1 Fields

Fields are an abstract structure that describes sets of "numbers" and their operations.

Definition 1.1 (Fields)
A set $\mathcal{F}$ together with two binary operations

- $+ : \mathcal{F} \times \mathcal{F} \to \mathcal{F}$ (Addition)
- $\cdot : \mathcal{F} \times \mathcal{F} \to \mathcal{F}$ (Multiplication)

$\forall x, y, z \in \mathcal{F}$

- $x + (y + z) = (x + y) + z$ (Associativity)
- $x(yz) = (xy)z$ (Associativity)
- $x + y = y + x$ (Commutatitivity)
- $xy = yx$ (Commutatitivity)
- $\exists 0 \in \mathcal{F}$ such that $x + 0 = x \ \forall x \in \mathcal{F}$ (Neutral additive element)
- $\exists 1 \in \mathcal{F}$ such that $x \cdot 1 = x \ \forall x \in \mathcal{F}$ (Neutral scalar multiplication element)
- $\forall x \in \mathcal{F} \ \exists -x \in \mathcal{F} \quad x + (-x) = 0$ (Additive Inverse)
- $\forall y \in \mathcal{F}\backslash\{0\} \exists y^{-1} \exists \mathcal{F} \quad yy^{-1} = 1$ (Multiplicative inverse)
- $x(y + z) = xy + xz$ (Distributivity)

Example of fields: Rational numbers $\mathbb{Q}$, real numbers $\mathbb{R}$ and complex numbers $\mathbb{C}$. Another example is the set $\mathcal{F} = \{0, 1\}$. The set $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$ is also an example.

Definition 1.2
For any field $\mathcal{F}$ we denote the set of n-tuples by $\mathcal{F}^n$. We define two important operations on $\mathcal{F}^n$.

- Addition: Given two elements $x = (x_1, \ldots, x_n)$ $y = (y_1, \ldots, y_n)$ $x + y = (x_1 + y_1, \ldots, x_n + y_n)$
- Scalar multiplication: Given an element $\lambda \in \mathcal{F}$ and a n-tuple $x \in \mathcal{F}^n$ we define $\lambda x := (\lambda x_1, \ldots, \lambda x_n)$

We often write $0 \in \mathcal{F}^n$ for the n-tuple consisting of n zeros.
Let $x \in \mathcal{F}^n$ we define $-x := (-x_1, \ldots, -x_n)$ and we see that $x + (-x) = 0$.

For any $x \in \mathcal{F}^n$ we have $0 + x = x$

2 Vector Spaces

Definition 2.1 (Vector Space)
Let V be a set and $\mathcal{F}$ be a field.

Let $+ : V \times V \to V$ and $\cdot : \mathcal{F} \times V \to V$ be two binary operations. We say V is a vector space (with respect to these operations) over $\mathcal{F}$, or an $\mathcal{F}$-vector space (VS) if

- Addition is commutative: $\forall u, v \in V \quad u + v = v + u$.
- Addition is associative: $\forall u, v, w \in V \quad u + (v + w) = (u + v) + w$.
- Multiplication is associative: $\forall \lambda, \mu \in \mathcal{F} \, \forall v \in V \quad (\lambda\mu)v = \lambda(\mu v)$.
- Neutral additive: $\exists 0 \in V$ such that $\forall v \in V \quad 0 + v = v$.
- Inverse addition: $\forall v \in V \exists - v \in V \quad v + (-v) = 0$.
- Neutral scalar multiplication: $1 \in \mathcal{F}$ it holds that $1 \cdot v = v \quad v \in V$.
- Distributivity: $\forall u, v \in V \ \forall \lambda\mu \in \mathcal{F} \quad \lambda(u + v) = \lambda u + \lambda v$ and $(\lambda + \mu)v = \lambda v + \mu v$.

Example 2.1
a) $\mathcal{F}^n$ is an $\mathcal{F}$-VS, it holds that $\forall n \in \mathbb{N}$ especially $\mathcal{F}$ is an $\mathcal{F}$-VS.

b) $V = \{0\} \subseteq \mathcal{F}$ is an $\mathcal{F}$-VS.

c) $\mathcal{F}^\infty := \{(x_1, x_2, \dots) : x_i \in \mathcal{F} \ i \in \mathbb{N}\}$, the set of all infinite sequences is an $\mathcal{F}$-VS

d) Let $V := \{f : S \to \mathcal{F}\}$ be the set of functions from a set S into $\mathcal{F}$ then V is an $\mathcal{F}$-VS with $f, g \in V$ for which $(f + g)(s) := f(s) + g(s) \forall s \in S$. Similiarly $\forall \lambda \in \mathcal{F} \quad (\lambda f)(s) = \lambda(f(s))$. Sometimes you will see this notation:

$$V = \mathcal{F}^S$$

Example. $\mathbb{R}^{[0,1]}$.

Theorem 2.1
Let V be an $\mathcal{F}$-VS. Then the additive neutral element is unique.

Proof 2.1
Suppose there is another additive neutral element: 0 and $0'$ are both neutral. Then

$$\begin{aligned} 0 &= 0 + 0' \text{ since } 0' \text{ is neutral} \\ &= 0' \text{ since } 0 \text{ is neutral} \end{aligned}$$

Hence $0 = 0'$ and there is an unique neutral.

∎

Theorem 2.2
Let V be an $\mathcal{F}$-VS. Then every element in V has a unique additive inverse.

Proof 2.2
Let $v \in V$ and suppose w and w' are both additive inverse for v.

$$w' = 0 + w' = (w + v) + w' = w + (v + w') = w + 0 = w$$

We will from now on decide the unique inverse of v be $-v$ and write $w + (-v) := w - v$.

Theorem 2.3
Let V be an $\mathcal{F}$-VS. Then $\forall v \in V$

$$0 \in \mathcal{F} \quad v = 0 \in V$$

Proof 2.3
We see that

$$0 \cdot v = (0 + 0)v = 0v + 0v$$

Add $-0v$ on both sides

$$0 = 0v.$$

$\blacksquare$

Theorem 2.4
Let V be an $\mathcal{F}$-VS. Then $\forall \lambda \in \mathcal{F}$

$$\lambda \cdot 0 = 0.$$

Proof 2.4

$$\lambda 0 = \lambda(0 + 0) = \lambda 0 + \lambda 0$$

Add $-\lambda 0$ on both sides

$$0 = \lambda 0.$$

$\blacksquare$

Theorem 2.5
Let V be an $\mathcal{F}$-VS and $-1 \in \mathcal{F}$ is the additive inverse of the multiplicative neutral in $\mathcal{F}$. Then

$$(-1)v = -v \quad \forall v \in V$$

Proof 2.5

$$1 \cdot v + (-1)v = (1-1) \cdot v = 0v = 0$$

by Theorem 2.3.

■

For any VS V the subset $\{0\}$ is also a VS. We generalise this notion.

Definition 2.2 (Subspaces)
Let V be an $\mathcal{F}$-VS then a subset $U \subseteq V$ is called a subspace if U is also an $\mathcal{F}$-VS with respect to the same operations.

Theorem 2.6 (Proposition)
A subset $U \subseteq V$ of an $\mathcal{F}$-VS V is a subspace iff (=if and only if)

- $0 \in U$
- $\forall u, w \in U \quad u + w \in U$
- $\forall \lambda \in \mathcal{F} \quad \forall u \in U \quad \lambda u \in U$

Proof 2.6
$\Rightarrow$ If U is a VS then all these conditions hold.

$\Leftarrow$ Condition 1 implies neutral additive of VS.

By condition 3 we know that $(-1)u \in U$, $(-1)u = -u$ and thereby implies the additive inverse of VS.

■

Example 2.2
1) For any VS V, $\{0\}$ and V itself are subspaces.

2) The set of all polynomials with coefficients in some field $\mathcal{F}$ is a VS, called $\mathcal{F}[x]$. For every $0 \leq d \in \mathbb{N}_0$ the set of polynomials of degree at most d is a subspace.

3) We have seen that $\mathbb{R}^{[0,1]}$ is a $\mathbb{R}$-VS. The sets of continuous or differentiable functions form subspaces.

4) We can classify all subspaces of $\mathbb{R}^3$ in a hierarchy: $\mathbb{R}^3 >$ planes containing the origin $>$ lines going through the origin $> \{0\}$.

Definition 2.3
Let $U_1, U_2, \dots, U_m$ be subspaces of a VS V. Then we define their sum.

$$U_1 + \dots + U_m := \{u_1 + \dots + u_m : u_i \in U_i\}$$

Theorem 2.7 (Proposition)
Finite sums of subspaces are subspaces again.

Proof 2.7
We only need to show this for two subspaces.

Let $U_1, U_2 \subseteq V$ be subspaces. Then since $0 \in U_1$ and $0 \in U_2 \Rightarrow 0 + 0 = 0 \in U_1 + U_2$.
Let $u_1 + u_2, u_1' + u_2' \in U_1 + U_2$ then $u_1 + u_2 + u_1' + u_2' = (u_1 + u_1') + (u_2 + u_2') \in U_1 + U_2$
Let $\lambda \in \mathcal{F}$ then

$$\lambda(u_1 + u_2) = \lambda u_1 + \lambda u_2 \in U_1 + U_2.$$

$\blacksquare$

Theorem 2.8 (Proposition)
Let $U_1, U_2 \subseteq V$ be subspaces, then $U_1 + U_2$ is the smallest subspace of V containing both.

Proof 2.8
We see that $U_1 \subseteq U_1 + U_2$ because $\forall u_1 \in U_1 \quad u_1 + 0 = u_1 \in U_1 + U_2$, the same applies to U_2.

Assume there exists $W \subseteq U_1 + U_2$ that contains U_1 and U_2. Then there must exist an element $u_1 + u_2 \notin W$. But $u_1 \in W$ and $u_2 \in W \rightarrow W$ is not a subspace.

EX: Functions and reals can be split into subspaces of even and odd reals.

EX: L_1, L_2 lies in $\mathbb{R}^n$ $L_1 + L_2 = \begin{cases} P \text{ plane} \\ L_1 \text{ if } L_1 = L_2 \end{cases}$

EX: P is a plane in $\mathbb{R}^3$ and L is a line in $\mathbb{R}^3$:

$$P + L = \begin{cases} \mathbb{R}^3 \text{ if } L \subsetneq P \\ P \text{ if } L \subseteq P \end{cases}$$

Definition 2.4 (Direct Sum)
Let $U_1, \dots, U_m \subseteq V$ be subspaces. Then their sum is called a direct sum if $U_1 + \dots + U_m$ has a unique representation as a sum $u_1 + \dots + u_m$. We then write $U_1 \oplus \dots \oplus U_m$ for this sum.

Theorem 2.9 (Proposition)
The sum $U_1 + \dots + U_m$ is direct iff there is a unique way to write 0 as a sum $u_1 + \dots + u_m$.

Proof 2.9
$\Rightarrow$ check

$\Leftarrow$

If the sum is not direct then there exists an element that has two different representations

$$u_1 + \dots + u_m = u'_1 + \dots + u'_m$$

where not all $u_i = u'_i$. Then

$$(u_1 - u'_1) + (u_2 - u'_2) + \dots + (u_m - u'_m) = 0$$

and at least one different $u_i - u'_i \neq 0$.

∎

Theorem 2.10 (Lemma)
$U + W$ is direct iff

$$U \cap W = \{0\}$$

Proof 2.10
$\Rightarrow$: Let $v \in U \cap W$ and $v \neq 0$ then

$$0 + 0 = 0 = v + (-v)$$

and hence the sum is not direct.

$\Leftarrow$: $0 = u + w \Rightarrow -u \in U - w \in W \Rightarrow u = w = 0$

3 Bases and Dimension

A list is an n-tuple.

Definition 3.1 (2.3 and 2.5)
Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be a list of vectors in an $\mathcal{F}$-VS. Then for any $\lambda_i \in \mathcal{F}$ we call $\lambda_1\mathbf{v}_1 + \cdots + \lambda_m\mathbf{v}_m$ a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m$. (Note that λ_i can be zero)

The set of all linear combinations is called the span of $\mathbf{v}_1, \ldots, \mathbf{v}_m$ and denoted $\text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$. For consistency we let $\text{span}() = \{0\}$.

Theorem 3.1 (Proposition 2.7)
Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is a list of vectors. Then $\text{span}(\mathbf{v}_i)$ is a subspace and it is the smallest subspace containing all $\mathbf{v}_i$.

Proof 3.1
We show that span is a subspace.

1. $0 \in \text{span}(\mathbf{v}_i)$, just let $\lambda_i = 0 \ \forall\, i$
2. $\sum_{i=1}^m \lambda_i\mathbf{v}_i + \sum \mu_i\mathbf{v}_i = \sum(\lambda_i + \mu_i)\mathbf{v}_i \in \text{span}(\mathbf{v}_i)$
3. $\sum \lambda_i\mathbf{v}_i = \sum(\mu\lambda_i)\mathbf{v}_i \in \text{span}(\mathbf{v}_i)$

$$\mathbf{v}_i = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 1\mathbf{v}_i + \cdots + 0\mathbf{v}_m \in \text{span}(\mathbf{v}_i)$$

Assume $W \subseteq \text{span}(\mathbf{v}_i)$ such that $\mathbf{v}_i \in W \ \forall\, i$. Then $\exists x \in \text{span}(\mathbf{v}_i) \backslash W \quad x = \sum \lambda_i\mathbf{v}_i \in W$ which is a contradiction.

■

Definition 3.2 (2.17)
We say a list $\mathbf{v}_i$ of vectors is <u>linearly independent</u> if

$$0 = \lambda_1\mathbf{v}_1 + \cdots + \lambda_m\mathbf{v}_m \Rightarrow \forall\, \lambda_i = 0.$$

Theorem 3.2 (Lemma 20)
A list $\mathbf{v}_i$ is linearly independent iff every vector in $\text{span}(\mathbf{v}_i)$ has a unique representation as a linear combination.

Proof 3.2
$\Rightarrow$(direct proof) Assume that $\sum \lambda_i \mathbf{v}_i = \sum \mu_i \mathbf{v}_i$ then

$$\sum(\lambda_i - \mu_i)\mathbf{v}_i = 0 \Rightarrow \lambda_i - \mu_i = 0 \Rightarrow \lambda_i = \mu_i$$

because $\mathbf{v}_i$ is linear independent.

■

Remark:

1. If a list $\mathbf{v}_i$ is linearly dependant then there exist λ_i not all zero, such that $\sum \lambda_i \mathbf{v}_i = 0$
2. A single $\mathbf{v}$ is linearly dependant iff $\mathbf{v} = 0$. Because then $1\mathbf{v} = 1 \cdot 0 = 0$, note that $1 \in \mathcal{F}, v \in V, 0 \in V$.

Definition 3.3 (2.27)
Let V be an $\mathcal{F}$-VS. Then

1. A list $\mathbf{v}_i$ such that $V = \text{span}(\mathbf{v}_i)$ is called a generating set (spanning set). If the list is finite (always assumed here) then we say V is finitely generated.
2. A list $\mathbf{v}_i$ is called a basis for V if it is a linearly independent generating set.

Example 3.1

1. $$\lambda_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \lambda_1 = 0 = \lambda_2$$
2. Let $V = \mathcal{F}^n$ and denote by e_i the vector with a one in the i-th coordinate and zero elsewhere.
3. Let $V = \mathbb{R}[x]^{\leq m}(= \mathcal{P}_m(\mathbb{R}))$ then $1, x, x^2, \dots, x^m$ are a basis with $m+1$ elements.

The $e_1, \dots e_n$ are the so-called standard basis vectors.

Theorem 3.3 (Lemma)
Let $\mathbf{v}_1, \dots, \mathbf{v}_m$ be linearly dependent. Then $\exists j$ such that $\mathbf{v}_j \in \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_{j-1})$ and $\mathbf{v}_1, \dots, \mathbf{v}_m$ without $\mathbf{v}_j$ spans the same space.

Proof 3.3
Since $\mathbf{v}_i$ is linearly dependent $\exists \lambda_i \in \mathcal{F}$, not all zero such that $\sum_{i=1}^{m} \lambda_i \mathbf{v}_i = 0$. Let j be the max index such that $\lambda_j \neq 0$. Then

$$\sum_{i=1}^{j-1} -\frac{\lambda_i}{\lambda_j}\mathbf{v}_i = \mathbf{v}_j \Rightarrow \mathbf{v}_j \in \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_{j-1}). \tag{1}$$

Let $\sum_{i=1}^{m} \mu_i \mathbf{v}_i \in \operatorname{span}(\mathbf{v}_i)$. Substitute equation (1) for $\mathbf{v}_j$

$$\begin{aligned}\mu_1\mathbf{v}_1 + \cdots + \mu_j\mathbf{v}_j + \cdots + \mu_m\mathbf{v}_m &= \mu_1\mathbf{v}_1 + \cdots + \mu_j\left(\sum_{i=1}^{j-1} -\frac{\lambda_i}{\lambda_j}\mathbf{v}_i\right) + \cdots + \mu_m\mathbf{v}_m \\ &= \left(\mu_1 - \frac{\mu_j\lambda_1}{\lambda_j}\right)\mathbf{v}_1 + \left(\mu_2 - \frac{\mu_j\lambda_2}{\lambda_j}\right)\mathbf{v}_2 + \cdots + \left(\mu_{j-1} - \frac{\mu_j\lambda_{j-1}}{\lambda_j}\right)\mathbf{v}_{j-1} + \mu_{j+}\end{aligned}$$

Theorem 3.4 (Steinitz)
Let V be a finitely generated VS. Then the length of any linear independent list is smaller or equal to the length of any generating list.

Proof 3.4
Let $\mathbf{u}_1, \dots, \mathbf{u}_m$ be linearly independent and $\mathbf{w}_1, \dots, \mathbf{w}_n$ a generating set.

$$\operatorname{span}(\mathbf{w}_j) = V, \mathbf{u}_1 \in V.$$

Then $(\mathbf{u}_1\mathbf{w}_1, \dots, \mathbf{u}_1\mathbf{w}_m)$ is linearly dependent. Then for $\sum \lambda_j \mathbf{w}_j = \mathbf{u}_1 \quad wlog \quad \lambda_1 \neq 0 \Rightarrow \frac{1}{\lambda_1}\mathbf{u}_1 - \frac{\lambda_2}{\lambda_1}\mathbf{w}_2 - \frac{\lambda_m}{\lambda_1} = \mathbf{w}_1$ (without loss of generality) point being

$$\operatorname{span}(\mathbf{u}_1, \mathbf{w}_2, \dots, \mathbf{w}_n) = \operatorname{span}(\mathbf{w}_1, \dots, \mathbf{w}_n) = V.$$

The new list $S_1 = (\mathbf{u}_1, \mathbf{w}_2, \dots, \mathbf{w}_n)$ also spans V. Then $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{w}_2, \dots, \mathbf{w}_n)$ and

$$\mathbf{u}_2 = \lambda_1\mathbf{u}_1 + \lambda_2\mathbf{w}_2 + \cdots + \lambda_n\mathbf{w}_n$$

assume $\lambda_2 \neq 0$ and thus an element $\mathbf{w}_2$ can be pulled out of the set without loss:

$$\Rightarrow S_2 = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{w}_3, \dots, \mathbf{w}_n)$$

also spans V and we can keep going.

Remark: This shows that no list that is bigger than a generating set can be linearly independent. Also any list that is shorter than a linearly independent list can not generate the whole space.

Theorem 3.5 (Basis)
A list of vectors is a basis for V iff every $\mathbf{v} \in V$ can be uniquely be written as a linear combination.

Proof 3.5
Lemma 20. If you can write every element uniquely then you can write zero uniquely.

Theorem 3.6
Let $\text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n) = V$. Then there is a subset of $\mathbf{v}_i$ that is a basis.

Proof 3.6
We construct the basis in n-steps.

We add a vector $\mathbf{v}_i$ to our basis if $\mathbf{v}_i \notin \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1})$. Let $\mathbf{w}_1, \ldots, \mathbf{w}_m$ be the basis acquired this way. Assume

$$\sum \lambda_i \mathbf{w}_i = 0.$$

Let j be max such that $\lambda_j \neq 0$ then $\sum_{i=1}^{j-1} \lambda_i \mathbf{w}_i = \lambda_j \mathbf{w}_j$, contradiction.

Therefore $\mathbf{w}_i$ is linearly independent and it still spans V.

Theorem 3.7 (Corollary)
Every finitely generated VS has a basis.

Theorem 3.8 (Corollary)
Every linearly independent set can be extended to a basis.

Proof 3.7
Let $(\mathbf{u}_1, \ldots, \mathbf{u}_m)$ be linearly independent and let $(\mathbf{w}_1, \ldots, \mathbf{w}_n)$ be a generated set. Then $(\mathbf{u}_1, \ldots, \mathbf{u}_m, \mathbf{w}_1, \ldots, \mathbf{w}_n)$ is a geenerating set. Use Theorem 3.6 to acquire a basis.

Theorem 3.9 (2.35)
Every basis of a finitely generated VS has the same length.

Proof 3.8
Let B_1 and B_2 be two bases. Since B_1 is linearly independent and B_2 generates V.

$$\begin{aligned} |B_1| &\leq |B_2| \\ |B_2| &\leq |B_1| \\ \Rightarrow |B_1| &= |B_2| \end{aligned}$$

Definition 3.4 (Dimension)
Let V be an $\mathcal{F}$-VS. Then we define dimension as

$$\dim_{\mathcal{F}}(V) = \begin{cases} \text{length of the basis if } V \text{ is finitely generated} \\ \infty \quad \text{otherwise} \end{cases}$$

Theorem 3.10 (Corollary)
Let $U \subseteq V$ be a subspace. Then $\dim(U) \leq \dim(V)$.

Proof 3.9
A basis of U is a linear set in V. Hence it is shorter or equal in length to any generating set of V, especially a basis of V.

Theorem 3.11 (Corollary 2.39)
A linearly independent list of size $\dim(V)$ is already a basis.

Proof 3.10
We can extend the list to a basis. But it is already of length $\dim(V)$ hence nothing is added.

Theorem 3.12 (Corollary 2.42)
Let $\dim(V) = n$ then every generating set of length n is already a basis.

Two sets A, B with size $|A|, |B|$. The union has size: $|A \cup B| = |A| + |B| - |A \cap B|$

Theorem 3.13
Let A, B be subspaces of a finite dimensional space V. Then $\dim(A+B) = \dim(A) + \dim(B) - \dim(A \cap B)$.

Proof 3.11
Let $c_1, \ldots, c_l$ be a basis for $A \cap B$. We extend to a basis $c_1, \ldots, c_l, a_1, \ldots, a_m$ of A and to a basis $c_1, \ldots, c_l, b_1, \ldots, b_n$ of B.

We want to show that $c_i a_j b_k$ is a basis for $A + B$. This is a generating set, now we need to check that it is linearly independent.

Now let

$$\begin{aligned}
0 &= \sum \alpha_i a_i + \sum \beta_j b_j + \sum \mu_k c_k \\
-\sum \alpha_i a_i &= \sum \beta_j b_j + \sum \mu_k c_k \in A \cap B \\
-\sum \alpha_i a_i &= \sum \delta_k c_k \\
0 &= \sum \alpha_i a_i + \sum \delta_k c_k \\
\Rightarrow \alpha_i &= 0 \quad \delta_k = 0 \\
\Rightarrow 0 &= \sum (\beta_j b_j + \sum \gamma_k c_k) \\
\Rightarrow \beta_j &= 0 \quad \gamma_k = 0
\end{aligned}$$

4 Maps

Definition 4.1 (3.2/3.8)
Let $V,\ W$ be two $\mathcal{F}$-VS. A map $T : V \to W$ is called linear if

$$T(\mathbf{v} + \mathbf{v}') = T(\mathbf{v}) + T(\mathbf{v}') \quad \forall \mathbf{v}, \mathbf{v}' \in V$$
$$T(\lambda \mathbf{v}) = \lambda T(\mathbf{v}) \quad \forall \lambda \in \mathcal{F}\ \mathbf{v} \in V.$$

The set of all linear maps from V into W is denoted $\mathrm{Hom}_{\mathcal{F}}(V,\ W)$ meaning homomorphism (in the book: $\mathcal{L}(V,\ W)$). If $V = W$ we also write $\mathrm{End}_{\mathcal{F}}(\mathbf{v}) = \mathrm{Hom}(V,\ V)$.

Example 4.1

$$0 \in \mathrm{Hom}(V,\ W) \quad 0 \in \mathcal{F}\mathbf{v} = 0 \in W$$

Another example is the identity (id):

$$id \in \mathrm{End}(V) \quad id\ \mathbf{v} = \mathbf{v}.$$

Differentiating a polynomial is a linear map. The same applies to integration.
Multiplication by x^2 is a linear map in $\mathrm{Hom}(\mathbb{R}[x],\ \mathbb{R}[x])$.
Most commonly though:

$$T(x,\ y,\ z) = (2x - y, 3y + z)$$

Theorem 4.1
Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis for V and $\mathbf{w}_1, \ldots, \mathbf{w}_n$ some vectors in W. Then there exists a unique linear map T such that $T(\mathbf{v}_i) = \mathbf{w}_i$.

Proof 4.1
We show uniqueness and existence by explicitly calculating images of T. Let $\mathbf{v} \in V$ then exists unique $\lambda \in \mathcal{F}$ such that

$$\mathbf{v} = \sum \lambda_i \mathbf{v}_i.$$

Now

$$T(\mathbf{v}) = T\Big(\sum \lambda_i \mathbf{v}_i\Big) = \sum T(\lambda_i \mathbf{v}_i) = \sum \lambda_i T(\mathbf{v}_i)$$

Theorem 4.2 (Proposition 3.7)
$\mathrm{Hom}(V, W)$ is itself a $\mathcal{F}$-VS with usual addition and scalar multiplication

$$\forall S, T \in \mathrm{Hom}(V, W)$$
$$(S + T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v})$$
$$\forall \lambda \in \mathcal{F} \quad (\lambda \cdot T)(\mathbf{v}) = \lambda(T(\mathbf{v}))$$

Proof 4.2

1. $0 \in \mathrm{Hom}(V, W)$
2. $S, T \in \mathrm{Hom}(V, W) \Rightarrow S + T \in \mathrm{Hom}(V, W)$
3. $T \in \mathrm{Hom} \Rightarrow \lambda T \in \mathrm{Hom}$

Definition 4.2 (3.8)
Let $T \in \mathrm{Hom}(U, V)$ and $S \in \mathrm{Hom}(V, W)$. Then we define $ST \in \mathrm{Hom}(U, W)$ $(U \underbrace{\to}_{T} V \underbrace{\to}_{S} W)$. As $ST(\mathbf{u}) = S(T(\mathbf{u})) = S \circ T(\mathbf{u})$. We see that for three suitable maps

$$\begin{aligned}(ST)U &= S(TU)\\ id\,T &= T\,id = T\\ (S+T)U &= SU + TU\\ S(T+U) &= ST + SU\end{aligned}$$

Note! Composition of linear maps is not commutative: $T, D \in \mathrm{End}(\mathbb{R}[x])$

$$T(p) = x^2 p \quad D(p) = p' \quad TD(p) = x^2 p' \quad DT(p) = x^2 p' + 2xp.$$

Definition 4.3 (3.12 / 3.17)
Let $T \in \mathrm{Hom}(V, W)$. We define the image(range) of T as $\mathrm{im}(T) = \{T\mathbf{v} : \mathbf{v} \in V\} \subseteq W$ and its kernal(nullspace) as $\ker(T) = \{\mathbf{v} \in V : T\mathbf{v} = 0\} \subseteq V$.

Theorem 4.3
Image and kernel are subspaces.

Proof 4.3
We start with the image:

1. $0 \in \mathrm{im}(T)$ $\quad T(0) = T(0+0) = T(0) + T(0) \Rightarrow T(0) = 0$ and bonus $T(0) = T(0 \cdot \mathbf{v}) = 0 \cdot T(\mathbf{v}) = 0$
2. $\mathbf{w}, \mathbf{w}' \in \mathrm{im}(T) \Rightarrow T(\mathbf{v}) = \mathbf{w},\ T(\mathbf{v}') = \mathbf{w}'$

$$\mathbf{w} + \mathbf{w}' = T(\mathbf{v}) + T(\mathbf{v}') = T(\mathbf{v} + \mathbf{v}')$$

3. $\mathbf{w} \in \mathrm{im}(T) \quad \lambda \in \mathcal{F}$
$$T(\lambda\mathbf{v}) = \lambda T(\mathbf{v}) = \lambda\mathbf{w}$$

Now the kernal:

1. By (*) $0 \in \ker$
2. $\mathbf{v}, \mathbf{v}' \in \ker \quad T(\mathbf{v}+\mathbf{v}') = T(\mathbf{v}) + T(\mathbf{v}') = 0 + 0 = 0$
3. $\mathbf{v} \in \ker \quad T(\lambda\mathbf{v}) = \lambda T(\mathbf{v}) = \lambda 0 = 0$

Definition 4.4
Let $f : A \to B$. We say f is injective if $f(a) = f(b) \Rightarrow a = b$ and surjective if $\forall b \; \exists a$ such that $f(a) = b$

Theorem 4.4 (Proposition)
$T \in \mathrm{Hom}(V, U)$ is injective iff $\ker(T) = \{0\}$ and surjective if $\mathrm{im}(T) = W$.

Proof 4.4
Injective: $\Rightarrow$ proof. Assume T is injective. Let $\mathbf{v} \in \ker(T)$ then

$$T(\mathbf{v}) = 0 = T(0) \Rightarrow \mathbf{v} = 0$$

by injectivity.

$\Leftarrow$ proof. Assume $\ker(T) = \{0\}$ and

$$T(a) = T(b) \Rightarrow T(a) - T(b) = 0 \Rightarrow T(a-b) = 0 \Rightarrow a - b = 0 \Rightarrow a = b$$

Surjective is automatically done as it literally means it is the whole thing.

Theorem 4.5
Let V be a finite dimensional VS and $T \in \mathrm{Hom}(V, W)$. Then $\mathrm{im}(T)$ is also finite dimensional and
$$\dim(V) = \dim(\ker(T)) + \dim(\mathrm{im}(T))$$

Proof 4.5
Let $\mathbf{u}_1, \ldots, \mathbf{u}_m$ be a basis of $\ker(T)$ which is a subspace of V and we can extend this to a basis of V by adding $\mathbf{v}_1, \ldots, \mathbf{v}_n$. The $\dim(V) = m + n \quad \dim(\ker(T)) = m$. Need to show that $\dim(\mathrm{im}(T)) = n$.

Let $\mathbf{v} \in V$ then

$$\mathbf{v} = \sum \lambda_i \mathbf{u}_i + \lambda \mu_j \mathbf{v}_j$$

and

$$T(\mathbf{v}) = T\Big(\sum \lambda_i \mathbf{u}_i + \sum \mu_j \mathbf{v}_j\Big) = T\Big(\sum \mu_j \mathbf{v}_j\Big) = \sum \mu_j T(\mathbf{v}_j)$$

which implies ($\Rightarrow$) the set of vectors $T(\mathbf{v}_j)$ generates/spans the image of T. Now we need to show that they are linear independent.

Assume

$$\sum \alpha_j T(\mathbf{v}_j) = 0$$

if they are linear independent then all $\alpha_j = 0$:

$$\begin{aligned} \sum T(\alpha \mathbf{v}_j) &= T\left(\underbrace{\sum \alpha_j \mathbf{v}_j}_{\in \ker(T)}\right) \\ \sum \alpha_j \mathbf{v}_j &= \sum \beta_i \mathbf{u}_i \\ \sum \alpha_j \mathbf{v}_j + \sum (-\beta_i) \mathbf{u}_i &= 0 \\ \Rightarrow \alpha_j &= 0 \end{aligned}$$

because $\mathbf{u}_1, \ldots, \mathbf{u}_m, \mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis and hence linear independent.

■

Theorem 4.6 (Corollary)
If $\dim(V) > \dim(W)$ then no $T \in \mathrm{Hom}(V, W)$ is injective.

Proof 4.6

$$\begin{aligned} \dim(V) &= \dim(\ker(T)) + \dim(\mathrm{im}(T)) \\ \Rightarrow \dim(V) - \dim(\mathrm{im}(T)) &= \dim(\ker(T)) \\ \mathrm{im}(T) &\leq W \\ \dim(\mathrm{im}(T)) &\leq \dim(W) < \dim(V) \\ \Rightarrow 1 &\leq \dim(\ker(T)) \\ \Rightarrow \ker(T) &\neq \{0\} \end{aligned}$$

which implies T is not injective.

Theorem 4.7 (Corollary)
If $\dim(V) < \dim(W)$ no $T \in \mathrm{Hom}(V, W)$ is surjective.

Proof 4.7

$$
\begin{aligned}
\dim(V) &= \dim(\ker(T)) + \dim(\operatorname{im}(T)) \\
\dim(W) > \dim(V) &\geq \dim(\operatorname{im}(T)) \\
\operatorname{im}(T) &\subsetneq W
\end{aligned}
$$

Definition 4.5
Let $T \in \operatorname{Hom}(V, W)$ and $\mathbf{v}_1, \ldots, \mathbf{v}_m$ basis of V and $\mathbf{w}_1, \ldots, \mathbf{w}_n$ basis of W. Then the matrix of T with respect to these bases is given by the entries A_{jk} defined by

$$T\mathbf{v}_k = \sum A_{jk}\mathbf{w}_j.$$

$A = (A_{jk}) = \mathcal{M}(T)$.

Example 4.2
$\mathbb{R}[x]^{<4}$ with the differentiation mapping, $D \in \operatorname{Hom}(\mathbb{R}[x]^{<4}, \mathbb{R}[x]^{<3})$, the basis of $\mathbb{R}[x]^{<4}$ is $1, x, x^2, x^3$ and for $\mathbb{R}[x]^{<3}$ it is $1, x, x^2$. We get the entries of the matrix by

$$
\begin{aligned}
D(1) &= 0 \\
D(x) &= 1 = 1 \cdot 1 + 0x + 0x^2 \\
D(x^2) &= 2x = 0 \cdot 1 + 2 \cdot x + 0x^2 \\
D(x^3) &= 3x^2 = 0 \cdot 1 + 0 \cdot x + 3 \cdot x^2
\end{aligned}
$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

Another example changes the first basis to $1 + x, x + x^2, x^2 + x^3, x^3$, and the entries of the matrix are

$$
\begin{aligned}
D(1 + x) &= 1 \\
D(x + x^2) &= 1 + 2x \\
D(x^2 + x^3) &= 2x + 3x^2 \\
D(x^3) &= 3x^2
\end{aligned}
$$

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 3 & 3 \end{pmatrix}$$

With two linear maps T, S such that $S \circ T$ makes sense, then $\mathbf{u} \to \mathbf{v} \to \mathbf{w}$ and the matrix of the combined bases is $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$.

5 Invertibility and Isomorphisms

Definition 5.1
Let $T \in \mathrm{Hom}(V, W)$, then we say T is invertible if there exists an $S \in \mathrm{Hom}(W, V)$ such that $ST = \mathrm{id}_V$ and $TS = \mathrm{id}_W$. We call S the inverse of T.

Theorem 5.1 (Proposition)
The inverse of an invertible map is unique.

Proof 5.1
Suppose $T \in \mathrm{Hom}(V, W)$ is invertible and S and S' are both inverses. then

$$S = S\mathrm{id}_W = S(TS') = (ST)S' = \mathrm{id}_V S' = S'.$$

We decide $T^{-1} = S$ from now on.

Theorem 5.2
A linear map T is invertible iff it is injective and surjective.

Proof 5.2
Direct proof "$\Rightarrow$".

We want to show it is injective: Assume $T\mathbf{v} = T\mathbf{v}'$. Since it is invertible it has an inverse $T^{-1}T\mathbf{v} = T^{-1}T\mathbf{v}'$ and thus $\mathbf{v} = \mathbf{v}'$.

To show it is surjective we have $\mathbf{w} \in W$ and $T^{-1}\mathbf{w}$ is a method to get it back into V. We do this by $TT^{-1}\mathbf{w} = \mathbf{w}$.

Now indirect proof "$\Leftarrow$".

We construct inverse $S : W \to V$ by defining $S\mathbf{w} = \mathbf{v}$ where $T\mathbf{v} = \mathbf{w}$. This $\mathbf{v}$ exosts because T is surjective and $\mathbf{v}$ is unique because T is injective. Obviously $TS = \mathrm{id}_W$. Consider

$$T(ST) = (TS)T = T$$

and now we want to show that ST is the identity of V:

$$T(ST)\mathbf{v} = (TS)T\mathbf{v} = T\mathbf{v} \Rightarrow ST = \mathrm{id}_V$$

because T is injective.

Need to check that it is closed under addition and multiplication for it to be linear:

$$TS(x + y) = x + y = TSx + TSy = T(Sx + Sy)$$

By injectivity of T we have that $S(x + y) = Sx + Sy$. Now for multiplication:

$$TS(\lambda x) = \lambda x = \lambda TSx = T(\lambda Sx)$$

and we are good. Thus it is linear.

Definition 5.2
We say two VS are isomorphic if there exists an invertible linear map $T : V \to W$. We write $V \cong W$ and call T a isomorphism.

Theorem 5.3
Any two finite dimensional $\mathcal{F}$-VSs are isomprhic iff they have the same dimension.

Proof 5.3
We have seen that maps between VSs of different dimensions are either not injective or not surjective. Therefore if they are $V \cong W \Rightarrow \dim(V) = \dim(W)$. Now assume $\dim(V) = \dim(W) = n$ and let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and $\mathbf{w}_1, \ldots, \mathbf{w}_n$ be bases for V and W respectively. Define $T : V \to W$ by $T(\mathbf{v}) = T(\sum \lambda_i \mathbf{v}_i) = \sum \lambda_i T(\mathbf{v}_i) = \sum \lambda_i \mathbf{w}_i$, then T is an isomorphism. Let $T(\mathbf{v}) = 0$, thus $\mathbf{v} \in \ker(T)$:

$$T(\mathbf{v}) = T(\sum \lambda_i \mathbf{v}_i) = \sum \lambda_i \mathbf{w}_i = 0 \Rightarrow \lambda_i = 0 \;\forall i$$

because the $\mathbf{w}_i$ are linearly independent as they are a basis. This implies $\mathbf{v} = 0 \Rightarrow \ker(T) = \{0\}$.

Say $\mathbf{w} = \sum \mu_i \mathbf{w}_i \in W$. Then

$$T(\sum \mu_i \mathbf{v}_i) = \mathbf{w}$$

$\Rightarrow T$ is surjective.

After fixing bases for V and W we have a map $\mathcal{M} : \mathrm{Hom}_{\mathcal{F}}(V, W) \to \mathcal{F}^{m \times n}$.
<u>One checks that $\mathcal{M}$ is indeed linear.</u>

Theorem 5.4
The map $\mathcal{M}$ is an isomorphism.

Proof 5.4
Need to show it is injective and surjective. We start with showing it is injective:

$$\mathcal{M}(T) = 0$$

each column represents a basis vector of V, and if these are all 0 then $T(\mathbf{v}_i) = 0 \;\forall i$ where $\mathbf{v}_i$ is a basis. Thus $T\mathbf{v} = 0 \;\forall \mathbf{v} \in V$ and thus $T = 0$ is the linear map that maps all vectors to the zerovector. Injectivity is then shown.

Now to show surjectivity we have $A \in \mathcal{F}^{m \times n}$ then we define T such that

$$T\mathbf{v}_k = \sum_{j=1}^{m} A_{jk} \mathbf{w}_j$$

and it follows that $\mathcal{M}(T) = A$.

Theorem 5.5 (Corollary)

$$\dim(\mathrm{Hom}_{\mathcal{F}}(V, W)) = \dim(V) + \dim(W)$$

Proof 5.5
$\mathcal{F}^{m\times n}$ with $E_{i,j}$ which has zeros everywhere except row i and column j where there is a 1. These are a basis.

Definition 5.3

$$\mathrm{End}(V) = \mathrm{Hom}(V, V)$$

is the set of linear maps from V into V, called the endomorphisms.

Theorem 5.6
Let V be a finite dimensional VS and $T \in \mathrm{End}(V)$. Then the following statements are equivalent:

1. T is injective.
2. T is surjective.
3. T is invertible.

Proof 5.6
$\Leftrightarrow$ proof:
The kernal of T is just zero, this implies that $\dim(V) = \dim(\mathrm{im}(T)) + \dim(\ker(T))$ but the dimension of the kernal is zero. Which can only happen if $V = \mathrm{im}(T)$.

No need to check for injectivity and surjectivity if it maps only to zero.

Definition 5.4
Let $V_1, \ldots, V_m$ be $\mathcal{F}$-VS then we define a new VS as

$$V_1 \cdot \cdots \cdot V_m = \{(\mathbf{v}_1, \ldots, \mathbf{v}_m) : \mathbf{v}_i \in V_i\}$$

with obvious addition and scalar multiplication.

Theorem 5.7

$$\dim(V_1 \cdot \cdots \cdot V_m) = \dim(V_1) + \cdots + \dim(V_m).$$

Proof 5.7
We provide a basis. Choose a basis for each component V_i and for each basis vector. Consider the element in $V_1 \cdot \cdots \cdot V_m$ that has $\mathbf{v}_k$ in the ith position and zeros elsewhere. That forms a basis.

Theorem 5.8 (Proposition)
Let $U_1, \ldots, U_m$ are subspaces of V. Define the map $\Gamma : U_1 \cdot \cdots \cdot U_m \to U_1 + \cdots + U_m$. Then $U_1 + \cdots + U_m$ is defined if Γ is injective.

Proof 5.8
Γ is injective: $0 = \Gamma(\mathbf{u}_1, \ldots, \mathbf{u}_m) \Rightarrow \mathbf{v} = (\mathbf{u}_1, \ldots, \mathbf{u}_m) = (0, \ldots, 0)$ which is the same as saying $0 = \sum \mathbf{u}_i$. If the sum is direct then the sum implies $\mathbf{u}_i = 0 \; \forall i$.

Theorem 5.9 (Corollary)
Let $U_1, \ldots, U_m$ be subspaces of a finite dimensional VS V. Then $U_1 + \cdots + U_m$ is direct iff

$$\dim(U_1 + \cdots + U_m) = \dim(U_1) + \cdots + \dim(U_m).$$

Proof 5.9
Γ is obviously surjective. By Theorem 5.6 Γ is an isomorphism iff the dimensions match. Hence if the sum is direct Γ is injective by the Theorem 5.8 and have isomorphism, therefore the dimensions match.

Definition 5.5
Let $U \subset V$ and $\mathbf{v} \in V$ then the set $\mathbf{v} + U := \{\mathbf{v} + \mathbf{u} \; \mathbf{u} \in U\}$ is called an affine subset.

Theorem 5.10 (Lemma)
Let $U \subset V$ then the following are equivalent:

1. $\mathbf{v} - \mathbf{w} \in U$
2. $\mathbf{v} + U = \mathbf{w} + U$
3. $\mathbf{v} + U \cup \mathbf{w} + U \neq \emptyset$

Definition 5.6
Let $U \subset V$ and consider the set of all affine subsets:

$$V/U := \{\mathbf{v} + U : \mathbf{v} \in V\}.$$

This is called the quotient space.

Theorem 5.11
V/U is a VS with additivity given by $(\mathbf{v}+U)+(\mathbf{w}+U) = (\mathbf{v}+\mathbf{w})+U$ and multiplication given by $\lambda(\mathbf{v} + U) = \lambda\mathbf{v} + U$.

Proof 5.10
We need to check the this:
Let $\mathbf{v} + U = \mathbf{v}' + U$ and $\mathbf{w} + U = \mathbf{w}' + U$ then we want

$$(\mathbf{v} + \mathbf{w}) + U = (\mathbf{v}' + \mathbf{w}') + U$$
$$(\mathbf{v} + \mathbf{w}) - (\mathbf{v}' + \mathbf{w}') = (\mathbf{v} - \mathbf{v}') + (\mathbf{w} - \mathbf{w}') \in U$$

since $\mathbf{v} - \mathbf{v}' \in U$ and the same applies to the $\mathbf{w}$'s.

Theorem 5.12

$$\dim(V/U) = \dim(V) - \dim(U).$$

Proof 5.11
Consider the projection map $\pi : V \to V/U$ by $\mathbf{v} \to \mathbf{v} + U$. The kernal $\ker(\pi) = \{\mathbf{v} \in V : \pi(\mathbf{v}) = 0 + U\}$ which is everything in U that is mapped to zero, which is just U. The image is $\mathrm{im}(\pi) = V/U$, and the dimension follow

$$\dim(V) = \dim(V/U) + \dim(U) \Rightarrow \dim(V/U) = \dim(V) - \dim(U).$$

6 Eigenvalues and Eigenspaces

Definition 6.1 (5.5/5.7/5.34)
Let V be an $\mathcal{F}$-VS and $T \in \text{End}(V)$. An element $\lambda \in \mathcal{F}$ is called an eigenvalue if there is a vector $\mathbf{v} \in V$, $\mathbf{v} \neq 0$, such that $T\mathbf{v} = \lambda\mathbf{v}$. Then $\mathbf{v}$ is called an eigenvector for λ, and the subspace $E(\lambda, T) = \ker(T - \lambda I)$ is called the eigenspace.

Example 6.1

- Let $T \in \text{End}(V)$ with $\ker(T) \neq \{0\}$ then 0 is an eigenvalue.
- Let $A = \begin{pmatrix} 1 & -1 \\ -2 & 0 \end{pmatrix}$, $A : \mathbb{R}^2 \to \mathbb{R}^2$. We have

$$A \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \end{pmatrix} = -1\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \Rightarrow \text{-1 is an eigenvalue.}$$
$$A \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \end{pmatrix} = 2\begin{pmatrix} -1 \\ 1 \end{pmatrix}, \Rightarrow \text{2 is an eigenvalue.}$$

Theorem 6.1 (Corollary 5.6)
Let $T \in \text{End}(V)$, $\lambda \in \mathcal{F}$ then the following are equivalent:

1. λ is an eigenvalue.
2. $\ker(T - \lambda I) \neq \{0\} \Leftrightarrow T - \lambda I$ is not injective.

If V is finite dimensional then the following are also equivalent to the above:

1. $T - \lambda I$ is not surjective.
2. $T - \lambda I$ is not invertible.

Proof 6.1

$$T\mathbf{v} = \lambda\mathbf{v} \Leftrightarrow T\mathbf{v} - \lambda\mathbf{v} = (T - \lambda I)\mathbf{v} = 0.$$

Theorem 6.2 (5.10)
Let $T \in \text{End}(V)$ and suppose $\lambda_1, \ldots, \lambda_n$ are distinct eigenvalues, with corresponding eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Then $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent.

Proof 6.2
Assume $\exists \mu_i$ such that $\sum_{i=1}^{n} \mu_i \mathbf{v}_i = 0$ and not all $\mu_i = 0$. Wlog let us assume that $\mu_n \neq 0$, then
$$0 = T(\sum \mu_i \mathbf{v}_i) = \sum \mu_i T\mathbf{v}_i = \sum \mu_i \lambda_i \mathbf{v}_i.$$
Subtract $\lambda_n \cdot \sum \mu_i \mathbf{v}_i$ for this
$$0 = \sum \mu_i \lambda_i \mathbf{v}_i - \sum \mu_i \lambda_n \mathbf{v}_i = \sum_{i=1}^{n-1} \mu_i(\lambda_i - \lambda_n)\mathbf{v}_i.$$
Repeating this procedure shows that all $\mu_i = 0$. Hence $\mathbf{v}_1, \dots, \mathbf{v}_n$ is linearly independent.

Theorem 6.3 (Corollary 5.13)
Let V be finite dimensional, then any endomorphism has at most $\dim(V)$ distinct eigenvalues.

Remark: $T \in \mathrm{End}(V)$ then $T^2 = T \circ T$ and $T^3 = T \circ T^2$. Also $T^\circ = \mathrm{id}$, if T is invertible then $T^{-m} = (T^{-1})^m$. Now we can build polynomials: Let $p(x) \in \mathcal{F}[x]$, $p(T) = c_0\mathrm{id} + c_1 T + c_2 T^2 + \cdots + c_m T^m$.

Theorem 6.4 (5.21)
Let V be a finite dimensional $\mathbb{C}$-VS then every $T \in \mathrm{End}(V)$ has an eigenvalue.

Proof 6.3
Let $\mathbf{v} \in V$, $\mathbf{v} \neq 0$, then the following list $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \dots, T^n\mathbf{v}$, where $n = \dim(V)$, is linearly dependent (there is $n+1$ vectors). This implies $\exists c_i \in \mathbb{C}$ such that
$$\sum c_i T^i \mathbf{v} = 0$$
not all $c_i = 0$. Consider
$$p(x) = \sum_{i=0}^{n} c_i x^i = c \prod_{i=1}^{n} (x - \lambda_i).$$
Now we take
$$p(T) = c \cdot \prod_{i=1}^{n} (T - \lambda_i I)$$
$$p(T)\mathbf{v} = c \cdot \prod_{i=1}^{n} (T - \lambda_i I)\mathbf{v} = 0.$$
Somewhere along the way a vector is mapped to zero. Hence at least one of $(T - \lambda_i I)$ is not injective and therefore T has an eigenvalue.

Theorem 6.5 (5.26)
Let $T \in \text{End}(V)$ and $\mathbf{v}_1, \dots, \mathbf{v}_n$ is a basis of V. Then the following are equivalent:

1. The matrix $A = \mathcal{M}(T)$ is upper triangular: $A_{ij} = 0,\ i > j$.
2. $T\mathbf{v}_j \in \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_j)$.
3. $\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_j)$ is invariant under T for each j.

Proof 6.4

1. $$T\mathbf{v}_j = \sum_{i=1}^{n} A_{ij}\mathbf{v}_i = \sum_{i=1}^{j} A_{ij}\mathbf{v}_i$$
   as the image of $\mathbf{v}_j$ is a linear combination of the vectors $\mathbf{v}_1, \dots, \mathbf{v}_j$
2. (3) $\Rightarrow$ (2): $\mathbf{v}_j \in \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_j)$
   $$T\mathbf{v}_j \in \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_j)$$
3. (2) $\Rightarrow$ (3): $T\left(\sum_{i=1}^{j} \lambda_i \mathbf{v}_i\right) = \sum_{i=1}^{j} \lambda_i T(\mathbf{v}_i)$ which sits in $\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_i)$.

Theorem 6.6 (5.27)
Let V be a finite dimensional $\mathbb{C}$-VS. For every $T \in \text{End}V$ there exists a basis such that $\mathcal{M}(T)$ is upper triangular with respect to this basis.

With upper triangular matrices it is ensured that the first basis vector is an eigenvector.

Proof 6.5
We use induction on the dimension of V. For $\dim(V) = 1$ this is true.

Now assume that the theorem holds for all $\mathbb{C}$-VS of dimension lower than $\dim(V)$. T has at least one eigenvalue λ by theorem ?? and we consider the subspace $U = \text{im}(T - \lambda I)$, which is smaller than V. We see that U is invariant under T since

$$\mathbf{u} \in U, \quad T\mathbf{u} = T\mathbf{u} - \lambda\mathbf{u} + \lambda\mathbf{u} = \underbrace{(T - \lambda I)\mathbf{u}}_{\in U} + \underbrace{\lambda\mathbf{u}}_{\in U} \in U.$$

Hence we can consider the restriction of T onto U, $T|_U \in \text{End}U$. Since $\dim(U) < \dim(V)$ there exists a basis $\mathbf{u}_1, \dots, \mathbf{u}_n$ of U such that $T|_U$ is upper triangular. Extend the basis to a basis of V $\mathbf{u}_1, \dots, \mathbf{u}_n, \mathbf{v}_1, \dots, \mathbf{v}_m$. We see that

$$T\mathbf{v}_i = \underbrace{T\mathbf{v}_i - \lambda\mathbf{v}_i}_{\in U} + \lambda\mathbf{v}_i) \in \text{span}(\mathbf{u}_1, \dots, \mathbf{u}_m\mathbf{v}_i) \subseteq \text{span}(\mathbf{u}_1, \dots, \mathbf{u}_n, \mathbf{v}_1, \dots, \mathbf{v}_m).$$

And by Proposition ?? $\mathcal{M}(T)$ is upper triangular with respect to this basis.

Theorem 6.7 (5.30)
Suppose $T \in \text{End}V$ and there is a basis such that $\mathcal{M}(T)$ is upper triangular. Then T is invertible iff all diagonal elements are nonzero.

Proof 6.6
"$\Leftarrow$" proof: Let $\mathcal{M}(T) = A = \begin{pmatrix} \lambda_1 & & \\ & dots & \\ & & \lambda_n \end{pmatrix}$ and all $\lambda_i \neq 0$. We show that T is surjective. Since $T\mathbf{v}_1 = \lambda_1 \mathbf{v}_1 \Rightarrow T\frac{1}{\lambda_1}\mathbf{v}_1 = \mathbf{v}_1$ we see that $\mathbf{v}_1 \in \text{im}(T)$. Also

$$T\mathbf{v}_2 = A_{12}\mathbf{v}_1 + \lambda_2 \mathbf{v}_2 \Rightarrow T\left(\mathbf{v}_2 - \frac{A_{12}}{\lambda_1}\mathbf{v}_1\right) = A_{12}\mathbf{v}_1 + \lambda_2\mathbf{v}_2 - A_{12}\mathbf{v}_1 = \lambda_2\mathbf{v}_2.$$

Continuing like this shows that $\mathbf{v}_i \in \text{im}(T)\ \forall i \Rightarrow T$ is surjective which implies that T is invertible.

"$\Rightarrow$" proof: Assume $\lambda_i = 0$ then the subspace $\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_i)$ is mapped by T onto $\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_{i-1})$. Hence T can not be injective as a bigger space is mapped into a smaller space (i space to $i-1$ space). $\exists \mathbf{v} \in \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_i)$ such that $T\mathbf{v} = 0,\ \mathbf{v} \neq 0$.

Theorem 6.8 (5.32)
Let $\mathcal{M}(T)$ be upper triangular then the eigenvalues of T are the diagonal entries.

Proof 6.7
An element $\lambda \in \mathcal{F}$ is an eigenvalue iff $T - \lambda I$ is not invertible. The matrix $\mathcal{M}(T - \lambda I)$ has diagonal elements $(\lambda_i - \lambda)$ where λ_i are the diagonal entries of $\mathcal{M}(T)$. $T - \lambda I$ not invertible iff $\lambda_i - \lambda = 0$ for at least one $i \Leftrightarrow \lambda = \lambda_i$ for some i.

Theorem 6.9 (Proposition)
Let V be a finite dimensional VS and $T \in \text{End}V$. Then $E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ is a direct sum where λ_i are distinct eigenvalues, and

$$\sum \dim(E(\lambda_m, T)) \leq \dim(V).$$

Proof 6.8
Assume there are $\mathbf{u}_i \in E(\lambda_i, T)$ such that not all $\mathbf{u}_i = 0$ and $\sum \mathbf{u}_i = 0$. Every $\mathbf{u}_i$ is an eigenvector to a different eigenvalue(or $\mathbf{u}_i) = 0$) but these are linearly independent. For the sum to be zero, all $\mathbf{u}_i$ must be zero. Hence

$$\dim(E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)) = \sum \dim(E(\lambda_i, T)) \leq \dim(V).$$

Definition 6.2
We say $T \in \text{End}V$ is diagonalizable if there exists a basis such that $\mathcal{M}(T)$ is diagonal.

Theorem 6.10 (Proposition)
Let V be finite dimensional, $T \in \text{End}V$, and $\lambda_1, \dots, \lambda_m$ are the eigenvalues of T. Then the following are equivalent:

1. T is diagonalizable.
2. V has a basis of eigenvectors.
3. $\dim(V) = \sum \dim(E(\lambda_1, T))$.

Proof 6.9

1. Implies (2). Every vector is mapped to a multiple of itself. And (2) implies (1).
2. Implies (3). Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be a basis of eigenvectors. Each of $\mathbf{v}_i \in E(\lambda_j, T)$ for some j. This implies that
$$\sum \dim(E(\lambda_j, T)) \geq n = \dim(V) \Rightarrow \sum \dim(E(\lambda_j, T)) = \dim(V)$$
by Proposition 6.9.
3. Implies (2). Assume $\sum \dim(E) = \dim(V)$. Choose a basis for each $E(\lambda_j, T)$ then the union of these bases is a basis for V $\mathbf{v}_1, \dots, \mathbf{v}_n$. To show linearly independence we assume that they are not: Assume $\sum \mu_i \mathbf{v}_i = 0$ for not all $\mu_i = 0$. Rearrange the sum by corresponding eigenvalues
$$\sum_{i=1}^{n} \mu_i \mathbf{v}_i = \sum_{j=1}^{m} \mathbf{u}_j = 0,$$
where $\mathbf{u}_j \in E(\lambda_j)$. Each $\mathbf{u}_j$ is either an eigenvector to a distinct eigenvalue or 0. Since eigenvectors to distinct eigenvalues are linearly independent all $\mathbf{u}_j = 0 = \sum_{i \in s_j} \mu_i \mathbf{v}_i$, but $\mathbf{v}_i$ are basis for $E(\lambda_j) \Rightarrow \mu_i = 0$.

Theorem 6.11 (Lemma)
If λ is an eigenvalue then $\dim(E(\lambda, T)) \geq 1$.

Theorem 6.12 (Corollary 5.44)
If $T \in \mathrm{End}V$ has $\dim(V)$ distinct eigenvalues, then T is diagonalizable.

7 Inner Product Spaces

We're going to talk about geometry now with length and angles of vectors.

Definition 7.1 (6.3 Inner Product)
Let V be an $\mathbb{R}$- or $\mathbb{C}$-VS, then a function $\langle \cdot | \cdot \rangle : V \times V \to \mathcal{F}$ is called an inner product if

1. $\langle \mathbf{u} | \mathbf{v} \rangle = \overline{\langle \mathbf{v} | \mathbf{u} \rangle}$
2. $\langle \lambda \mathbf{u} + \mu \mathbf{w} | \mathbf{v} \rangle = \lambda \langle \mathbf{u} | \mathbf{v} \rangle + \mu \langle \mathbf{w} | \mathbf{v} \rangle$
3. $\langle \mathbf{v} | \mathbf{v} \rangle \geq 0$ with equality iff $\mathbf{v} = 0$. Item one ensures that this one makes sense, as it only applies to the reals otherwise.

Example 7.1
The typical Euclidean Spaces: $\mathbb{R}^n, \mathbb{C}^n$:

$$\langle (x_1, \dots, x_n) | (y_1, \dots, y_n) \rangle = \sum_{i=1}^{n} x_i \bar{y}_i = \mathbf{x} \cdot \bar{\mathbf{y}}^{\intercal}$$

which can also be scaled by a scalar

$$\sum_{i=1}^{n} c_i x_i \bar{y}_i.$$

Another example is a VS of a continous function, real-valued $[-1, 1]$:

$$\langle f | g \rangle = \int_{-1}^{1} f(x) g(x) \, \mathrm{d}x$$

Theorem 7.1 (Proposition 6.7)
Let V be an inner product space (a VS with an inner product), then

1. $\forall \mathbf{u} \in V, \quad \varphi_{\mathbf{u}} : V \to \mathcal{F}, \ \mathbf{v} \to \langle \mathbf{v} | \mathbf{u} \rangle, \quad \varphi_{\mathbf{u}} \in \mathrm{Hom}_{\mathcal{F}}(V, \mathcal{F})$
2. $\langle 0 | \mathbf{u} \rangle = \langle \mathbf{u} | 0 \rangle = 0 \ \forall \mathbf{u} \in V$
3. $\langle \mathbf{u} | \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u} | \mathbf{v} \rangle + \langle \mathbf{u} | \mathbf{w} \rangle$
4. $\langle \mathbf{u} | \lambda \mathbf{v} \rangle = \bar{\lambda} \langle \mathbf{u} | \mathbf{v} \rangle \ \forall \lambda \in \mathcal{F} \ \mathbf{u}, \mathbf{v} \in V$

Definition 7.2
We say $\mathbf{u}$ and $\mathbf{v}$ are orthogonal, in symbols $\mathbf{u} \perp \mathbf{v}$, if $\langle \mathbf{u} | \mathbf{v} \rangle = 0$.

Remark: 0 is orthogonal to everything. Over $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ 0 is the only vector that is orthogonal to itself.

Definition 7.3 (Norm)
Let V be a VS then a function $\|\cdot\| \to \mathbb{R}_{\geq 0}$ is a norm if

1. $\|\mathbf{v}\| = 0$ iff $\mathbf{v} = 0$
2. $\|\lambda\mathbf{v}\| = |\lambda|\|\mathbf{v}\|, \quad \lambda \in \mathcal{F}$
3. $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

Definition 7.4 (6.8)
Let V be an inner product space then we can define a norm by $\|\mathbf{v}\| = \sqrt{\langle\mathbf{v}|\mathbf{v}\rangle}$.

Proof 7.1
We will now prove the first two conditions of a norm (Definition 7.3)

1. $\|\mathbf{v}\| = 0 = \sqrt{\langle\mathbf{v}|\mathbf{v}\rangle}$, by condition 3 of inner product.
2. $\|\lambda\mathbf{v}\|^2 = \langle\lambda\mathbf{v}|\lambda\mathbf{v}\rangle = \lambda\bar{\lambda}\langle\mathbf{v}|\mathbf{v}\rangle = |\lambda|^2\|\mathbf{v}\|^2$.
3. $\sqrt{\langle\mathbf{v}|\mathbf{v}\rangle}$ is a norm condition and we have

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\| &\leq \|\mathbf{u}\| + \|\mathbf{v}\| \\
\|\mathbf{u}+\mathbf{v}\|^2 &= \langle\mathbf{u}+\mathbf{v}|\mathbf{u}+\mathbf{v}\rangle \\
&= \|\mathbf{u}\|^2 + \langle\mathbf{u}|\mathbf{v}\rangle + \langle\mathbf{u}\bar{|}\mathbf{v}\rangle + \|\mathbf{v}\|^2 \\
&= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathrm{Re}(\langle\mathbf{u}|\mathbf{v}\rangle) \leq \\
&\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2|\langle\mathbf{u}|\mathbf{v}\rangle| \leq \\
\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2 \\
\|\mathbf{u}+\mathbf{v}\| &\leq \|\mathbf{u}\| + \|\mathbf{v}\|.
\end{aligned}$$

Theorem 7.2 (Pythagorean Theorem)
Suppose $\mathbf{u} \perp \mathbf{v}$ then

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$

Proof 7.2

$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2 &= \langle \mathbf{u}+\mathbf{v}|\mathbf{u}+\mathbf{v}\rangle \\ &= \langle \mathbf{u}|\mathbf{u}\rangle + \langle \mathbf{u}|\mathbf{v}\rangle + \langle \mathbf{v}|\mathbf{u}\rangle + \langle \mathbf{v}|\mathbf{v}\rangle \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.\end{aligned}$$

Theorem 7.3 (Lemma)
Let $\mathbf{u}, \mathbf{v} \in V,\ \mathbf{v} \neq 0$ we have

$$\begin{gathered} c = \frac{\langle \mathbf{u}|\mathbf{v}\rangle}{\|\mathbf{v}\|^2} \\ \mathbf{w} = \mathbf{u} - c\mathbf{v}, \end{gathered}$$

then $\mathbf{u} = \mathbf{w} + c\mathbf{v}$ and $\mathbf{w} \perp \mathbf{v}$.

Proof 7.3
Calculate their inner product to prove that they are orthogonal:

$$\begin{aligned}\langle \mathbf{v}|\mathbf{w}\rangle &= \left\langle \mathbf{v} \middle| \mathbf{u} - \frac{\langle \mathbf{u}|\mathbf{v}\rangle}{\|\mathbf{v}\|^2}\mathbf{v} \right\rangle \\ &= \langle \mathbf{v}|\mathbf{u}\rangle - \left\langle \mathbf{v} \middle| \frac{\langle \mathbf{u}|\mathbf{v}\rangle}{\|\mathbf{v}\|^2}\mathbf{v} \right\rangle \\ &= \langle \mathbf{v}|\mathbf{u}\rangle - \frac{\overline{\langle \mathbf{u}|\mathbf{v}\rangle}}{\|\mathbf{v}\|^2}\|\mathbf{v}\|^2 = \langle \mathbf{v}|\mathbf{u}\rangle - \langle \mathbf{v}|\mathbf{u}\rangle = 0.\end{aligned}$$

We call $c\mathbf{v} = \mathrm{proj}_{\mathbf{v}}(\mathbf{u})$.

Theorem 7.4 (Cauchy-Schwarz Inequality)
Let $\mathbf{u}, \mathbf{v} \in V$ then

$$|\langle \mathbf{u}|\mathbf{v}\rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|.$$

Proof 7.4
If $\mathbf{v} = 0$ then the inequality holds.

Assume now $\mathbf{v} \neq 0$ and consider the orthogonal decomposition

$$\begin{aligned}
\mathbf{u} &= \frac{\langle \mathbf{u}|\mathbf{v}\rangle}{\|\mathbf{v}\|^2}\mathbf{v} + \mathbf{w} \\
\|\mathbf{u}\|^2 &= \left\| \frac{\langle \mathbf{u}|\mathbf{v}\rangle}{\|\mathbf{v}\|^2}\mathbf{v} + \mathbf{w} \right\|^2 = \left\| \frac{\langle \mathbf{u}|\mathbf{v}\rangle}{\|\mathbf{v}\|^2}\mathbf{v} \right\|^2 + \|\mathbf{w}\|^2 \\
&= \frac{\langle \mathbf{u}|\mathbf{v}\rangle^2}{\|\mathbf{v}\|^4}\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 \geq \frac{\langle \mathbf{u}|\mathbf{v}\rangle^2}{\|\mathbf{v}\|^2} \\
\|\mathbf{u}\|^2 &\geq \frac{\langle \mathbf{u}|\mathbf{v}\rangle}{\|\mathbf{v}\|^2} \\
\|\mathbf{v}\|^2\|\mathbf{u}\|^2 &\geq \langle \mathbf{u}|\mathbf{v}\rangle^2 \\
\|\mathbf{v}\|\|\mathbf{u}\| &= |\langle \mathbf{u}|\mathbf{v}\rangle|.
\end{aligned}$$

Here we used Pythagorean theorem and that $\|\mathbf{w}\|^2 \geq 0$.

Theorem 7.5 (Parallellogram)

$$\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = 2\Big(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2\Big).$$

Definition 7.5 (6.27 / 6.23)
Æet $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ where V is an inner product space. We say $\mathbf{v}_1, \dots, \mathbf{v}_n$ are orthonormal if

1. $\langle \mathbf{v}_i|\mathbf{v}_j\rangle = 0$ or $\mathbf{v}_i \perp \mathbf{v}_j,\ i \neq j$.
2. $\langle \mathbf{v}_i|\mathbf{v}_i\rangle = 1$ equivalent to $\|\mathbf{v}_i\| = 1$.

(Might also see $\langle \mathbf{v}_i|\mathbf{v}_j\rangle = \delta_{i.j}$ which is the Kronecker delta.)
If $\mathbf{v}_1, \dots, \mathbf{v}_n$ is also a basis we call it an orthonormal basis.

Remark: Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be an orthonormal basis of a VS V. Then $\forall \mathbf{v} \in V$:

$$\mathbf{v} = \langle \mathbf{v}_1|\mathbf{v}_1\rangle \mathbf{v}_1 + \langle \mathbf{v}_1|\mathbf{v}_2\rangle \mathbf{v}_2 + \cdots + \langle \mathbf{v}_1|\mathbf{v}_n\rangle \mathbf{v}_n.$$

Lemma 7.1
Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be a list of orthonormal vectors. Then $\|\lambda_1\mathbf{v}_1 + \cdots + \lambda_m\mathbf{v}_m\|^2 = |\lambda|^2 + \cdots + |\lambda_m|^2$.

Proof 7.5

$$\|\lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m\|^2$$

Note that all vectors are orthogonal to each other as well as any linear combintion of the others. Thus

$$\lambda \langle \mathbf{v}_2 | \mathbf{v}_m \rangle + \mu \langle \mathbf{v}_1 | \mathbf{v}_m \rangle = 0.$$

The sum is now:

$$\|\lambda_1 \mathbf{v}_1 + \cdots + \lambda_{m-1} \mathbf{v}_{m-1}\|^2 + \|\lambda_m \mathbf{v}_m\|^2 = \|\lambda_1 \mathbf{v}_1 + \cdots + \lambda_{m-1} \mathbf{v}_{m-1}\|^2 + |\lambda_m|^2 \|\mathbf{v}_m\|^2$$

but since the vectors are orthonormal $\|\mathbf{v}_m\|^2 = 1$ and we can continue doing this to obtain

$$\sum |\lambda_i|^2 \|\mathbf{v}_i\|^2 = \sum_{i=1}^{m} |\lambda_i|^2.$$

Notice that the standard basis vectors $\mathbf{e}_i$ are an orthonormal basis.

Lemma 7.2 (6.26)
Any list of orthonormals is linear independent.

Proof 7.6
We start by assuming that the linear combination gives zero.

$$\lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m = 0$$

which only happens if all $\lambda_i = 0$.

$$\begin{aligned} \|\lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m\|^2 &= 0 \\ |\lambda_1|^2 + \cdots + |\lambda_m|^2 &= 0 \Rightarrow \lambda_i = 0. \end{aligned}$$

Corollary 7.1 (6.28)
An orthonormal list of length $\dim(V) < \infty$ is a basis.

Gram-Schmidt Orthonormalization

Algorithm for turning a basis into an orthonormal basis: Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be a basis of V. We will construct another, orthonormal, basis $\mathbf{w}_1, \dots, \mathbf{w}_n$.

$$
\begin{aligned}
\mathbf{w}_1 &= \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} \text{hence } \|\mathbf{w}_1\| = 1 \\
\tilde{\mathbf{w}}_2 &= \mathbf{v}_2 - \mathrm{proj}_{\mathbf{w}_1}(\mathbf{v}_2) = \mathbf{v}_2 - \langle \mathbf{v}_2 | \mathbf{w}_1 \rangle \, \mathbf{w}_1 \\
\mathbf{w}_2 &= \frac{\tilde{\mathbf{w}}_2}{\|\tilde{\mathbf{w}}_2\|} \\
\tilde{\mathbf{w}}_3 &= \mathbf{v}_3 - \mathrm{proj}_{\mathbf{w}_1}(\mathbf{v}_3) - \mathrm{proj}_{\mathbf{w}_2}(\mathbf{v}_3) = \mathbf{v}_3 - \langle \mathbf{v}_3 | \mathbf{w}_1 \rangle \, \mathbf{w}_1 - \langle \mathbf{v}_3 | \mathbf{w}_2 \rangle \, \mathbf{w}_2 \\
\mathbf{w}_3 &= \frac{\tilde{\mathbf{w}}_3}{\|\tilde{\mathbf{w}}_3\|} \\
\tilde{\mathbf{w}}_i &= \mathbf{v}_i - \mathrm{proj}_{\mathbf{w}_1}(\mathbf{v}_i) - \mathrm{proj}_{\mathbf{w}_2}(\mathbf{v}_i) - \cdots - \mathrm{proj}_{\mathbf{w}_{i-1}}(\mathbf{v}_i) \\
\mathbf{w}_i &= \frac{\tilde{\mathbf{w}}_i}{\|\tilde{\mathbf{w}}_i\|}
\end{aligned}
$$

8 Determinants

We start with change of basis. Remark: We write $\mathcal{M}(T)$ for the matrix representation of a linear map T, implicitly we are assuming that bases have been fixed. Let $T : V \to W$ and $(\mathbf{v}_i)$ is a basis for V and $(\mathbf{w}_j)$ is a basis for W then we will write from now on $\mathcal{M}(T, (\mathbf{v}_i), (\mathbf{w}_j))$ for the matrix representation of T with respect to the bases $(\mathbf{v}_i)$ and $(\mathbf{w}_j)$.

We have seen that $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$, to write it explicitly we write $\mathcal{M}(ST, (\mathbf{u}_i), (\mathbf{w}_j)) = \mathcal{M}(S, (\mathbf{v}_j), (\mathbf{w}_k))\mathcal{M}(T, (\mathbf{u}_i), (\mathbf{v}_j))$, hence the map is $U \to V \to W$ with $T : U \to V$, $S : V \to W$. We will write $\mathcal{M}(T, (\mathbf{v}_i))$ for $\mathcal{M}(T, (\mathbf{w}_i), (\mathbf{v}_i))$ when $T \in \mathrm{End}(V)$.

Lemma 8.1 (10.5)
Let $(\mathbf{u}_i), (\mathbf{v}_i)$ be the bases for V. Then

$$\mathcal{M}(\mathrm{id}, (\mathbf{u}_i), (\mathbf{v}_i))^{-1} = \mathcal{M}(\mathrm{id}, (\mathbf{v}_i), (\mathbf{u}_i)).$$

Proof 8.1

$$\mathcal{M}(\mathrm{id}, (\mathbf{u}_1), (\mathbf{v}_i))\mathcal{M}(\mathrm{id}, (\mathbf{v}_i), (\mathbf{u}_i)) = \mathcal{M}(\mathrm{id}, (\mathbf{v}_i), (\mathbf{v}_i)) = I_n.$$

Theorem 8.1 (10.7)
Let $T \in \mathrm{End}(V)$ and $(\mathbf{u}_i), (\mathbf{v}_i)$ bases of V. Then

$$\mathcal{M}(T, (\mathbf{u}_i)) = A^{-1}\mathcal{M}(T, (\mathbf{v}_i))A,$$

where $A = \mathcal{M}(\mathrm{id}, (\mathbf{u}_i), (\mathbf{v}_i))$.

Proof 8.2

$$\begin{aligned} A^{-1}\mathcal{M}(T, \mathbf{v}_i)A &= \mathcal{M}(\mathrm{id}, \mathbf{v}_i, \mathbf{u}_i)\mathcal{M}(T, \mathbf{v}_i)\mathcal{M}(\mathrm{id}, \mathbf{u}_i, \mathbf{v}_i) \\ &= \mathcal{M}(T, \mathbf{v}_i, \mathbf{u}_i)\mathcal{M}(T, \mathbf{u}_i, \mathbf{v}_i) = \mathcal{M}(T, \mathbf{u}_i, \mathbf{u}_i). \end{aligned}$$

Definition 8.1
A map $\det : \mathcal{F}^{n\times n} \to \mathcal{F}$ is called a determinant map if

1. $\det\begin{pmatrix} a_1 \\ \vdots \\ \lambda a_i \\ \vdots \\ a_n \end{pmatrix} = \lambda \det\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ and $\det\begin{pmatrix} a_1 \\ \vdots \\ a_i + a_i' \\ \vdots \\ a_n \end{pmatrix} = \det\begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} + \det\begin{pmatrix} a_1 \\ \vdots \\ a_i' \\ \vdots \\ a_n \end{pmatrix}$

2. If two rows are identical then det $= 0$

3. $\det(I_n) = 1$.

Theorem 8.2
For every determinant map it holds that

1. $\det(\lambda A) = \lambda^n \det(A)$

2. If a row of A is zero then $\det(A) = 0$

3. If B results from swappning two rows of A, then $\det(B) = -\det(A)$

4. $\det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ \vdots \\ a_i + \lambda a_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix}$

5. $\det \begin{pmatrix} \lambda_i & & \\ 0 & \ddots & \\ 0 & 0 & \lambda_n \end{pmatrix} = \prod_{i=1}^n \lambda_i$.

Proof 8.3

1. Apply the first item of Definition 8.1 n-times.

2.
$$\det \begin{pmatrix} a_1 \\ \vdots \\ 0 \cdot 0 \\ \vdots \\ a_n \end{pmatrix} = 0\det \begin{pmatrix} a_1 \\ \vdots \\ 0 \\ \vdots \\ a_n \end{pmatrix} = 0.$$

3.
$$0 = \det \begin{pmatrix} a_1 \\ \vdots \\ a_i + a_j \\ \vdots \\ a_i + a_j \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix}$$

   two of the terms are zero as they have two identical rows.

4.

$$\det\begin{pmatrix} a_1 \\ \vdots \\ a_i + \lambda a_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} = \det\begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} + \lambda\det\begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix}$$

where the last term is zero as it has two identical rows.

5. If $\lambda_i \neq 0\ \forall i$ then we can use elementary row operations to transform A into diagonal form in which case we know

$$\det\begin{pmatrix} \lambda & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_n \end{pmatrix} = \det\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix} = \prod_{i=1}^{n} \lambda_i \det(I).$$

If $\lambda_i = 0$ for all i then let i be the largest index such that $\lambda_i = 0$. Then we use row operations to make the ith row all zeroes and the determinant is zero.

Lemma 8.2
$\det(A) \neq 0$ if and only if A is invertible.

Proof 8.4
Using row operations we can transform A into an upper triangular matrix A'. (Equivalently, if $A = \mathcal{M}(T, \mathbf{v}_i)$ then $A' = \mathcal{M}(T, \mathbf{v}_i, \mathbf{u}_i)$). Now we have $\det(A) = \pm\det(A')$ but also the map of such an upper triangular matrix A' is only surjective if all the diagonal values are nonzero and $\det(A') \neq 0$ which is identical to saying that A' is invertible, equivalent to T being invertible and A is invertible: A' is surjective $\Leftrightarrow \lambda_i \neq 0 \Leftrightarrow \det(A') \neq 0 \Leftrightarrow A'$ is invertible $\Leftrightarrow T$ is invertible $\Leftrightarrow A$ is invertible.

Corollary 8.1
If there exists a determinant map then it is unique.

Proof 8.5
Let $A \in \mathcal{F}^{n\times n}$ then there are row operations that transform A into an upper diagonal matrix A' then

$$\det(A) = (-1)^k \det(A')$$

where k is the number of row swaps that were performed. Then we know

$$\det(A) = (-1)^k \prod_{i=1}^{n} \lambda_i$$

which only has one set of λ_i.

Theorem 8.3
There is exactly one determinant map for every field $\mathcal{F}$ and integer $n \geq 1$.

Proof 8.6
By induktion on n:

$$n = 1, \quad \det((a_{11})) = a_{11}$$

For $n > 1$ and $A \in \mathcal{F}^{n \times n}$ consider the submatrices $\hat{A}_{ij}$ given by removing the ith row and jth column of A. Then let

$$\det_n(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det_{n-1}(\hat{A}_{ij}).$$

We claim this is a determinant map (and is the same for any j). To do show we show the items of Definition 8.1.

1. Let $A' = \begin{pmatrix} a_1 \\ \vdots \\ \lambda a_k \\ \vdots \\ a_n \end{pmatrix}$ then

$$\begin{aligned} \det_n(A') &= \sum_{i=1, i \neq k}^{n} (-1)^{i+j} a_{ij} \det_{n-1}(\hat{A}_{ij}) + (-1)^{k+j} \lambda a_{kj} \det(\hat{A}'_{kj}) \\ &= \lambda \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det_{n-1}(\hat{A}_{ij}). \end{aligned}$$

2. Let $a_{kj} = a_{lj}\ \forall j$ and assume $k < l$. Then $\det_{n-1}(\hat{A}_{ij}) = 0$ where $i \neq k$ and $i \neq l$. We're left with

$$\det_n(A) = (-1)^{k+j} a_{kj} \det_{n-1}(\hat{A}_{kj}) + (-1)^{l+j} a_{kj} \det_{n-1}(\hat{A}_{lj})$$

We can get $\hat{A}_{lj}$ from $\hat{A}_{kj}$ by swapping rows $l - k - 1$ times. Then

$$\det_{n-1}(\hat{A}_{kj}) = (-1)^{l-k-1} \det(\hat{A}_{lj}).$$

Now we get

$$\begin{aligned} &(-1)^{k+j+l-k-1} a_{kj} \det(\hat{A}_{lj}) + (-1)^{l+j} a_{kj} \det(\hat{A}_{lj}) \\ &(-1)^{l+j-1} a_{kj} \det(\hat{A}_{lj}) + (-1)^{l+j} a_{kj} \det(\hat{A}_{lj}) = 0. \end{aligned}$$

3. s

$$\begin{aligned}\det_n(I) &= \sum_{i=1}^{n}(-1)^{i+j}S_{ij}\det_{n-1}(\hat{I}_{nij}) \\ &= \det_{n-1}(I_{n-1}) = 1.\end{aligned}$$

Corollary 8.2

$$\det(A) = \det(A^{\intercal})$$

Proof 8.7
We show that

$$\tilde{\det} : \mathcal{F}^{n\times n} \to \mathcal{F},\ A \to \det(A^{\intercal})$$

is a determinant map by checking the conditions of a determinant map and then use uniqueness. This is now an exercise.

Corollary 8.3

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

Proof 8.8
If $\det(B) = 0$, then B is not invertible and it follows that AB is not invertible implying that $\det(AB) = 0$. Assume $\det(B) \neq 0$. Define

$$\tilde{\det}(A) = \frac{\det(AB)}{\det(B)}$$

and show that it is a determinant map.

1.

$$\tilde{\det}(\lambda_i I \cdot A) = \tilde{\det}\begin{pmatrix} a_1 \\ \vdots \\ \lambda a_i \\ \vdots \\ a_n \end{pmatrix}$$

$$\tilde{\det}\begin{pmatrix} a_1 \\ \vdots \\ \lambda a_i \\ \vdots \\ a_n \end{pmatrix} = \frac{\det(\lambda_i I \cdot AB)}{\det(B)} = \frac{\det\begin{pmatrix} a_1 b_1 & a_1 b_n \\ \lambda a_i b_1 & \lambda a_i b_n \\ a_n b_1 & a_n b_n \end{pmatrix}}{\det(B)} = \lambda \tilde{\det}(A)$$

jeg gider ikke mere