambda v) = \lambda v = \lambda I(v)$, for all $v, u \in V$ and $\lambda \in \mathbb{F}$. Generally, we call I the **inclusion map** *from V into W. In the case $V = W$, we call I the* **identity map** *on V.*

(3) We have $T(0_V) = 0_W$ and $T(-v) = -T(v)$, for all $v \in U$, since $T(0_V) = T(0_{\mathbb{F}} v) = 0_{\mathbb{F}} T(v) = 0_W$ and $T(-v) = T((-1)v) = (-1)T(v) = -T(v)$.

(4) Let $m, n \geq 1$ be integers and $A \in \mathcal{M}_{m \times n}(\mathbb{F})$. Then $T_A : \mathbb{F}^n \to \mathbb{F}^m : v \mapsto Av$, is a linear map. Note that Av only makes sense if we take v as a column vector. For instance, for $A = \begin{pmatrix} 5 & 1+3i & -i \\ 1-3i & 2 & \frac{3-i}{5} \end{pmatrix} \in \mathcal{M}_{2 \times 3}(\mathbb{C})$, then

$$T_A : \mathbb{C}^3 \to \mathbb{C}^2 : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 5x + (1+3i)y - iz \\ (1-3i)x + 2y + \left(\frac{3-i}{5}\right) z \end{pmatrix}$$

(5) For $\alpha \in \mathbb{F}$ the map $T_\alpha : \mathbb{F}[X] \to \mathbb{F} : f = \sum_{i=0}^n \lambda_i X^i \mapsto f(\alpha) := \sum_{i=0}^n \lambda_i \alpha^i$ is linear, called the **evaluation homomorphism** *of α.*

(6) The map $T : \mathbb{F}[X] \to \mathbb{F}[X] : \sum_{i=0}^n \alpha_i X^i \mapsto \sum_{i=1}^n (i \cdot \alpha_i) X^{i-1}$ is linear, mapping each polynomial onto its derivative. (Here $i \cdot \alpha = \sum_{k=1}^i \alpha$). For instance $T(3 + 7X^2 - 6X^8) = 14X - 48X^7$.
Likewise for the subspace $\mathcal{D}(\mathbb{R}, \mathbb{R})$ of differentiable functions in $\mathcal{F}(\mathbb{R}, \mathbb{R})$, we have that $T : \mathcal{D}(\mathbb{R}, \mathbb{R}) \to \mathcal{F}(\mathbb{R}, \mathbb{R}) : f \mapsto f'$ is a linear map.

(7) In $\mathbb{F}^n$, reflections and rotations that fix the origin are linear maps $\mathbb{F}^n \to \mathbb{F}^n$.

Lemma 3.3. *Let V, W be $\mathbb{F}$-vector spaces. Then* $\operatorname{Hom}(V, W)$ *is a subspace of $\mathcal{F}(V, W)$ (see Example 2.2 (9)). In particular,* $\operatorname{Hom}(V, W)$ *is a vector space, where for all $S, T \in \operatorname{Hom}(V, W)$ and $\lambda \in \mathbb{F}$.*

(i) $S + T : V \to W : v \mapsto S(v) + T(v)$
(ii) $\lambda T : V \to W : v \mapsto \lambda T(v)$

Proof. As $\mathrm{Hom}(V,W)$ contains the zero map, it is a non-empty subset of $\mathcal{F}(V,W)$. Next let $S,T \in \mathrm{Hom}(V,W)$ and $\lambda \in \mathbb{F}$. Are $S+T, \lambda T \in \mathrm{Hom}(V,W)$? Let $v,u \in V$ and $\mu \in \mathbb{F}$. Then

$$\begin{aligned}(S+T)(u+v) &= S(u+v)+T(u+v) = (S(u)+S(v))+(T(u)+T(v))\\ &\overset{(V1)}{=} (S(u)+T(u))+(S(v)+T(v)) = (S+T)(u)+(S+T)(v)\\ (S+T)(\mu v) &= S(\mu v)+T(\mu v) = \mu S(v)+\mu T(v) \overset{(V5)}{=} \mu(S(v)+T(v))\\ &= \mu(S+T)(v)\end{aligned}$$

Hence $S+T \in \mathrm{Hom}(V,W)$. The rest is homework. $\square$

Lemma 3.4. *Let V, W be $\mathbb{F}$-vector spaces and $\mathcal{B}$ a basis of V. Then every map $T : \mathcal{B} \to W$ extends uniquely to a $T \in \mathrm{Hom}(V,W)$. In particular, every $T \in \mathrm{Hom}(V,W)$ is uniquely determined by its behaviour on $\mathcal{B}$.*

Proof. Let $T : \mathcal{B} \to W$ be given and let $v \in V$. Then $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$, for $\lambda_i \in \mathbb{F}$ and $v_i \in \mathcal{B}$. Now set $T(v) := \lambda_1 T(v_1)+\ldots+\lambda_n T(v_n)$. It is easy to show that $T \in \mathrm{Hom}(V,W)$. Furthermore note that $T(v)$ is uniquely determined by $T(v_1),\ldots,T(v_n)$. $\square$

Example 3.5. *(1) If V or W are the zero vector space, then $\mathrm{Hom}(V,W)$ only contains the zero map. In particular, $\dim(\mathrm{Hom}(V,W)) = 0$.*

(2) How many elements are there is $\mathrm{Hom}((\mathbb{F}_3)^3,(\mathbb{F}_3)^2)$? Note that $(\mathbb{F}_3)^3$ has the standard basis $\mathcal{B} = (e_1,e_2,e_3)$. Also, there are nine vectors in $(\mathbb{F}_3)^2$. Thus each e_i can be mapped on one of nine vectors, and so there are $9^3 = 729$ different ways to define a function T on $\mathcal{B}$. Each such T extends to a unique element in $\mathrm{Hom}((\mathbb{F}_3)^3,(\mathbb{F}_3)^2)$ and each element in $\mathrm{Hom}((\mathbb{F}_3)^3,(\mathbb{F}_3)^2)$ arises in such a way. Hence $\mathrm{Hom}((\mathbb{F}_3)^3,(\mathbb{F}_3)^2)$ contains 729 linear maps.

Lemma 3.6. *Let V, W be non-zero $\mathbb{F}$-vector spaces. Then*

$$\dim(\mathrm{Hom}(V,W)) = \dim(V)\cdot\dim(W).$$

Proof. Let $\mathcal{B}_V$ and $\mathcal{B}_W$ be respective bases for V and W. For each pair $(v,w) \in \mathcal{B}_V \times \mathcal{B}_W$ we define, for all $s \in \mathcal{B}_V$,

$$E_{v,w}(s) := \begin{cases} w, & \text{if } s = v.\\ 0_W, & \text{otherwise.}\end{cases}$$

By Lemma 3.4, each $E_{v,w}$ extends to a homomorphism $V \to W$, i.e.

$$E_{v,w}\left(\sum_{s\in\mathcal{B}_V} \lambda_s s\right) = \sum_{s\in\mathcal{B}_V} \lambda_s E_{v,w}(s) = \lambda_v w.$$

Assume that $\mathcal{B} := \{E_{v,w} : (v,w) \in \mathcal{B}_V \times \mathcal{B}_W\}$ is linearly dependent. Then there is a finite subset J of $\mathcal{B}_V \times \mathcal{B}_W$ and non-zero $\lambda_{v,w} \in \mathbb{F}$, for $(v,w) \in J$ such that

$$0 = \sum_{(v,w)\in J} \lambda_{v,w} E_{v,w}.$$

Choose $s \in \mathcal{B}_V$ such that $(s, w') \in J$, for some $w' \in \mathcal{B}_W$. Then

$$0_W = \sum_{(v,w)\in J} \lambda_{v,w} E_{v,w}(s) = \sum_{w:(s,w)\in J} \lambda_{s,w} w,$$

and so all $\lambda_{u,w} = 0$, a contradiction. Hence $\mathcal{B}$ is linearly independent in $\mathrm{Hom}(V,W)$. Thus $\dim(\mathrm{Hom}(V,W)) = \dim(V) \cdot \dim(W)$ follows, if either V or W are infinite-dimensional.

Henceforth let $\mathcal{B}_V$ and $\mathcal{B}_W$ be finite, and let $T \in \mathrm{Hom}(V,W)$. Then for each $v \in \mathcal{B}_V$ there are $\lambda_w \in \mathbb{F}$, for all $w \in \mathcal{B}_W$ such that

$$T(v) = \sum_{w\in\mathcal{B}_W} \lambda_w w = \sum_{w\in\mathcal{B}_W} \lambda_w E_{v,w}(v).$$

Thus $\mathcal{B}$ spans $\mathrm{Hom}(V,W)$ and hence is a basis. Thus the result follows. □

Definition 3.7. *For $\mathbb{F}$-vector spaces V, W and $T \in \mathrm{Hom}(V,W)$, we call*

(1) $\ker(T) := \{v \in V : T(v) = 0_W\}$ *the* **kernel** *of T and*
(2) $\mathrm{im}(T) := \{T(v) : v \in V\}$ *the* **image** *of T.*

Lemma 3.8. *Let V, W be $\mathbb{F}$-vector spaces and $T \in \mathrm{Hom}(V,W)$. Then* $\ker(T)$ *is a subspace of V and* $\mathrm{im}(T)$ *is a subspace of W.*

Proof. Since $T(0_V) = 0_W$, we have $0_V \in \ker(T)$. Next let $u, v \in \ker(T)$ and $\lambda \in \mathbb{F}$. Then $T(u+v) = T(u) + T(v) = 0_W + 0_W = 0_W$ and $T(\lambda v) = \lambda T(v) = \lambda 0_W = 0_W$. Hence $u + v, \lambda v \in \ker(T)$, and so $\ker(T)$ is a subspace of V. The rest is homework. □

Remark/Example 3.9. *(1) Note that T is injective if and only if* $\ker(T) = \{0_V\}$ *and T is surjective if and only if* $\mathrm{im}(T) = W$.

(2) If $T : V \to W$ is the zero map, then $\ker(T) = V$ *and* $\mathrm{im}(T) = \{0_W\}$. *If $V \subseteq W$ and $I : V \to W : v \mapsto v$, then* $\ker(I) = \{0_V\}$ *and* $\mathrm{im}(I) = V$.

(3) Let $T : \mathbb{F}[X] \to \mathbb{F}[X] : f \mapsto f'$. If $\mathbb{F} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$, then $\mathrm{im}(T) = \mathbb{F}[X]$ *and* $\ker(T) = \{aX^0 : a \in \mathbb{F}\}$.

(4) Let V be an $\mathbb{F}$-vector space with subspaces U and W such that $V = U \oplus W$. Then $T : V \to V : u + w \mapsto w$ is a linear map, with $\ker(T) = U$ *and* $\mathrm{im}(T) = W$.

Definition 3.10. *Let $T \in \mathrm{Hom}(V,W)$ for finite-dimensional $\mathbb{F}$-vector spaces V and W. We call*

(1) *the dimension of* $\ker(T)$, *the* **nullity** *of* T *and write* $\text{null}(T)$.
(2) *the dimension of* $\text{im}(T)$, *the* **rank** *of* T *and write* $\text{rank}(T)$.

Theorem 3.11. *Let* $T \in \text{Hom}(V, W)$ *for finite-dimensional* $\mathbb{F}$*-vector spaces* V *and* W. *Then*

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

Proof. Let $\{v_1, \ldots, v_p\}$ be a basis of $\ker(T)$ and extend it to a basis $\{v_1, \ldots, v_n\}$ of V. Note that the result follows if we prove that $\mathcal{B} := \{T(v_{p+1}), \ldots, T(v_n)\}$ is a basis of $\text{im}(T)$. For $w \in \text{im}(T)$ there is $v \in V$ such that $T(v) = w$ and there are $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ such that $v := \sum_{i=1}^n \lambda_i v_i$. Then

$$w = T(v) = T\left(\sum_{i=1}^n \lambda_i v_i\right) = \sum_{i=1}^n \lambda_i T(v_i) = \sum_{i=p+1}^n \lambda_i T(v_i).$$

Hence $\mathcal{B}$ spans $\text{im}(T)$. If $w = 0_W$ in the above equation, then $v \in \ker(T)$ and so v is a linear combination of $\{v_1, \ldots, v_p\}$. Hence $\lambda_i = 0$, for $i = p+1, \ldots, n$. Thus $\mathcal{B}$ is linearly independent. Overall $\mathcal{B}$ is a basis of $\text{im}(T)$. $\square$

Example 3.12. *Let* $A = \begin{pmatrix} 5 & 1+3i & -i \\ 1-3i & 2 & \frac{3-i}{5} \end{pmatrix} \in \mathcal{M}_{2\times 3}(\mathbb{C})$ *from Example 3.2. What are* $\ker(T_A)$ *and* $\text{im}(T_A)$? *We study* $Ax = v$, *for* $x = (x_1, x_2, x_3) \in \mathbb{C}^3$ *and* $v = (a, b) \in \mathbb{C}^2$

$$(A|v) = \begin{pmatrix} 5 & 1+3i & -i & | & a \\ 1-3i & 2 & \frac{3-i}{5} & | & b \end{pmatrix} \overset{R2-\left(\frac{1-3i}{5}\right)R1}{\to} \begin{pmatrix} 5 & 1+3i & -i & | & a \\ 0 & 0 & 0 & | & b - \left(\frac{1-3i}{5}\right)a \end{pmatrix}$$

We have $\ker(T_A) = \{x \in \mathbb{C}^3 : T_A(x) = 0\} = \{x \in \mathbb{C}^3 : Ax = 0\} = \{(x_1, x_2, x_3) \in \mathbb{C}^3 : 5x_1 + (1+3i)x_2 - ix_3 = 0\}$, *which is a plane in* $\mathbb{C}^3$. *Hence* $\dim(\ker(T_A)) = 2$, *i.e.* $\text{null}(T_A) = 2$. *Also note that* $Ax = v$ *has a solution if and only if* $b - \left(\frac{1-3i}{5}\right)a = 0$. *Hence* $\text{im}(T_A) = \{(a, b) \in \mathbb{C}^2 : b = \left(\frac{1-3i}{5}\right)a\} = \mathbb{C}(1, \frac{1-3i}{5}) = \mathbb{C}(5, 1-3i)$, *which is a line in* $\mathbb{C}^2$. *Hence* $\dim(\text{im}(T_A)) = 1$, *i.e.* $\text{rank}(T_A) = 1$.

Lemma 3.13. *Let* V, W *be* $\mathbb{F}$*-vector spaces,* $T \in \text{Hom}(V, W)$ *and* S *a subset of* V. *Set* $T(S) := \{T(s) : s \in S\}$. *Then*

(1) *If* T *is injective and* S *is linearly independent in* V, *then* $T(S)$ *is linearly independent in* W.

(2) *If* T *is surjective and* S *spans* V, *then* $T(S)$ *spans* W.

Proof. (1) For $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ and $s_1, \ldots, s_n \in S$ let $\lambda_1 T(s_1) + \ldots + \lambda_n T(s_n) = 0$. Then $T(\lambda_1 s_1 + \ldots + \lambda_n s_n) = 0$, as T is linear. Then $\lambda_1 s_1 + \ldots + \lambda_n s_n = 0$, as T is injective. Then $\lambda_1 = \ldots = \lambda_n = 0$, as S is linearly independent. Overall $T(S)$ is linearly independent.

(2) Let $w \in W$. As T is surjective, there is $v \in V$ such that $T(v) = w$. Since S spans V, we have $v = \lambda_1 s_1 + \ldots + \lambda_n s_n$, for some $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ and some $s_1, \ldots, s_n \in S$. Now $w = \lambda_1 T(s_1) + \ldots + \lambda_n T(s_n) \in \operatorname{span}(T(S))$. In particular, $T(S)$ spans W. □

Remark 3.14. *Let V, W, U be $\mathbb{F}$-vector spaces. For $T \in \mathcal{F}(V, W)$ and $S \in \mathcal{F}(W, U)$ we have the composition $TS : V \to U : v \mapsto T(S(v))$. If S and T are homomorphism, then so is TS. Furthermore composition satisfies*

(i) $(RS)T = R(ST)$

(ii) $R(S + T) = RS + RT$ *and* $(R + S)T = RT + ST$

(iii) $R(\lambda S) = (\lambda R)S = \lambda(RS)$

where R, T, S are functions between $\mathbb{F}$-vector spaces, such that the compositions make sense.

We call a function $T \in \mathcal{F}(V, W)$ **invertible** *if there is some $S \in \mathcal{F}(W, V)$ such that $ST = I_V$ and $TS = I_W$, where I_V and I_W are the identity maps on V and W, respectively. In this case, S is unique, we denote it by T^{-1} and call it the* **inverse** *of T. Otherwise $T \in \mathcal{F}(V)$ is called* **non-invertible**. *It is well-know that T is invertible if and only if T is both injective and surjective.*

Lemma 3.15. *Let V, W be $\mathbb{F}$-vector spaces. If $T \in \operatorname{Hom}(V, W)$ is invertible, then $T^{-1} \in \operatorname{Hom}(W, V)$.*

Proof. Let $w, u \in W$. Then

$$w + u = T(T^{-1}(w)) + T(T^{-1}(u)) = T(T^{-1}(w)) + T^{-1}(u)).$$

Taking T^{-1} of both sides, gives $T^{-1}(w + u) = T^{-1}(w) + T^{-1}(u)$. It remains to show that $T^{-1}(\lambda w) = \lambda T^{-1}(w)$, for all $\lambda \in \mathbb{F}$. □

Theorem 3.16. *Let V, W be $\mathbb{F}$-vector space with respective dimensions n and m and let $T \in \operatorname{Hom}(V, W)$. Then the following are equivalent:*

(1) T is bijective, i.e. invertible

(2) $n = m$ and T is injective

(3) $n = m$ and T is surjective

Proof. Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis of V. Note that (1) implies (2) and (3), where $n = m$ follows from Theorem 3.11. Next let (2) be true. Then $\{T(v_1), \ldots, T(v_n)\}$ is linearly independent, by Lemma 3.13, and hence a basis of W, by Lemma 2.31. Thus for every $w \in W$ there are $\mu_1, \ldots, \mu_n \in \mathbb{F}$ so that

$$w = \mu_1 T(v_1) + \ldots + \mu_n T(v_n) = T(\mu_1 v_1 + \ldots + \mu_n v_n).$$

Hence T is also surjective. In particular, (2) implies (1).

Now let (3) be true and assume that $v \in \ker(T)$. If $v \neq 0_V$, then by Steinitz, $\{v, v_2, \ldots, v_n\}$ spans V. As T is surjective, $\{T(v), T(v_2), \ldots, T(v_n)\}$ spans W, by Lemma 3.13, and thus is a basis of W, by Lemma 2.31. But $T(v) = 0_W$ cannot be part of a basis. Hence $v = 0_V$ and $\ker T = \{0_V\}$, i.e T is also injective. Therefore (3) implies (1). $\square$

Definition 3.17. *Let V, W be $\mathbb{F}$-vector spaces. A bijective (i.e invertible) linear map $T : V \to W$ is called an* **isomorphism** *between V and W. In this case we say that V and W are* **isomorphic** *(as $\mathbb{F}$-vector spaces) and write $V \cong W$ or $V \cong_{\mathbb{F}} W$.*

The homomorphisms in $\mathcal{F}(V)$ are called **endomorphisms** *and we write $\mathrm{End}(V) := \mathrm{Hom}(V, V)$. Invertible endomorphisms are called* **automorphisms** *and we write $\mathrm{Aut}(V)$.*

Remark/Example 3.18. *(1) Note that $\dim(\mathrm{End}(V)) = \dim(V)^2$. Also there are three operations on $\mathrm{End}(V)$: (i) vector addition $S+T$, (ii) scalar multiplication λT and (iii) composition TS. One calls such an object an $\mathbb{F}$-algebra.*

(2) Let $S = \{s_1, \ldots, s_n\}$. Then $\mathcal{F}(S, \mathbb{F}) \cong \mathbb{F}^n$, via $T(f) = (f(s_1), \ldots, f(s_n))$. Check that $T \in \mathrm{Hom}(\mathcal{F}(S, \mathbb{F}), \mathbb{F}^n)$ and T is injective and surjective.

(3) $\mathbb{F}^\infty \cong \mathcal{F}(\mathbb{N}^*, \mathbb{F})$ via the map $(x_1, x_2, x_3, \ldots) \mapsto (f : \mathbb{N}^* \to \mathbb{F} : i \mapsto x_i)$.

(4) $\mathbb{F}[X]$ is isomorphic to the subspace in $\mathcal{F}(\mathbb{N}, \mathbb{F})$ of those functions $f : \mathbb{N} \to \mathbb{F}$ such that $f(n) = 0$ for all but finitely many $n \in \mathbb{N}$, via $f \mapsto \sum_{i=0}^{\infty} f(i)X^i$.

Definition 3.19. *Let V be an $\mathbb{F}$-vector space with basis $\mathcal{B} = \{v_1, \ldots, v_n\}$. For every $v \in V$ there are unique $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ such that $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$. We define*

$$v_{\mathcal{B}} := \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{F}^n,$$

called **column vector** *of v with respect to $\mathcal{B}$.*

Lemma 3.20. *Let V be an n-dimensional $\mathbb{F}$-vector space with basis $\mathcal{B}$. Then $T : V \to \mathbb{F}^n : v \mapsto v_{\mathcal{B}}$ is an isomorphism.*

Proof. Homework! $\square$

Theorem 3.21. *Let V, W be finite-dimensional $\mathbb{F}$-vector spaces. Then $V \cong W$ if and only if $\dim(V) = \dim(W)$.*

Proof. "$\Rightarrow$": Let $\{v_1, \ldots, v_n\}$ be a basis of V. If $T : V \to W$ is bijective, then $\{T(v_1), \ldots, T(v_n)\}$ is a basis of W, by Lemma 3.13, and thus $\dim(V) = \dim(W)$.

"$\Leftarrow$": If $\dim(V) = \dim(W)$, then there are isomorphisms $T : V \to \mathbb{F}^n$ and $S : W \to \mathbb{F}^n$, by Lemma 3.20, and so $S^{-1}T : V \to W$ is an isomorphism. $\square$

Lemma 3.22. *Let V be an n-dimensional $\mathbb{F}$-vector space with basis $\mathcal{B}$. Furthermore let $u_1, \dots, u_n$ be vectors in V. Then the following are equivalent:*

(1) $\{u_1, \dots, u_n\}$ *is a basis of* V

(2) the matrix $A := ((u_1)_{\mathcal{B}} \dots (u_n)_{\mathcal{B}}) \in \mathcal{M}_{n\times n}(\mathbb{F})$ *is invertible*

(3) $\det(A) \neq 0$

Proof. "(1) $\Leftrightarrow$ (2)": One checks that $u_1, \dots, u_n$ are linearly independent in V if and only if $(u_1)_{\mathcal{B}} \dots (u_n)_{\mathcal{B}}$ are linearly independent in $\mathbb{F}^n$. The latter holds if and only if the REF of A is the identity matrix, i.e. A is invertible.
"(2) $\Leftrightarrow$ (3)": standard fact about matrices. $\square$

Example 3.23. *Let $\mathbb{F} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{F}_{p^n}, p \neq 2\}$. Then the subspace $P_2 = \{a + bX + cX^2 : \ a, b, c \in \mathbb{F}\}$ of $\mathbb{F}[X]$ has the basis $\{1, X, X^2\}$. For the set $\{u_1 = 1 + X, u_2 = 1 + X^2, u_3 = X + X^2\}$ we have*

$$A = ((u_1)_{\mathcal{B}} \ (u_2)_{\mathcal{B}} \ (u_3)_{\mathcal{B}})] = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

We have $\det(A) = -2 \neq 0$. *So A is invertible and $\{u_1, u_2, u_3\}$ is a basis of P_2.*

4. Linear Maps and Matrices

Definition 4.1. *Let V, W be finite-dimensional $\mathbb{F}$-vector spaces with respective bases $\mathcal{B} := \{v_1, \dots, v_n\}$ and $\mathcal{C} := \{w_1, \dots, w_m\}$. Also let $T \in \mathrm{Hom}(V, W)$. Then, for all $j = 1, \dots, n$, there are unique $a_{ij} \in \mathbb{F}$, for $i = 1, \dots, m$ such that*

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

Then $M_{\mathcal{B},\mathcal{C}}(T) := (a_{ij}) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \in \mathcal{M}_{m\times n}(\mathbb{F})$ *is called the matrix of T with respect to the bases $\mathcal{B}$ and $\mathcal{C}$.*

Remark 4.2. *Note that*

$$T\left(\sum_{j=1}^{n} \lambda_j v_j\right) = \sum_{j=1}^{n} \lambda_j T(v_j) = \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \lambda_j a_{ij}\right) w_i.$$

Hence

$$T\left(\sum_{j=1}^{n} \lambda_j v_j\right) = \sum_{i=1}^{m} \mu_i w_i \Leftrightarrow \sum_{j=1}^{n} \lambda_j a_{ij} = \mu_i, \forall i \Leftrightarrow M_{\mathcal{B},\mathcal{C}}(T) \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix}.$$

Thus

$$\ker(T) = \left\{ \sum_{j=1}^{n} \lambda_j v_j : \ M_{\mathcal{B},\mathcal{C}}(T) \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = 0 \right\}$$

$$\operatorname{im}(T) = \left\{ \sum_{i=1}^{m} \mu_i w_i : \ M_{\mathcal{B},\mathcal{C}}(T) \cdot x = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix}, \ \textit{for some } x \in \mathbb{F}^n \right\}$$

Example 4.3. *(1) Let* $V = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$ *with basis* $\mathcal{B} = \{(1, 0, -1), (0, 1, -1)\}$ *and* $W = \{(x, y, z, w) \in \mathbb{R}^4 : x + y + z + w = 0\}$ *with basis* $\mathcal{C} = \{w_1 := (1, 0, 0, -1), w_2 := (0, 1, 0, -1), w_3 := (0, 0, 1, -1)\}$. *Next let*

$$T(x, y, z) = (x - 2y - z, 2x - y - z, -x - y, -6x - 2z),$$

and check that $T \in \operatorname{Hom}(V, W)$. *Now*

$$\begin{aligned} T(1, 0, -1) &= (2, 3, -1, -4) = 2w_1 + 3w_2 - w_3 \\ T(0, 1, -1) &= (-1, 0, -1, 2) = -w_1 - w_3. \end{aligned}$$

Hence

$$\left(\begin{array}{c|c} & \mu_1 \\ M_{\mathcal{B},\mathcal{C}}(T) & \mu_2 \\ & \mu_3 \end{array} \right) = \left(\begin{array}{cc|c} 2 & -1 & \mu_1 \\ 3 & 0 & \mu_2 \\ -1 & -1 & \mu_3 \end{array} \right) \to \left(\begin{array}{cc|c} 0 & -3 & \mu_1 + 2\mu_3 \\ 0 & -3 & \mu_2 + 3\mu_3 \\ -1 & -1 & \mu_3 \end{array} \right)$$

$$\to \left(\begin{array}{cc|c} 0 & 0 & \mu_1 - \mu_2 - \mu_3 \\ 0 & -3 & \mu_2 + 3\mu_3 \\ -1 & -1 & \mu_3 \end{array} \right) \to \left(\begin{array}{cc|c} 1 & 1 & -\mu_3 \\ 0 & 1 & -\frac{\mu_2}{3} - \mu_3 \\ 0 & 0 & \mu_1 - \mu_2 - \mu_3 \end{array} \right)$$

So $\ker(T) = \{0_V\}$. *Also* $\mu_1 w_1 + \mu_2 w_2 + \mu_3 w_3 \in \operatorname{im}(T)$ *if and only if* $\mu_1 = \mu_2 + \mu_3$. *Hence*

$$\begin{aligned} \operatorname{im}(T) &= \{(\mu_2 + \mu_3) w_1 + \mu_2 w_2 + \mu_3 w_3 : \ \mu_2, \mu_3 \in \mathbb{R}\} \\ &= \{\mu_2 (w_1 + w_2) + \mu_3 (w_1 + w_3) : \ \mu_2, \mu_3 \in \mathbb{R}\} \end{aligned}$$

(2) Let $A = (a_{ij}) \in \mathcal{M}_{m \times n}(\mathbb{F})$. *Then* $T_A : \mathbb{F}^n \to \mathbb{F}^m : v \mapsto Av$ *is linear, by Example 3.2(4). Now*

$$T_A(e_j) = A e_j = \sum_{i=1}^{m} a_{ij} e_i,$$

for all $j = 1, \ldots, n$. *Thus* $M_{\mathcal{SB}_n, \mathcal{SB}_m}(T_A) = A$, *where* $\mathcal{SB}_k$ *denotes the standard bases in* $\mathbb{F}^k$.

Theorem 4.4. *Let V, W be finite-dimensional $\mathbb{F}$-vector spaces with bases $\mathcal{B}$ and $\mathcal{C}$ and dimensions n and m, respectively. Then*

$$M_{\mathcal{B},\mathcal{C}} : \mathrm{Hom}(V, W) \to \mathcal{M}_{m\times n}(\mathbb{F}) : T \mapsto M_{\mathcal{B},\mathcal{C}}(T),$$

is an isomorphism of $\mathbb{F}$-vector spaces.

Proof. Let $\mathcal{B} := \{v_1, \ldots, v_n\}$ and $\mathcal{C} := \{w_1, \ldots, w_m\}$ and set $\Delta := M_{\mathcal{B},\mathcal{C}}$. For $T, S \in \mathrm{Hom}(V, W)$ we have

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i, \quad \text{and} \quad S(v_j) = \sum_{i=1}^{m} b_{ij} w_i,$$

for all $j = 1, \ldots, n$. So $\Delta(T) = (a_{ij})$ and $\Delta(S) = (b_{ij})$. Next, for $\lambda \in \mathbb{F}$,

$$(T + S)(v_j) = T(v_j) + S(v_j) = \sum_{i=1}^{m} (a_{ij} + b_{ij}) w_i, \quad \text{and}$$

$$(\lambda T)(v_j) = \lambda T(v_j) = \sum_{i=1}^{m} (\lambda a_{ij}) w_i,$$

for all $j = 1, \ldots, n$. Hence $\Delta(T + S) = \Delta(T) + \Delta(S)$ and $\Delta(\lambda T) = \lambda \Delta(T)$, i.e. Δ is a homomorphism.

Next observe that if $\Delta(T) = \Delta(S)$, then T and S are identical on $\mathcal{B}$, and so by Lemma 3.4, $T = S$. Hence Δ is injective. As $\mathrm{Hom}(V, W)$ and $\mathcal{M}_{m\times n}(\mathbb{F})$ have both dimension mn, it follows with Theorem 3.16 that Δ is bijective. $\square$

Remark 4.5. *Let V, W, U be finite-dimensional $\mathbb{F}$-vector spaces with respective bases $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$. Also let $n = \dim(V)$.*

(1) Let $T \in \mathrm{Hom}(V, W)$ and $S \in \mathrm{Hom}(W, U)$. Then a careful calculation shows that

$$M_{\mathcal{B},\mathcal{D}}(ST) = M_{\mathcal{C},\mathcal{D}}(S) \cdot M_{\mathcal{B},\mathcal{C}}(T).$$

(2) In the case $V = W$ and $\mathcal{B} = \mathcal{C}$, we write $M_{\mathcal{B}}(T)$ for the matrix of $T \in \mathrm{End}(V)$ with respect to $\mathcal{B}$. In particular, $\mathrm{End}(V) \cong \mathcal{M}_n(\mathbb{F})$.

(3) Let $T \in \mathrm{End}(V)$. Then $M_{\mathcal{B}}(T)$ is the identity matrix I_n, i.e the $n \times n$-matrix with ones on the main diagonal and zeros elsewhere, if and only if T is the identity map id_V on V.

Theorem 4.6. *(Change of Bases) Let V, W be finite-dimensional $\mathbb{F}$-vector spaces, such that $\mathcal{B}$ and $\mathcal{B}'$ are bases of V and $\mathcal{C}$ and $\mathcal{C}'$ are bases of W. Also let $T \in \mathrm{Hom}(V, W)$. Then there matrices $X \in \mathcal{M}_{\dim(W)}(\mathbb{F})$ and $Y \in \mathcal{M}_{\dim(V)}(\mathbb{F})$ such that*

$$M_{\mathcal{B},\mathcal{C}}(T) = X \cdot M_{\mathcal{B}',\mathcal{C}'}(T) \cdot Y.$$

In fact, $X = M_{\mathcal{C}',\mathcal{C}}(\mathrm{id}_W)$ and $Y = M_{\mathcal{B},\mathcal{B}'}(\mathrm{id}_V)$, which are thus invertible.

Proof. This follows from Remark 4.5 and the fact that $T = \mathrm{id}_W \circ T \circ \mathrm{id}_V$. $\square$

Theorem 4.7. *Let V, W be n-dimensional $\mathbb{F}$-vector spaces with respective bases $\mathcal{B}$, $\mathcal{C}$ and let $T \in \mathrm{Hom}(V, W)$. Then T is invertible if and only if $M_{\mathcal{B},\mathcal{C}}(T)$ is invertible. In this case $M_{\mathcal{C},\mathcal{B}}(T^{-1}) = M_{\mathcal{B},\mathcal{C}}(T)^{-1}$.*

Proof. "$\Rightarrow$": If T is invertible, then $T^{-1} \in \mathrm{Hom}(W, V)$. Then $I_n = M_{\mathcal{B}}(T^{-1}T) = M_{\mathcal{C},\mathcal{B}}(T^{-1}) \cdot M_{\mathcal{B},\mathcal{C}}(T)$ and $I_n = M_{\mathcal{C}}(TT^{-1}) = M_{\mathcal{B},\mathcal{C}}(T) \cdot M_{\mathcal{C},\mathcal{B}}(T^{-1})$.

"$\Leftarrow$": Let $X \in \mathcal{M}_n(\mathbb{F})$ be the inverse of $M_{\mathcal{B},\mathcal{C}}(T)$. By Theorem 4.4 there is some $S \in \mathrm{Hom}(W, V)$ such that $M_{\mathcal{C},\mathcal{B}}(S) = X$. Now

$$M_{\mathcal{B},\mathcal{B}}(ST) = M_{\mathcal{C},\mathcal{B}}(S) \cdot M_{\mathcal{B},\mathcal{C}}(T) = X \cdot M_{\mathcal{B},\mathcal{C}}(T) = I.$$

Thus ST is the identity map on V. Analogous, TS is the identity map on W. $\square$

Example 4.8. *(1) The standard basis $\mathcal{SB} := \{e_1, e_2\}$ and $\mathcal{B} := \{(1,1), (1,2)\}$ are two bases of $\mathbb{R}^2$. Next let $T(x, y) = (y, x)$, for all $(x, y) \in \mathbb{R}^2$. The $T \in \mathrm{Hom}(\mathbb{R}^2, \mathbb{R}^2)$. We have*

$$M_{\mathcal{B},\mathcal{SB}}(T) = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}, \ M_{\mathcal{B},\mathcal{SB}}(\mathrm{id}_{\mathbb{R}^2}) = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \text{ and}$$

$$M_{\mathcal{SB},\mathcal{B}}(\mathrm{id}_{\mathbb{R}^2}) = M_{\mathcal{B},\mathcal{SB}}(\mathrm{id}_{\mathbb{R}^2})^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

Hence

$$M_{\mathcal{SB},\mathcal{B}}(T) = M_{\mathcal{SB},\mathcal{B}}(\mathrm{id}_{\mathbb{R}^2}) \cdot M_{\mathcal{B},\mathcal{SB}}(T) \cdot M_{\mathcal{SB},\mathcal{B}}(\mathrm{id}_{\mathbb{R}^2}) = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$$

Indeed, $T(e_1) = e_2 = -(1,1) + (1,2)$ and $T(e_2) = e_1 = 2(1,1) - (1,2)$.

(2) Note the subspace $P_2 = \{a + bX + cX^2 : \ a, b, c \in \mathbb{R}\}$ of $\mathbb{R}[X]$, with basis $\mathcal{B} := \{1, X, X^2\}$ and the subspace $\mathrm{Sym}_2(\mathbb{R}) := \left\{ \begin{pmatrix} a & b \\ b & c \end{pmatrix} : \ a, b, c \in \mathbb{R} \right\}$ of $\mathcal{M}_2(\mathbb{R})$, with basis $\mathcal{C} := \left\{ E_1 := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_2 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, E_3 := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$. Next let $T \in \mathrm{Hom}(P_2, \mathrm{Sym}_2(\mathbb{R}))$ be the linear map uniquely defined by

$$T(1) := E_1, \quad T(X) := 2E_1 + E_2, \quad T(X^2) := 3E_1 + 2E_2 + E_3,$$

that is, $T(a + bX + cX^2) = \begin{pmatrix} a + 2b + 3c & b + 2c \\ b + 2c & c \end{pmatrix}$. Hence

$$M_{\mathcal{B},\mathcal{C}}(T) = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{with inverse} \quad \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

So T is invertible and

$$T^{-1}\begin{pmatrix} a & b \\ b & c \end{pmatrix} = a \cdot T^{-1}(E_1) + b \cdot T^{-1}(E_2) + c \cdot T^{-1}(E_3)$$
$$= a \cdot 1 + b \cdot (-2 + X) + c \cdot (1 - 2X + X^2)$$
$$= (a - 2b + c) + (b - 2c)X + cX^2$$

Check that indeed, $T^{-1}\begin{pmatrix} a + 2b + 3c & b + 2c \\ b + 2c & c \end{pmatrix} = a + bX + cX^2$.

Remark 4.9. *Let V be an n-dimensional $\mathbb{F}$-vector space and $f \in \mathrm{End}(V)$.*

(1) If $X = M_{\mathcal{B},\mathcal{B}'}(\mathrm{id}_V)$, for bases $\mathcal{B}$ and $\mathcal{B}'$ of V, then by Theorems 4.6 and 4.7,

$$M_{\mathcal{B}}(T) = X^{-1} \cdot M_{\mathcal{B}'}(T) \cdot X.$$

(2) Let $A, B \in \mathcal{M}_n(\mathbb{F})$. We say A is a **diagonal** *matrix if all non-zero entries of A lie on the main diagonal. We call A and B* **similar** *if there is some invertible $X \in \mathcal{M}_n(\mathbb{F})$ such that $A = X^{-1} \cdot B \cdot X$. We call A* **diagonalisable** *if A is similar to a diagonal matrix.*

(3) We call T **diagonalisable**, *if there is a basis $\mathcal{B}$ such that $M_{\mathcal{B}}(T)$ is a diagonal matrix. In general this is not possible. Take for instance the linear map $T : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \mapsto (0, x)$. Note that $T \circ T = 0$. Now assume that $M_{\mathcal{B}}(T) = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$, for some basis $\mathcal{B}$, where $\alpha, \beta \in \mathbb{F}$. Then $M_{\mathcal{B}}(T \circ T)$ equals both the zero matrix and $(M_{\mathcal{B}}(T))^2 = \begin{pmatrix} \alpha^2 & 0 \\ 0 & \beta^2 \end{pmatrix}$. As this forces $\alpha = \beta = 0$, we get $T = 0$, which is false.*

(4) Note that T is diagonalisable if and only if $M_{\mathcal{B}}(T)$ is diagonalisable for any basis $\mathcal{B}$. In this case there is an invertible $X \in \mathcal{M}_n(\mathbb{F})$ and a diagonal matrix $D \in \mathcal{M}_n(\mathbb{F})$ such that

$$M_{\mathcal{B}}(T) = X^{-1}DX = X^{-1} \cdot \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix} \cdot X.$$

Then for all integers $k \geq 1$,

$$M_{\mathcal{B}}(T^k) = (X^{-1}DX)^k = X^{-1}D^kX = X^{-1} \cdot \begin{pmatrix} \alpha_1^k & & 0 \\ & \ddots & \\ 0 & & \alpha_n^k \end{pmatrix} \cdot X.$$

This allows for a quick way to determine powers of T.

Definition 4.10. *Let V be a finite-dimensional $\mathbb{F}$-vector space and $T \in \operatorname{End}(V)$. We call $\lambda \in \mathbb{F}$ an* **eigenvalue** *of T if there exists a non-zero $v \in V$ such that $T(v) = \lambda v$. In this case v is called* **eigenvector** *of T with respect to λ. Finally*

$$E(T, \lambda) := \{v \in V : \ T(v) = \lambda v\},$$

denotes the **eigenspace** *of T with respect to λ.*

Lemma 4.11. *Let V be a finite-dimensional $\mathbb{F}$-vector space, $T \in \operatorname{End}(V)$ and $\lambda \in \mathbb{F}$ an eigenvalue of T. Then $E(T, \lambda)$ is a subspace of V.*

Proof. Homework! $\square$

Theorem 4.12. *Let V be a finite-dimensional $\mathbb{F}$-vector space, $T \in \operatorname{End}(V)$ and $\lambda \in \mathbb{F}$. Then the following are equivalent:*

(1) λ is an eigenvalue of T
(2) $T(v) = \lambda v$, for some non-zero $v \in V$
(3) $(T - \lambda \operatorname{id}_V)(v) = 0_V$, for some non-zero $v \in V$
(4) $\ker(T - \lambda \operatorname{id}_V)$ is non-trivial
(5) $T - \lambda \operatorname{id}_V$ is not injective
(6) $T - \lambda \operatorname{id}_V$ is not bijective
(7) $X := M_{\mathcal{B}}(T - \lambda \operatorname{id}_V)$ is not invertible, for any basis $\mathcal{B}$ of V
(8) $\det(X) = 0$
(9) λ is an eigenvalue of $S^{-1}TS$, for all $S \in \operatorname{Aut}(V)$.

Proof. We have (6) $\Rightarrow$ (5) by Theorem 3.16 and (6) $\Leftrightarrow$ (7) by Theorem 4.7. Next we prove (1) $\Rightarrow$ (9). If (1), then $T(v) = \lambda v$, for a non-zero $v \in V$. Next let $S \in \operatorname{Aut}(V)$. Then $S^{-1}(v) \neq 0_V$ and $(S^{-1}TS)(S^{-1}(v)) = S^{-1}(T(v)) = S^{-1}(\lambda v) = \lambda(S^{-1}(v))$, i.e. $S^{-1}(v)$ is an eigenvector w.r.t. the eigenvalue λ of $S^{-1}TS$. Thus (9) holds. Finally (9) $\Rightarrow$ (1) follows with $S = \operatorname{id}_V$. $\square$

Example 4.13. *(1) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $T(x, y) = (-y, x + y)$, i.e $T \in \operatorname{End}_{\mathbb{F}}(\mathbb{F}^2)$. We work w.r.t the standard basis $\mathcal{SB}$. Then $T(1, 0) = (0, 1)$ and $T(0, 1) = (-1, 1)$. Now*

$$X := \mathcal{M}_{\mathcal{SB}}(T - \lambda \operatorname{id}) = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} -\lambda & -1 \\ 1 & 1 - \lambda \end{pmatrix}$$

and so $\det(X) = (-\lambda)(1 - \lambda) - ((-1) \cdot 1) = \lambda^2 - \lambda + 1$. Hence $\det(X) = 0$ if and only if $\lambda = \frac{1 \pm \sqrt{3}i}{2}$. Thus if $\mathbb{F} = \mathbb{R}$, then $\lambda^2 - \lambda + 1 = 0$ has no real solutions and so T has no eigenvalues. If $\mathbb{F} = \mathbb{C}$, then $\lambda_{1|2} = \frac{1 \pm \sqrt{3}i}{2}$ are the eigenvalues of T. Next we calculate $E(T, \lambda) = \ker(T - \lambda \operatorname{id})$, for $\lambda := \lambda_i$,

$$\begin{pmatrix} -\lambda & -1 \\ 1 & 1 - \lambda \end{pmatrix} \overset{R1 \leftrightarrow R2}{\to} \begin{pmatrix} 1 & 1 - \lambda \\ -\lambda & -1 \end{pmatrix} \overset{R2 + \lambda R1}{\to} \begin{pmatrix} 1 & 1 - \lambda \\ 0 & -1 + \lambda - \lambda^2 \end{pmatrix} = \begin{pmatrix} 1 & 1 - \lambda \\ 0 & 0 \end{pmatrix}$$

So $E(T,\lambda)=\{(x,y)\in\mathbb{C}^2 : x=(\lambda-1)y\}=\{(x,y)\in\mathbb{C}^2 : \lambda x=y\}=\mathbb{C}(\lambda,1)$.

(2) *For* $\mathbb{F}=\mathbb{R}$ *find all eigenvalues of* $T : P_3 \to P_3$, *where* $T(a+bX+cX^2+dX^3)$ $=(a+13d)+(-25a+7b+11c-6d)X+(18a+c+5d)X^2+(-2d)X^3$. *W.r.t the standard basis* $\mathcal{B}=\{1,X,X^2,X^3\}$ *we have*

$$A := M_{\mathcal{B}}(T) = \begin{pmatrix} 1 & 0 & 0 & 13 \\ -25 & 7 & 11 & -6 \\ 18 & 0 & 1 & 5 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

Next

$$\begin{aligned}\det(A-\lambda I_4) &= \det\begin{pmatrix} 1-\lambda & 0 & 0 & 13 \\ -25 & 7-\lambda & 11 & -6 \\ 18 & 0 & 1-\lambda & 5 \\ 0 & 0 & 0 & -2-\lambda \end{pmatrix} \\ &= (-2-\lambda)\cdot\det\begin{pmatrix} 1-\lambda & 0 & 0 \\ -25 & 7-\lambda & 11 \\ 18 & 0 & 1-\lambda \end{pmatrix} \\ &= (-2-\lambda)\cdot(7-\lambda)\cdot\det\begin{pmatrix} 1-\lambda & 0 \\ 18 & 1-\lambda \end{pmatrix} \\ &= (-2-\lambda)\cdot(7-\lambda)\cdot(1-\lambda)^2\end{aligned}$$

Hence T *has three eigenvalues* -2, 7 *and* 1. *For* $\lambda=1$, *we have*

$$A-\lambda I_4 = A-I_4 = \begin{pmatrix} 0 & 0 & 0 & 13 \\ -25 & 6 & 11 & -6 \\ 18 & 0 & 0 & 5 \\ 0 & 0 & 0 & -3 \end{pmatrix} \to \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 6 & 11 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Its nullspace is $\mathbb{R}(0,11,-6,0)$, *i.e.* $E(T,1)=\mathbb{R}(11X-6X^2)$. *The nullspaces of* $A+2I_4$ *and* $A-7I_4$ *are* $\mathbb{R}(39,370,-219,-9)$ *and* $\mathbb{R}(0,1,0,0)$, *respectively. So* $E(T,-2)=\mathbb{R}(39+370X-219X^2-9X^3)$ *and* $E(T,7)=\mathbb{R}(X)$.

Theorem 4.14. *Let* V *be a finite-dimensional* $\mathbb{F}$*-vector space,* $T\in\operatorname{End}(V)$ *and* $\lambda_1,\ldots,\lambda_k$ *distinct eigenvalues of* T. *Then*

$$E(T,\lambda_1)+\ldots+E(T,\lambda_k)=E(T,\lambda_1)\oplus\ldots\oplus E(T,\lambda_k)$$

Proof. The statement holds if $k=1$. Next let $k>1$ and assume the statement holds for $k-1$. Let $v\in X := E(T,\lambda_k)\cap(E(T,\lambda_1)+\ldots+E(T,\lambda_{k-1}))$, i.e. $v=v_1+\ldots+v_{k-1}$, for $v_i\in E(T,\lambda_i)$. Then

$$\lambda_k v_1+\ldots+\lambda_k v_{k-1}=\lambda_k v=T(v)=\lambda_1 v_1+\ldots+\lambda_{k-1}v_{k-1}.$$

Then $\lambda_k v_i = \lambda_i v_i$, for all $i = 1, \dots, k-1$, and since $\lambda_k \neq \lambda_i$, we get $v_i = 0$, for all $i = 1, \dots, k-1$. Hence $X = \{0_V\}$ and so the statement holds for k. $\square$

Corollary 4.15. *Let V be a finite-dimensional $\mathbb{F}$-vector space, $T \in \mathrm{End}(V)$, $\lambda_1, \dots, \lambda_k$ distinct eigenvalues of T and $v_1, \dots, v_k$ corresponding eigenvectors. Then $\{v_1, \dots, v_k\}$ are linearly independent. In particular, T has at most $\dim(V)$ distinct eigenvalues.*

Theorem 4.16. *Let V be a finite-dimensional $\mathbb{F}$-vector space and $T \in \mathrm{End}(V)$. Then the following are equivalent:*

(1) T is diagonalisable.

(2) V has a basis of eigenvectors of T.

(3) V is the direct sum of all eigenspaces of T.

(4) $\dim(V)$ is the sum of the dimensions of all eigenspaces of T.

Proof. "(1) $\Rightarrow$ (2)": Note that $M_{\mathcal{B}}(T) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, for some basis $\mathcal{B} = \{v_1, \dots, v_n\}$. Then $T(v_i) = \lambda_i v_i$, for all $i = 1, \dots, n$, i.e. v_i is an eigenvector of T for the eigenvalue λ_i.

"(2) $\Rightarrow$ (3)": Let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of T. Then by assumption $V = E(T, \lambda_1) + \dots + E(T, \lambda_k)$. Now (3) follows from Theorem 4.14.

"(3) $\Rightarrow$ (4)": trivial

"(4) $\Rightarrow$ (1)": Let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of T and let

$$\dim(V) = \sum_{i=1}^{k} \dim(E(T, \lambda_i)) = \dim[E(T, \lambda_1) \oplus \dots \oplus E(T, \lambda_k)]$$

Hence $V = E(T, \lambda_1) \oplus \dots \oplus E(T, \lambda_k)$. Now let $v_1, \dots, v_{s_1}$ be a basis of $E(T, \lambda_1)$, $v_{s_1+1}, \dots, v_{s_2}$ be a basis of $E(T, \lambda_2)$ and so on, then $\mathcal{B} := \{v_1, \dots, v_n\}$ is a basis of V. Then

$$M_{\mathcal{B}}(T) = \begin{pmatrix} \lambda_1 I_{s_1} & & & 0 \\ & \lambda_2 I_{s_2} & & \\ & & \ddots & \\ 0 & & & \lambda_k I_{s_k} \end{pmatrix}$$

$\square$

Corollary 4.17. *Let V be a finite-dimensional $\mathbb{F}$-vector space, $T \in \mathrm{End}(V)$ diagonalisable and $\mathcal{B}$ a basis of V of eigenvectors. Then $M_{\mathcal{B}}(T)$ is a diagonal matrix.*

Corollary 4.18. *Let V be an n-dimensional $\mathbb{F}$-vector space and let $T \in \operatorname{End}(V)$ have n distinct eigenvalues. Then T is diagonalisable.*

Example 4.19. *(1) Recall $T \in \operatorname{End}(P_3)$ from Example 4.13(2). We found that T has three eigenvalues with one-dimensional eigenspaces each. As $\dim(P_3) = 4$, it follows that T is not diagonalisable.*

(2) Let $T : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \mapsto (y, x + y)$ and show that it has eigenvalues $\lambda_{1|2} = \frac{1\pm\sqrt{5}}{2}$ and eigenspaces $E(T, \lambda_i) = \mathbb{C}(1, \lambda_i)$. Note that $\mathcal{B} := \{(1, \lambda_1), (1, \lambda_2)\}$ is a basis of $\mathbb{R}^2$ of eigenvectors of T. Furthermore $T(1, \lambda_i) = (\lambda_i, 1 + \lambda_i) = \lambda_i(1, \lambda_i)$, as $\lambda_i^2 = 1 + \lambda_i$. Also let $\mathcal{SB} = \{e_1, e_2\}$ be the standard basis of $\mathbb{R}^2$. Then

$$D := M_{\mathcal{B}}(T) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \; X := M_{\mathcal{B},\mathcal{SB}}(\mathrm{id}) = \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}$$

$$X^{-1} = M_{\mathcal{SB},\mathcal{B}}(\mathrm{id}) = \frac{-1}{\sqrt{5}} \cdot \begin{pmatrix} \lambda_2 & -1 \\ -\lambda_1 & 1 \end{pmatrix} \quad \textit{and} \quad M_{\mathcal{SB}}(T) = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix},$$

and so $X^{-1} \cdot M_{\mathcal{SB}}(T) \cdot X = M_{\mathcal{B}}(T)$ or $M_{\mathcal{SB}}(T) = X \cdot M_{\mathcal{B}}(T) \cdot X^{-1}$.

(3) The **Fibonacci-numbers** *are defined by*

$$F_0 := 0,\; F_1 := 1,\; F_{n+1} := F_n + F_{n-1},$$

for all integers $n \geq 1$. Hence $(F_n)_{n\geq 0} = (0, 1, 1, 2, 3, 5, 8, 13, 21, \ldots)$. We seek an explicit formula for F_n. We have

$$\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}}_{=:A} \begin{pmatrix} F_{n-1} \\ F_n \end{pmatrix} = \ldots = A^n \begin{pmatrix} F_0 \\ F_1 \end{pmatrix} = A^n \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Note that by (2) we have $A^n = (XDX^{-1})^n = XD^nX^{-1}$ and so

$$\begin{aligned} \begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} &= A^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} = XD^n \begin{pmatrix} \lambda_2 & -1 \\ -\lambda_1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = X \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \\ &= \frac{-1}{\sqrt{5}} \cdot \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix} \begin{pmatrix} -\lambda_1^n \\ \lambda_2^n \end{pmatrix} = \frac{1}{\sqrt{5}} \cdot \begin{pmatrix} \lambda_1^n - \lambda_2^n \\ \lambda_1^{n+1} - \lambda_2^{n+1} \end{pmatrix} \end{aligned}$$

Overall,

$$F_n = \frac{1}{\sqrt{5}} \cdot \left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}} \cdot \left(\frac{1-\sqrt{5}}{2}\right)^n.$$

Remark 4.20. *Let V be a $\mathbb{C}$-vector space and $T \in \operatorname{End}_{\mathbb{C}}(V)$. For any basis $\mathcal{B} = \{b_1, \ldots, b_n\}$ of V the equation $\det(M_{\mathcal{B}}(T - \lambda\,\mathrm{id}) = 0$ has at least one* *solution $\lambda \in \mathbb{C}$. Hence T has at least one eigenvalue λ_1 with eigenvector v_1. By Steinitz we have that $\mathcal{B}_1 = \{v_1, b_2, \ldots, b_n\}$ is a basis of V. Note that*

$$M_{\mathcal{B}_1}(T) = \left(\begin{array}{c|c} \lambda_1 & \star \\ \hline 0_{n-1\times 1} & \star \end{array}\right)$$

Next set $W := \operatorname{span}(b_2, \ldots, b_n)$. Then $\mathbb{C}^n = \mathbb{C}^n v_1 \oplus W$. Furthermore the map

$$\begin{array}{cccccc} \hat{T}: & W & \to & \mathbb{C}^n = \mathbb{C}v_1 \oplus W & \to & W \\ & w & \mapsto & T(w) = \alpha v_1 + w' & \mapsto & w' \end{array}$$

is an endomorphism on W. As before $\hat{T}$ has at least one eigenvalue λ_2 with eigenvector v_2. Note that λ_2 may or may not equal λ_1, but $\{v_1, v_2\}$ are linearly independent. Now by Steinitz $\mathcal{B}_1 = \{v_1, v_2, b_3, \ldots, b_n\}$ is a basis of V. Next note that there are $\alpha_i \in \mathbb{C}$ such that

$$T(v_2) = \alpha_1 v_1 + \sum_{i=2}^{n} \alpha_i b_i = \alpha_1 v_1 + \hat{T}(v_2) = \alpha_1 v_1 + \lambda_2 v_2.$$

Hence

$$M_{\mathcal{B}_2}(T) = \left(\begin{array}{cc|c} \lambda_1 & \star & \\ 0 & \lambda_2 & \star \\ \hline \multicolumn{2}{c|}{0_{n-2\times 2}} & \star \end{array}\right)$$

Now set $W := \operatorname{span}(b_3, \ldots, b_n)$ and repeat. In this way we construct a basis $\mathcal{B}$ of V such that $M_{\mathcal{B}}(T)$ is an upper-triangular matrix, i.e. a matrix with only zero entries below the main diagonal. The elements on the main diagonal are the eigenvalues of T with possible repetition. In particular, $M_{\mathcal{B}}(T)$ is invertible if and only if zero is not an eigenvalue of T.

5. Inner Product Spaces

Throughout this section let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Recall **complex conjugation** $\mathbb{C} \to \mathbb{C} : z = a + bi \mapsto \overline{z} := a - bi$, which is a **field automorphism** of $\mathbb{C}$ with fixed field $\mathbb{R}$. This means it is a bijection on $\mathbb{C}$, and $\overline{z+w} = \overline{z} + \overline{w}$ and $\overline{zw} = \overline{z}\,\overline{w}$, for all $z, w \in \mathbb{C}$, and $\overline{z} = z$ if and only if $z \in \mathbb{R}$. Also recall that the modulus $|z|$ of z is defined by $|z|^2 = z\overline{z}$.

Definition 5.1. *Let V be a finite-dimensional vector space over $\mathbb{F}$. An* **inner product** *on V is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ such that*

- *(i) $\langle v, v \rangle$ is a non-negative real number, for all $v \in V$, and $\langle v, v \rangle = 0$ if and only if $v = 0$, (i.e. $\langle \cdot, \cdot \rangle$ is positive definite)*
- *(ii) $\langle au + bv, w \rangle = a\langle u, w \rangle + b\langle v, w \rangle$ for all $a, b \in \mathbb{F}$ and $u, v, w \in V$, (i.e. $\langle \cdot, \cdot \rangle$ is linear in the first variable)*
- *(iii) $\langle v, w \rangle = \overline{\langle w, v \rangle}$ for all $v, w \in V$. (i.e. $\langle \cdot, \cdot \rangle$ is conjugate symmetric)*

The pair $(V, \langle \cdot, \cdot \rangle)$ *is called an* **inner product space**. *Furthermore we call* $u, v \in V$ **orthogonal** *if* $\langle u, v \rangle = 0$.

Remark/Example 5.2. *(1) For the remainder let* V *denote an inner product space* $(V, \langle \cdot, \cdot \rangle)$.

(2) $\langle u, av + bw \rangle = \overline{a}\langle u, v \rangle + \overline{b}\langle u, w \rangle$, *i.e.* $\langle \cdot, \cdot \rangle$ *is conjugate linear in the second variable*

(3) $\langle u, v \rangle = 0$ *if and only if* $\langle v, u \rangle = 0$

(4) $\langle 0_V, v \rangle = \langle v, 0_V \rangle = 0$.

(5) The **standard inner product** *on* $\mathbb{F}^n$ *is the* **dot-product**

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = x_1\overline{y_1} + \cdots + x_n\overline{y_n}.$$

Note that any two distinct elements e_i *and* e_j *of the standard basis of* $\mathbb{F}^n$ *are orthogonal. Also for instance* $(2, 1)$ *and* $(-1, 2)$ *are orthogonal in* $\mathbb{R}^2$.

(6) For real numbers $a < b$, *let* $V = \mathcal{C}[a, b]$ *be the* $\mathbb{R}$*-vector space of real valued continuous functions on the interval* $[a, b]$. *Then* $\langle f, g \rangle = \int_a^b f(x)g(x)\,dx$ *is an inner product on* V. *For instance* $f(x) = x$ *and* $g(x) = 1$ *are orthogonal in* $\mathcal{C}[-1, 1]$, *but not in* $\mathcal{C}[0, 1]$.

(7) Let $V = \mathcal{M}_n(\mathbb{F})$, *for some integer* $n \geq 1$. *The* **trace** $\mathrm{tr}(X)$ *of* $X \in V$ *is the sum of the elements on the main diagonal of* X. *Then* $\langle A, B \rangle = \mathrm{tr}(A\overline{B}^t)$, *for* $A, B \in V$, *is an inner product on* V, *where* B^t *denotes the transpose of* B *and* $\overline{B}$ *is the matrix obtained by conjugating the entries in* B, *e.g*

$$\left\langle \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ i & 3-i \end{pmatrix} \right\rangle = \mathrm{tr}\left(\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -i \\ 0 & 3+i \end{pmatrix} \right) = \mathrm{tr}\begin{pmatrix} 1 & -i \\ 1 & 3 \end{pmatrix} = 4$$

Lemma 5.3. *Let* $u, v \in V$, *with* $v \neq 0_V$. *Then there are unique* $\lambda \in \mathbb{F}$ *and* $w \in V$ *so that* $u = \lambda v + w$ *and* $\langle v, w \rangle = 0$. *Here* $\lambda = \frac{\langle u, v \rangle}{\langle v, v \rangle}$ *and* $w = u - \lambda v$.

Proof. First let $u = \lambda v + w$, for $\lambda \in \mathbb{F}$ and $w \in V$, where $\langle w, v \rangle = 0$. Then

$$\langle u, v \rangle = \lambda\langle v, v \rangle + \langle w, v \rangle = \lambda\langle v, v \rangle.$$

Hence $\lambda = \frac{\langle u, v \rangle}{\langle v, v \rangle}$. This gives uniqueness of λ and w. Conversely, with this value of λ, and $w := u - \lambda v$, we have $\langle w, v \rangle = \langle u, v \rangle - \lambda\langle v, v \rangle = 0$. □

Definition 5.4. *The function* $||v|| := \sqrt{\langle v, v \rangle}$, *for* $v \in V$, *is called the* **norm** *associated with* $(V, \langle \cdot, \cdot \rangle)$.

Remark 5.5. *(1) One may think of* $||v||$ *has the length of the vector* v. *Take for instance the* $\mathbb{R}$*-vector space* $V = \mathbb{R}^3$ *with the standard inner product and let* $v = (x, y, z) \in V$. *Then* $||v|| = \sqrt{\langle v, v \rangle} = \sqrt{x^2 + y^2 + z^2}$.

(2) $||\lambda v|| = |\lambda| \cdot ||v||$, *for all* $\lambda \in \mathbb{F}$ *and* $v \in V$.

(3) Let $u, v \in V$. *Then*

$$\begin{aligned} ||u+v||^2 &= \langle u+v, u+v\rangle = \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle \\ &= ||u||^2 + \langle u,v\rangle + \overline{\langle u,v\rangle} + ||v||^2 = ||u||^2 + 2\,\mathrm{Re}(\langle u,v\rangle) + ||v||^2 \end{aligned}$$

Note that if u *and* v *are orthogonal, then* $||u+v||^2 = ||u||^2 + ||v||^2$, *i.e we have Pythagoras' Theorem.*

Theorem 5.6. *Let* $u, v \in V$. *Then*

(1) (Cauchy-Schwarz Inequality)

$$|\langle u,v\rangle| \le ||u|| \cdot ||v||,$$

with equality if and only if one of u, v *is a scalar multiple of the other.*

(2) (Triangle Inequality)

$$||u+v|| \le ||u|| + ||v||,$$

with equality if and only if one of u, v *is a non-negative real multiple of the other.*

(3) (Parallelogram Equality)

$$||u+v||^2 + ||u-v||^2 = 2(||u||^2 + ||v||^2)$$

Proof. (1) Trivial, if $v = 0_V$. Hence assume $v \neq 0_V$. Then $u = \frac{\langle u,v\rangle}{\langle v,v\rangle} \cdot v + w$, where $\langle v,w\rangle = 0$, by Lemma 5.3. Then by Pythagoras' Theorem,

$$||u||^2 = \frac{\langle u,v\rangle^2}{||v||^4} \cdot ||v||^2 + ||w||^2 \ge \frac{\langle u,v\rangle^2}{||v||^2}.$$

Hence the inequality follows and equality holds if and only if $w = 0$, i.e u is a scalar multiple of v.

(2) Note that $\mathrm{Re}(\langle u,v\rangle) \le |\langle u,v\rangle| \le ||u|| \cdot ||v||$, with equality if and only if one of u, v is a non-negative real multiple of the other. The statement follows as

$$\begin{aligned} ||u+v||^2 &= ||u||^2 + 2\,\mathrm{Re}(\langle u,v\rangle) + ||v||^2 \le ||u||^2 + 2||u|| \cdot ||v|| + ||v||^2 \\ &= (||u|| + ||v||)^2. \end{aligned}$$

(3) Exercise. $\square$

Definition 5.7. *A tuple* $(v_1, \ldots, v_k)$ *be non-zero vectors in* V *is called*

(1) **orthogonal** *if* $\langle v_i, v_j\rangle = 0$, *for all* $i, j = 1, \ldots, n$ *with* $j \neq j$

(2) **orthonormal** *if* $\langle v_i, v_j\rangle = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$, *for all* $i, j = 1, \ldots, n$.

Lemma 5.8. *Every orthogonal tuple of vectors in V is linearly independent. In particular, every orthonormal tuple of vectors in V is linearly independent.*

Proof. Let $(v_1, \ldots, v_k)$ be an orthogonal tuple of vectors in V. Also let $\lambda_i \in \mathbb{F}$ such that $\lambda_1 v_1 + \ldots + \lambda_k v_k = 0_V$. Then for all $j = 1, \ldots, k$,

$$0 = \langle \sum_{i=1}^{k} \lambda_i v_i, v_j \rangle = \sum_{i=1}^{n} \lambda_i \langle v_i, v_j \rangle = \lambda_j ||v_j||^2$$

As $v_j \neq 0_V$, we have $\lambda_j = 0$. Hence the claim follows. $\square$

Definition 5.9. *A basis of V of orthogonal / orthonormal vectors is called* **orthogonal / orthonormal basis**.

Example 5.10. *(1) Every orthonormal basis is an orthogonal basis.*

(2) Consider $\mathbb{F}^n$ equipped with the standard inner product. Then $(e_1, \ldots, e_n)$ is an orthonormal basis in $\mathbb{F}^n$.

(3) Consider $(V = \mathcal{C}[0,1], \langle \cdot, \cdot \rangle)$, where $\langle f, g \rangle = \int_0^1 f(x)g(x)\, dx$, for $f, g \in V$. Then $\left(1,\ 2x-1,\ x^2 - x + \frac{1}{6}\right)$ is an orthogonal tuple.

(4) Consider $(V = \mathcal{M}_2(\mathbb{F}), \langle \cdot, \cdot \rangle)$, where $\langle A, B \rangle = \operatorname{tr}(A\overline{B}^t)$, for $A, B \in V$. The standard basis $(E_{1,1},\ E_{1,2},\ E_{2,1},\ E_{2,2})$ is an orthonormal basis, while $\left(I_2, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right)$ is an orthogonal basis of V.

Theorem 5.11. *(Gram-Schmidt process) If $(v_1, \ldots, v_k)$ is a linearly independent tuple of vectors in V, then there is an orthonormal tuple $(u_1, \ldots, u_k)$ in V such that for all $i = 1, \ldots, k$,*

$$\operatorname{sp}(v_1, \ldots, v_i) = \operatorname{sp}(u_1, \ldots, u_i).$$

Proof. If $k = 1$ we set $u_1 = \frac{v_1}{\|v_1\|}$. Check that $\langle u_1, u_1 \rangle = 1$ and $\operatorname{sp}(u_1) = \operatorname{sp}(v_1)$. Next let $k > 1$ and assume the statement holds for $k - 1$. Set

$$\widehat{u_k} = v_k - \sum_{j=1}^{k-1} \langle v_k, u_j \rangle u_j.$$

Note that $\widehat{u_k} \neq 0$, as otherwise $v_k \in \operatorname{sp}(u_1, \ldots, u_{k-1}) = \operatorname{sp}(v_1, \ldots, v_{k-1})$. Then set $u_k = \frac{\widehat{u_k}}{||\widehat{u_k}||}$. Check that $(u_1, \ldots, u_k)$ is an orthonormal tuple and $\operatorname{sp}(v_1, \ldots, v_k) = \operatorname{sp}(u_1, \ldots, u_k)$. $\square$

Corollary 5.12. *Every inner-product space has an orthonormal basis.*

Proof. Let $(v_1, \ldots, v_n)$ be any basis of V. Applying Gram-Schmidt produces an orthonormal tuple $(u_1, \ldots, u_n)$ in V, which is linearly independent and thus a basis of V. $\square$

Example 5.13. *In $\mathbb{R}^4$, equipped with the standard inner product, consider the vectors $v_1 = (4, 2, -2, -1)$, $v_2 = (2, 2, -4, -5)$ and $v_3 = (0, 8, -2, 5)$ and set $V = \operatorname{span}(v_1, v_2, v_3)$. We have $\langle v_1, v_1 \rangle = 16 + 4 + 4 + 1 = 25$ and so $||v_1|| = \sqrt{25} = 5$. Set $u_1 = \frac{v_1}{||v_1||} = \frac{1}{5} \cdot (4, 2, -2, -1)$. Next*

$$\langle v_2, u_1 \rangle = \frac{1}{5} \cdot (8 + 4 + 8 + 5) = \frac{25}{5} = 5$$

and so

$$\widehat{u_2} = v_2 - \langle v_2, u_1 \rangle u_1 = (2, 2, -4, -5) - (4, 2, -2, -1) = (-2, 0, -2, -4)$$

As $\langle \widehat{u_2}, \widehat{u_2} \rangle = 4 + 0 + 4 + 16 = 24$, we have $||\widehat{u_2}|| = \sqrt{24}$. Set $u_2 = \frac{\widehat{u_2}}{||\widehat{u_2}||} = \frac{1}{\sqrt{24}} \cdot (-2, 0, -2, -4)$. Next

$$\langle v_3, u_1 \rangle = \frac{1}{5} \cdot (0 + 16 + 4 - 5) = \frac{15}{5} = 3$$
$$\langle v_3, u_2 \rangle = \frac{1}{\sqrt{24}} \cdot (0 + 0 + 4 - 20) = -\frac{16}{\sqrt{24}}$$

Hence

$$\begin{aligned}
\widehat{u_3} &= v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2 \\
&= (0, 8, -2, 5) - \frac{3}{5} \cdot (4, 2, -2, -1) + \frac{2}{3}(-2, 0, -2, -4) \\
&= \frac{1}{15} \cdot \big(15 \cdot (0, 8, -2, 5) - 9 \cdot (4, 2, -2, -1) + 10 \cdot (-2, 0, -2, -4)\big) \\
&= \frac{1}{15} \cdot \big(-56, 102, -32, 44\big) = \frac{2}{15} \cdot \big(-28, 51, -16, 22\big)
\end{aligned}$$

Finally $\langle \widehat{u_3}, \widehat{u_3} \rangle = \frac{4}{225} \cdot (784 + 2601 + 256 + 484) = \frac{16500}{225} = \frac{220}{3}$ and so $||\widehat{u_3}|| = \sqrt{\frac{220}{3}}$. Set $u_3 = \frac{\widehat{u_3}}{||\widehat{u_3}||} = \frac{2\sqrt{3}}{15\sqrt{220}} \cdot (-28, 51, -16, 22)$. Hence (u_1, u_2, u_3) is an orthonormal basis of V.

Lemma 5.14. *Let V have an orthonormal basis $(u_1, \dots, u_n)$. Then for every $v \in V$,*

$$v = \sum_{i=1}^{n} \langle v, u_i \rangle u_i$$

Proof. There are $\alpha_i \in \mathbb{F}$ such that $v = \sum_{i=1}^{n} \alpha_i u_i$. Then for all $i = 1, \dots, n$,

$$\langle v, u_i \rangle = \langle \sum_{j=1}^{n} \alpha_j u_j, u_i \rangle = \sum_{j=1}^{n} \alpha_j \langle u_j, u_i \rangle = \alpha_i \cdot ||u_i||^2 = \alpha_i$$

□

Example 5.15. *We continue Example 5.13. Take* $v = v_1 + v_2 + v_3 = (6, 12, -8, -1)$*. Then*

$$\langle v, u_1 \rangle = \frac{1}{5} \cdot (24 + 24 + 16 + 1) = 13$$

$$\langle v, u_2 \rangle = \frac{1}{\sqrt{24}} \cdot (-12 + 16 + 4) = \frac{8}{\sqrt{24}} = \sqrt{\frac{8}{3}}$$

$$\langle v, u_3 \rangle = \frac{2\sqrt{3}}{15\sqrt{220}} \cdot (-168 + 612 + 128 - 22) = \frac{2\sqrt{3}}{15\sqrt{220}} \cdot 550 = \sqrt{\frac{220}{3}}$$

Hence $v = 13u_1 + \sqrt{\frac{8}{3}} u_2 + \sqrt{\frac{220}{3}} u_3$.

Lemma 5.16. *(Parseval's Identity) Let* $(u_1, \dots, u_n)$ *be an orthonormal basis of* V*. Then, for all* $v, w \in V$*,*

$$\langle v, w \rangle = \sum_{i=1}^{n} \langle v, u_i \rangle \overline{\langle w, u_i \rangle}$$

Proof. We have $v = \sum_{i=1}^{n} \langle v, u_i \rangle u_i$ and $w = \sum_{i=1}^{n} \langle w, u_i \rangle u_i$, by Lemma 5.14. Then

$$\begin{aligned} \langle v, w \rangle &= \left\langle \sum_{i=1}^{n} \langle v, u_i \rangle u_i, \sum_{j=1}^{n} \langle w, u_j \rangle u_j \right\rangle = \sum_{i,j=1}^{n} \langle v, u_i \rangle \overline{\langle w, u_j \rangle} \langle u_i, u_j \rangle \\ &= \sum_{i=1}^{n} \langle v, u_i \rangle \overline{\langle w, u_i \rangle} \end{aligned}$$

□

Theorem 5.17. *For* $w \in V$ *we define* $\varphi_w : V \to \mathbb{F} : v \mapsto \langle v, w \rangle$*. Then* $\varphi_w \in V^* := \mathrm{Hom}(V, F)$*. Furthermore* $\phi : V \to V^* : w \mapsto \varphi_w$*, is a bijection.*

Proof. We leave showing that φ_w is a homomorphism as an exercise. Next let $\varphi \in V^*$ and let $(u_1, \dots, u_n)$ be an orthonormal basis of V. Set

$$w := \sum_{i=1}^{n} \overline{\varphi(u_i)} u_i.$$

Then $\langle w, u_i \rangle = \overline{\varphi(u_i)}$, for all $i = 1, \dots, n$, by Lemma 5.14. Furthermore for every $v \in V$ we have $v = \sum_{i=1}^{n} \langle v, u_i \rangle u_i$. Now

$$\begin{aligned} \varphi(v) &= \varphi\left(\sum_{i=1}^{n} \langle v, u_i \rangle u_i \right) = \sum_{i=1}^{n} \langle v, u_i \rangle \varphi(u_i) = \sum_{i=1}^{n} \langle v, u_i \rangle \overline{\langle w, u_i \rangle} \\ &\underset{\text{Identity}}{\overset{\text{Parseval's}}{=}} \langle v, w \rangle = \varphi_w(v). \end{aligned}$$

Hence ϕ is surjective. Next assume that $\phi(w) = \phi(w')$, for $w, w' \in V$. Then, for all $v \in V$, we have $\langle v, w\rangle = \langle v, w'\rangle$, and so $\langle v, w - w'\rangle = 0$. In particular, $\langle w - w', w - w'\rangle = 0$ and so $w - w' = 0_V$, i.e. $w = w'$. Thus ϕ is injective. $\square$

Definition 5.18. *Let $T \in \mathrm{End}(V)$. Then there is some $T^* \in \mathrm{End}(V)$ such that $\langle T(v), u\rangle = \langle v, T^*(u)\rangle$, for all $v, u \in V$. We call T^* the* **adjoint** *of T.*

Proof. For $T \in \mathrm{End}(V)$ and $u \in V$, let $\varphi : V \to \mathbb{F} : v \mapsto \langle T(v), u\rangle$. Then $\varphi \in V^*$ (prove!) and so there is a unique $w \in V$ such that $\varphi = \varphi_w$, i.e $\langle T(v), u\rangle = \varphi_w(v) = \langle v, w\rangle$. We define $T^* : V \to V : u \mapsto w$. In particular, $\langle T(v), u\rangle = \langle v, T^*(u)\rangle$, for all $v, u \in V$. It remains to show that $T^* \in \mathrm{End}(V)$. For all $v \in V$, $u, u' \in V$ and $\lambda \in \mathbb{F}$ we have

$$\begin{aligned}\langle v, T^*(u + u')\rangle &= \langle T(v), u + u'\rangle = \langle T(v), u\rangle + \langle T(v), u'\rangle \\ &= \langle v, T^*(u)\rangle + \langle v, T^*(u')\rangle = \langle v, T^*(u) + T^*(u')\rangle\end{aligned}$$

$$\langle v, T^*(\lambda u)\rangle = \langle T(v), \lambda u\rangle = \overline{\lambda}\langle T(v), u\rangle = \overline{\lambda}\langle v, T^*(u)\rangle = \langle v, \lambda T^*(u)\rangle$$

Hence $T^*(u+u') = T^*(u)+T^*(u')$ and $T^*(\lambda u) = \lambda T^*(u)$, i.e. $T^* \in \mathrm{End}(V)$. $\square$

Lemma 5.19. *Let $S, T \in \mathrm{End}(V)$ with adjoints S^*, T^*.*

(a) $(T^*)^* = T$
(b) $(S + T)^* = S^* + T^*$
(c) $(\lambda T)^* = \overline{\lambda}T^*$, *for all* $\lambda \in \mathbb{F}$
(d) $(ST)^* = T^*S^*$

Proof. Exercise. $\square$

Remark 5.20. *Let $A \in \mathcal{M}_n(\mathbb{F})$. We call $A^* := \overline{A}^t$ the* **Hermitian transpose** *of A. Next let $T \in \mathrm{End}(V)$ and $\mathcal{B} = (u_1, \ldots, u_n)$ an orthonormal basis of V. Set $a_{i,j} := (M_{\mathcal{B}}(T))_{i,j}$. Since*

$$T(u_j) = \sum_{i=1}^{n} \langle T(u_j), u_i\rangle u_i,$$

for all $j = 1, \ldots, n$, by Lemma 5.14, it follows that $a_{i,j} = \langle T(u_j), u_i\rangle$. Hence the (i, j)-entry in $M_{\mathcal{B}}(T^*)$ is given by

$$\langle T^*(u_j), u_i\rangle = \overline{\langle u_i, T^*(u_j)\rangle} = \overline{\langle T(u_i), u_j\rangle} = \overline{a_{ji}},$$

i.e. $M_{\mathcal{B}}(T^*)$ is the Hermitian transpose of $M_{\mathcal{B}}(T)$.

Example 5.21. *(1) Consider $V = \mathbb{C}^3$ with the standard inner product. Also let $T(x, y, z) = (x + (2 + i)y,\ x + y - z,\ x - 3iz)$, for all $(x, y, z) \in \mathbb{C}^3$.*

Then $T \in End(V)$. The standard basis $\mathcal{SB} = \{e_1, e_2, e_3\}$ is orthonormal. We have

$$M_{\mathcal{SB}}(T) = \begin{pmatrix} 1 & 2+i & 0 \\ 1 & 1 & -1 \\ 1 & 0 & -3i \end{pmatrix} \text{ and so } M_{\mathcal{SB}}(T^*) = \begin{pmatrix} 1 & 1 & 1 \\ 2-i & 1 & 0 \\ 0 & -1 & 3i \end{pmatrix}$$

Hence $T^*(a,b,c) = (a+b+c, (2-i)a+b, -b+3ic)$, for all $(a,b,c) \in \mathbb{C}^3$. Indeed,

$$\begin{aligned} \langle T(x,y,z), (a,b,c)\rangle &= (x+(2+i)y)\cdot \overline{a} + (x+y-z)\cdot \overline{b} + (x-3iz)\cdot \overline{c} \\ &= x\cdot(\overline{a}+\overline{b}+\overline{c}) + y\cdot((2+i)\overline{a}+\overline{b}) + z\cdot(-\overline{b}-3i\overline{c}) \\ &= x\cdot(\overline{a+b+c}) + y\cdot(\overline{(2-i)a+b}) + z\cdot(\overline{-b+3ic}) \\ &= \langle (x,y,z),\ T^*(a,b,c)\rangle \end{aligned}$$

(2) *Let $(V = \mathcal{M}_n(\mathbb{F}), \langle\cdot,\cdot\rangle)$, where $\langle A, B\rangle = \operatorname{tr}(A\overline{B}^t)$, for $A, B \in V$. For a fixed $A \in V$ set $T(X) := AXA^t$, for all $X \in V$. Show that $T \in \operatorname{End}(V)$. What is its adjoint T^*? For all $X, Y \in V$,*

$$\begin{aligned} \langle X, T^*(Y)\rangle = \langle T(X), Y\rangle = \langle AXA^t, Y\rangle &= \operatorname{tr}(AXA^t\overline{Y}^t) \\ &= \operatorname{tr}(XA^t\overline{Y}^tA), \qquad \text{as } \operatorname{tr}(RS) = \operatorname{tr}(SR) \\ &= \operatorname{tr}(X(A^t\overline{Y}A)^t), \qquad \text{as } (RS)^t = S^tR^t \\ &= \operatorname{tr}(X(\overline{\overline{A}^tY\overline{A}})^t) = \langle X, \overline{A}^tY\overline{A}\rangle \end{aligned}$$

Hence $T^*(Y) = \overline{A}^tY\overline{A}$.

Definition 5.22. *Let $T \in \operatorname{End}(V)$. We call T*

(1) **self-adjoint**, *if $T = T^*$*
(2) **normal**, *if $TT^* = T^*T$*
(3) **unitary**, *if $TT^* = T^*T = \mathrm{id}_V$*

Remark/Example 5.23. *(1) If $T \in \operatorname{End}(V)$ is self-adjoint or unitary, then T is normal.*

(2) *If $T \in \operatorname{End}(V)$ is unitary, then T is invertible and $T^{-1} = T^*$.*

(3) *$T : \mathbb{C}^2 \to \mathbb{C}^2 : (x,y) \mapsto (x+iy, -ix+y)$ is self-adjoint, as $M_{\mathcal{SB}}(T) = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} = M_{\mathcal{SB}}(T)^* = M_{\mathcal{SB}}(T^*)$*

(4) *$T : \mathbb{C}^2 \to \mathbb{C}^2 : (x,y) \mapsto (x+(2+3i)y, (2+3i)x+y)$ and $S : \mathbb{C}^2 \to \mathbb{C}^2 : (x,y) \mapsto (x+y, -x+y)$ are normal, but not self-adjoint or unitary. For instance $M_{\mathcal{SB}}(S) = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \neq M_{\mathcal{SB}}(S^*) = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ and*

$$M_{\mathcal{SB}}(S) \cdot M_{\mathcal{SB}}(S^*) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = M_{\mathcal{SB}}(S^*) \cdot M_{\mathcal{SB}}(S) \neq I_2.$$

(5) $T : \mathbb{C}^2 \to \mathbb{C}^2 : (x, y) \mapsto \frac{1}{3} \cdot (2x + (2 + i)y, (-2 + i)x + 2y)$ *and* $S : \mathbb{C}^3 \to \mathbb{C}^3 : v \mapsto iv$ *are unitary, but not self-adjoint.*

Theorem 5.24. *Let* $T \in \mathrm{End}(V)$ *with eigenvalue* $\lambda \in \mathbb{C}$.

(1) If T *is self-adjoint, then* $\lambda \in \mathbb{R}$

(2) If T *is unitary, then* λ *lies on the unit circle in* $\mathbb{C}$*, i.e.* $\lambda = |1|$.

(3) If T *is normal and* $\mathbb{F} = \mathbb{C}$*, then there exists some* $v \in E(T, \lambda)$ *such* $T^*(v) = \overline{\lambda}v$*, i.e.* $\overline{\lambda}$ *is an eigenvalue of* T^*.

Proof. Let $v \in V$ be an eigenvector of T w.r.t λ. Note that $\langle v, v\rangle > 0$.

(1) As T is self-adjoint,

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle T(v), v\rangle = \langle v, T^*(v)\rangle = \langle v, T(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle,$$

and so $\lambda = \overline{\lambda}$.

(2) As T is unitary,

$$\langle v, v\rangle = \langle v, T^*(T(v))\rangle = \langle T(v), T(v)\rangle = \langle \lambda v, \lambda v\rangle = \lambda\overline{\lambda}\langle v, v\rangle = |\lambda|^2\langle v, v\rangle,$$

and so $|\lambda| = 1$.

(3) As T is normal,

$$T(T^*(v)) = (TT^*)(v) = (T^*T)(v) = T^*(T(v)) = T^*(\lambda v) = \lambda T^*(v).$$

Thus $T^*(v) \in E(T, \lambda)$, i.e. T^* as an endomorphism on $E(T, \lambda)$. As such it has an eigenvalue $\mu \in \mathbb{C}$ with corresponding $v \in E(T, \lambda)$. For this v

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle T(v), v\rangle = \langle v, T^*(v)\rangle = \langle v, \mu v\rangle = \overline{\mu}\langle v, v\rangle,$$

and so $\overline{\mu} = \lambda$, that is, $\mu = \overline{\lambda}$. □

Example 5.25. *(1) In Example 5.23 (3),* T *has eigenvalues* 0 *and* 2

(2) In Example 5.23 (5), T *has eigenvalues* $\frac{2-\sqrt{5}\ i}{3}$ *and* $\frac{2+\sqrt{5}\ i}{3}$ *and* S *has eigenvalue* i*, all of which lie on the unit circle in* $\mathbb{C}$.

(3) In Example 5.23 (4), T *has eigenvalues* $-1 - 3i$ *and* $3 + 3i$ *and* T^* *has eigenvalues* $-1+3i$ *and* $3-3i$*. Furthermore both* S *and* S^* *have eigenvalues* $1 - i$ *and* $1 + i$.

Theorem 5.26. *Let* $T \in \mathrm{End}(V)$ *be normal and* $\mathbb{F} = \mathbb{C}$*. Then there exists an orthonormal basis of* V *consisting of eigenvectors of* T*. In particular,* T *is diagonalisable.*

Proof. We use induction on $n = \dim(V)$. The statement is trivial if $n = 1$. Next let $n > 1$ and assume the statement holds for $n-1$. Over $\mathbb{C}$, T has at least one eigenvalue $\lambda \in \mathbb{C}$. As seen in the proof of Theorem 5.24 (3) there is some $v_1 \in V$ such that $v_1 \in E(T(\lambda) \cap E(T^*, \overline{\lambda})$. We set $W := \mathbb{C}v_1$ and $W^\perp := \{v \in V : \langle v, w\rangle = 0, \text{ for all } w \in W\}$. One checks that $W^\perp$ is a subspace of V. Furthermore it follows from Lemma 5.3 that $V = W \oplus W^\perp$. Next for $v \in W^\perp$ we have

$$\langle T(v), v_1\rangle = \langle v, T^*(v_1)\rangle = \langle v, \lambda v_1\rangle = \overline{\lambda}\langle w, v_1\rangle = 0$$

Consequently, $T(v) \in W^\perp$ and we can consider T as a normal endomorphism on $W^\perp$. Next note that $\dim(W^\perp) = n-1$. Thus by induction there is an orthonormal basis $(v_2, \ldots, v_n)$ of $W^\perp$ consisting of eigenvectors of T. Now $(v_1, v_2, \ldots, v_n)$ is an orthonormal basis of V consisting of eigenvectors of T. $\square$

Corollary 5.27. *Let $\mathbb{F} = \mathbb{C}$ and $T \in \operatorname{End}(V)$ be self-adjoint or unitary. Then T is diagonalisable.*

Example 5.28. *(1) In Example 5.23 (3),*

$$M_{\mathcal{SB}}(T) = \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} \text{ is similar to } M_{\mathcal{B}}(T) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix},$$

where $\mathcal{B} = \{(1, -i), (1, i)\}$ is an orthonormal basis of $\mathbb{C}^2$ consisting of eigenvectors of T.

(2) In Example 5.23 (4),

$$M_{\mathcal{SB}}(T) = \begin{pmatrix} 1 & 2+3i \\ 2+3i & 1 \end{pmatrix} \text{ is similar to } M_{\mathcal{B}}(T) = \begin{pmatrix} -1-3i & 0 \\ 0 & 3+3i \end{pmatrix},$$

where $\mathcal{B} = \{(-1, 1), (1, 1)\}$, and

$$M_{\mathcal{SB}}(S) = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \text{ is similar to } M_{\mathcal{B}}(S) = \begin{pmatrix} 1-i & 0 \\ 0 & 1+i \end{pmatrix},$$

where $\mathcal{B} = \{(-i, 1), (i, 1)\}$.

(3) In Example 5.23 (5),

$$M_{\mathcal{SB}}(T) = \begin{pmatrix} \frac{2}{3} & \frac{-2+i}{3} \\ \frac{2+i}{3} & \frac{2}{3} \end{pmatrix} \text{ is similar to } M_{\mathcal{B}}(T) = \begin{pmatrix} \frac{2-\sqrt{5}\ i}{3} & 0 \\ 0 & \frac{2+\sqrt{5}\ i}{3} \end{pmatrix},$$

where $\mathcal{B} = \left\{\left(\frac{-1-2i}{\sqrt{5}}, 1\right), \left(\frac{1+2i}{\sqrt{5}}, 1\right)\right\}$.

Finally $M_{\mathcal{SB}}(S) = \begin{pmatrix} i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & i \end{pmatrix}$ *is diagonal w.r.t. the standard basis.*