LINEAR ALGEBRA

JONAS T. HARTWIG

Revision	Date	Author(s)	Description
0.1	Feb 11, 2024	JH	First version
0.2	Feb 15, 2024	JH	Added Evaluating Polynomials, Real Jordan Form, Determinants

Contents

Date: Version 0.2 from Feb 15, 2024.

1. Fields and Vector Spaces

1.1. **Fields.**

Definition 1.1. A *binary operation* $*$ on a set A is a function $* : A \times A \to A$. Rather than $*(a, b)$ we write $a * b$.

Definition 1.2. An *abelian group* is a set A with a binary operation $*$ on it, satisfying

(i) *(Associative Law)* $(a * b) * c = a * (b * c)$ for all $a, b, c \in A$.
(ii) *(Commutative Law)* $a * b = b * a$ for all $a, b \in A$.
(iii) *(Existence of Identity)* There is an element $e \in A$ such that $e * a = a$ for all $a \in A$. Such an e is called an *identity element (relative to $*$).*
(iv) *(Existence of Inverses)* For each element $a \in A$ there is an element $a' \in A$ such that $a * a' = e$. The element a' is called an *inverse of a (relative to $*$).*

Proposition 1.3. *Let A be an abelian group with binary operation $*$.*

(a) (Uniqueness of Identity) *If e and e' are identity elements of A, then $e = e'$.*
(b) (Uniqueness of Inverse) *Let $a \in A$. If a' and a'' are inverses of A, then $a' = a''$*
(c) (Cancellation Law) *If $a * b = a * c$ then $b = c$.*

Proof. Exercise. □

Definition 1.4. A *field* is a set $\mathbb{F}$ together with two binary operations $+$ (called *addition*) and $\cdot$ (called multiplication) satisfying

(F1) $\mathbb{F}$ with $+$ is an abelian group; 0 denotes the identity element relative to $+$.
(F2) $\mathbb{F} \setminus \{0\}$ with $\cdot$ is an abelian group; $1 \in \mathbb{F} \setminus \{0\}$ denotes the identity element relative to $\cdot$.
(F3) *(Distributive Law)* $(\alpha + \beta) \cdot \gamma = \alpha \cdot \gamma + \beta \cdot \gamma$ for all $\alpha, \beta, \gamma \in \mathbb{F}$.

Rather than saying the "inverse relative to $+/\cdot$" we speak of the *additive/multiplicative inverse*. Similarly, we call 0 the *additive identity*, and 1 the *multiplicative identity*.

Example 1.5. $\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ are fields under usual operations.

Example 1.6. $\mathbb{Z}$ (with usual operations) is not a field because $2 \in \mathbb{Z}$ does not have a multiplicative inverse. Similarly, the set of positive real numbers $\mathbb{R}_{>0}$ (with usual operations) is not a field because there is no additive identity element.

1.2. **Vector Spaces.**

Definition 1.7. Let $\mathbb{F}$ be a field. An $\mathbb{F}$-*vector space* (or *vector space over* $\mathbb{F}$), V, is a set with two operations, a binary operation $+ : V \times V \to V$ called *addition*, and an operation $\mathbb{F} \times V \to V$, $(\lambda, v) \mapsto \lambda v$ called *scaling* such that

(V1) V is an abelian group with respect to $+$. The identity element is denoted by 0_V or just 0. The inverse of $v \in V$ is denoted $-v$.
(V2) $(\lambda\mu)v = \lambda(\mu v)$ for all $\lambda, \mu \in \mathbb{F}$ and $v \in V$.
(V3) $1v = v$ for all $v \in V$.
(V4) $\lambda(u + v) = \lambda u + \lambda v$ for all $\lambda \in \mathbb{F}$, $u, v \in V$.
(V5) $(\lambda + \mu)v = \lambda v + \mu v$ for all $\lambda, \mu \in \mathbb{F}$, $v \in V$.

The following properties follow from the axioms.

Proposition 1.8. *Let V be a vector space. Then*

(i) $0v = 0_V$ *for all* $v \in V$.
(ii) $(-1)v = -v$ *for all* $v \in V$.

Example 1.9. The set $\mathbb{F}^n$ of all column vectors $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$, $a_i \in \mathbb{F}$ is a vector space with operations

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix}, \qquad \lambda \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \lambda a_1 \\ \lambda a_2 \\ \vdots \\ \lambda a_n \end{bmatrix}$$

For positive integers n,

$$\underline{n} = \{1, 2, \ldots, n\}.$$

Example 1.10. Similarly, the set $\mathbb{F}^{m\times n} = M_{m\times n}(\mathbb{F})$ of all $m \times n$-matrices

$$(a_{ij})_{i\in\underline{m},j\in\underline{n}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

with entries $a_{ij} \in \mathbb{F}$, is a vector space with the obvious entry-wise operations.

Definition 1.11. The *transpose* of a matrix $A = (a_{ij})_{i\in\underline{m},j\in\underline{n}}$ is defined by $A^T = (a_{ji})_{j\in\underline{m},i\in\underline{n}}$. Thus it takes $m \times n$-matrices to $n \times m$-matrices.

A column vector can thus be written $\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}^T \in \mathbb{F}^n$.

Example 1.12. Let A be a set and V be a vector space. Let V^A denote the set of all functions from A to V. Then V^A is a vector space with respect to the so-called *pointwise* operations:

$$(f+g)(a) = f(a) + g(a), \qquad (\lambda f)(a) = \lambda f(a), \qquad \forall f, g \in V^A, a \in A, \lambda \in \mathbb{F}.$$

One may regard $\mathbb{F}^n$ as a special case of this construction by identifying a column $\begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}^T \in \mathbb{F}^n$ with the function $v \in \mathbb{F}^{\underline{n}}$ given by $v(i) = v_i$ for $i \in \underline{n} = \{1, 2, \ldots, n\}$. Similarly $\mathbb{F}^{m \times n}$ is really the same thing as $\mathbb{F}^{\underline{m} \times \underline{n}}$.

1.3. **Subspaces.**

Definition 1.13. A subset W of a subspace V is called a *subspace*, if W is itself a vector space with respect to the same operations as in V.

In particular, for W to be a subspace, it is necessary that $w + w' \in W$ for all $w, w' \in W$ (i.e. W is *closed under addition*) and $\lambda w \in W$ for all $w \in W$ and $\lambda \in \mathbb{F}$ (i.e. W is *closed under scaling.*) In fact, these conditions are also sufficient:

Proposition 1.14. *Let W be a subset of a vector space V. Suppose W is closed under the operations of addition and scaling in V. Then W is a subspace of V.*

Example 1.15. Let A be a set and V be a vector space. Let $V^{(A)}$ be the subset of V^A consisting of all functions $f : A \to V$ such that $f^{-1}(\{0_V\})$ is finite (or, equivalently, $f(a)$ is nonzero for at most finitely many $a \in A$). Then $V^{(A)}$ is a subspace of V^A.

Definition 1.16. Let V be a vector space. Let I be a (possibly infinite) index set. Let $(W_i)_{i \in I}$ be a family of subspaces of V. We define

$$\sum_{i \in I} W_i = \{w_{i_1} + w_{i_2} + \cdots + w_{i_k} \mid k \geq 0, i_j \in I, w_{i_j} \in W_{i_j}\}$$

As usual, we should interpret the expression $w_{i_1} + w_{i_2} + \cdots + w_{i_k}$ as the zero vector in V in the case when $k = 0$. In the case when I is a finite index set, such as $\{1, 2, \ldots, m\}$ we have

$$W_1 + W_2 + \cdots + W_m = \{w_1 + w_2 + \cdots + w_m \mid w_i \in W_i\}$$

Furthermore, we let $\bigcap_{i \in I} W_i$ denote the intersection of all the subspaces W_i.

Proposition 1.17. *Let V be a vector space. Let I be an index set. Let $(W_i)_{i \in I}$ be a family of subspaces of V.*

(i) *$\bigcap_{i \in I} W_i$ is a subspace of V.*
(ii) *$\sum_{i \in I} W_i$ is a subspace of V.*

There are two further notions related to the sum of subspaces.

Definition 1.18. Let V be a vector space. Let I be an index set. Let $(W_i)_{i\in I}$ be a family of subspaces of V. We say that the family of subspaces is *independent* if whenever we have

$$w_{i_1} + w_{i_2} + \cdots + w_{i_k} = 0$$

and $w_{i_j} \in W_{i_j}$ for $j = 1, 2, \ldots, k$, then $w_{i_1} = w_{i_2} = \cdots = w_{i_k} = 0$. In this case we denote the sum of subspaces by

$$\oplus_{i\in I} W_i$$

and say that the sum is *direct*.

1.4. **Bases and Coordinates.**

Definition 1.19. Let V be a vector space and A a subset of V.

(i) If A is finite, say[1] $A = \{v_1, v_2, \ldots, v_k\}$, we say that A is *linearly dependent* if there are scalars $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{F}$, not all of them zero, such that

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k = 0. \tag{1.1}$$

(ii) A is *linearly dependent* if there is a finite linearly dependent subset of A.

(iii) A is *linearly independent* if it is not linearly dependent.

(iv) A *linear combination* of elements of A is a vector in V of the form

$$\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_m a_m$$

for some $\lambda_i \in \mathbb{F}$, $a_i \in A$.

(v) The *span* of A is the set of all linear combinations of elements of A:

$$\operatorname{Span} A = \{\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_m a_m \mid m \geq 0, \lambda_i \in \mathbb{F}, a \in A\}$$

(vi) A *spans* V or *is a spanning set for* V if $\operatorname{Span} A = V$.

(vii) A is a *basis for* V if A is linearly independent and spans V.

Example 1.20. The set $A = \{ \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix} \}$ is a linearly independent subset of $\mathbb{F}^3$ but it does not span; the vector $\begin{bmatrix}0\\1\\0\end{bmatrix}$ does not belong to the span of A.

Exercise 1.1. Let A be a set. For each $a \in A$, define a function $e_a : A \to \mathbb{F}$ by

$$e_a(b) = \begin{cases} 1, & \text{if } b = a, \\ 0, & \text{otherwise.} \end{cases}$$

Show that $\{e_a \mid a \in A\}$ is a basis for $\mathbb{F}^{(A)}$.

[1] when naming the elements of a set like this we always mean $v_i \neq v_j \, \forall i \neq j$

Definition 1.21. $\mathbb{F}^{(A)}$ is called the *free vector space on* A and is also denoted by $\mathbb{F}A$.

1.5. **Linear Maps.**

Definition 1.22. Let V and W be vector spaces. A map $T : V \to W$ is *linear* if

$$T(v + v') = T(v) + T(v'), \qquad T(\lambda v) = \lambda T(v) \tag{1.2}$$

for all $v, v' \in V$ and $\lambda \in \mathbb{F}$. The set of all linear maps from V to W is denoted by

$$\mathrm{Hom}(V, W) = \mathrm{Hom}_{\mathbb{F}}(V, W). \tag{1.3}$$

We also put

$$\mathrm{End}(V) = \mathrm{End}_{\mathbb{F}}(V) = \mathrm{Hom}_{\mathbb{F}}(V, V). \tag{1.4}$$

Proposition 1.23. *The identity map* $\mathrm{Id}_V : V \to V$ *is a linear map. If* $T : V \to W$ *and* $S : W \to U$ *are linear maps, then the composition* $S \circ T : V \to U$ *is linear.*

1.6. **Change of Bases.**

Let $\mathcal{B} = (b_1, b_2, \ldots, b_n)$ and $\mathcal{B}' = (b'_1, b'_2, \ldots, b'_n)$ be two bases for a vector space V. We introduce the change-of-basis matrix

$$P = \begin{bmatrix} [b_1]_{\mathcal{B}'} & [b_2]_{\mathcal{B}'} & \cdots & [b_n]_{\mathcal{B}'} \end{bmatrix}$$

This is expressing the *"old" basis* ($\mathcal{B}$) *in the "new" basis* ($\mathcal{B}'$).

2. Quotient Spaces

2.1. **Cosets and their Operations.**

Definition 2.1. Fix a vector space V and a subspace $U \leq V$. A *coset of* U *in* V is a subset of V of the following form:

$$v + U = \{v + u \mid u \in U\} \tag{2.1}$$

The coset $0_V + U$ is called the *trivial coset*. The set of all cosets of U in V is denoted by U/V:

$$V/U = \{v + U \mid v \in V\}. \tag{2.2}$$

The *sum* of two cosets $v + U$ and $v' + U$ is defined by

$$(v + U) + (v' + U) = (v + v') + U. \tag{2.3}$$

The *scaling* of a coset $v + U$ by a scalar $\lambda \in \mathbb{F}$ is defined by

$$\lambda(v + U) = (\lambda v) + U. \tag{2.4}$$

Lemma 2.2. (a) *Two cosets* $v + U$ *and* $v' + U$ *are equals iff* $v - v' \in U$.

(b) *The sum of two cosets does not depend on the choice of representative. That is, if* $v + U = w + U$ *and* $v' + U = w' + U$, *then* $(v + v') + U = (w + w') + U$.

(c) *The scaling of a coset does not depend on the choice of representative. That is, if* $v + U = w + U$, *then* $(\lambda v) + U = (\lambda w) + U$.

Proposition 2.3. *Let V be a vector space and U be a subspace of V. Then V/U becomes a vector space using the sum and scaling of cosets defined above.*

2.2. **Applications.** Besides being a fundamental construction in linear algebra, quotient spaces have applications to multilinear algebra when defining the tensor product, exterior product, and symmetric product (see Section 3). The exterior product is furthermore used in differential geometry.

Below we give two examples from calculus.

Example 2.4. Let $\mathcal{C}(\mathbb{R})$ be the set of continuous function from $\mathbb{R}$ to $\mathbb{R}$ and let $\mathcal{C}^1(\mathbb{R})$ be the set of once continuously differentiable functions from $\mathbb{R}$ to $\mathbb{R}$. These are both vector spaces over $\mathbb{R}$ with pointwise operations $(f+g)(x) = f(x) + f(x)$, $(\lambda f)(x) = \lambda f(x)$. Furthermore, the derivative is a linear operator

$$\frac{d}{dx} : \mathcal{C}^1(\mathbb{R}) \to \mathcal{C}^0(\mathbb{R}).$$

The kernel of this linear map is the set of all functions whose derivative is zero. In other words, it's the constant functions from $\mathbb{R}$ to $\mathbb{R}$. Let us denote this set of constant functions by $\mathbb{R}1$. The map $\frac{d}{dx}$ is furthermore surjective, which follows from the fundamental theorem of calculus. Thus, by the first isomorphism theorem for vector spaces, we have an isomorphism

$$\mathcal{C}^1(\mathbb{R})/\mathbb{R}1 \to \mathcal{C}^0(\mathbb{R}).$$

The inverse of this map is given by the indefinite integral. We see from this that $\int f(x)dx$ is not a function but a coset $F(x) + \mathbb{R}1$, where $F \in \mathcal{C}^1(\mathbb{R})$, $F'(x) = f(x)$.

Example 2.5 (Big ordo notation)**.** Let $n > 0$ be an integer. Let $\mathcal{O}(x^n)$ denote the vector space of all continuous functions $f : \mathbb{R} \to \mathbb{R}$ for which there is a constant $C > 0$ such that $|f(x)| \leq Cx^n$ for all $x > 0$. Then $\mathcal{O}(x^n)$ is a subspace of the vector space $\mathcal{C}(\mathbb{R})$ of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$. An expression such as $g(x) + \mathcal{O}(x^n)$ is thus technically a coset of $\mathcal{O}(x^n)$ in $\mathcal{C}(\mathbb{R})$. So, instead of writing $f(x) = g(x) + \mathcal{O}(x^n)$, it would be better to write $f(x) \in g(x) + \mathcal{O}(x^n)$.

2.3. **Bases and Coordinates.**

Theorem 2.6. *Let $\mathcal{B}$ be a basis for U. Extend $\mathcal{B}$ to a basis $\mathcal{B} \sqcup \mathcal{C}$ for V. Then $\bar{\mathcal{C}} = \{c + U \mid c \in \mathcal{C}\}$ is a basis for V/U.*

2.4. **Isomorphism Theorems.**

Lemma 2.7. *Let $L : V \to W$ be a linear map, and $U \leq V$ such that $U \subset \ker L$. Then there is a well-defined linear map*

$$\bar{L} : V/U \to W, \qquad v + U \mapsto L(v) \tag{2.5}$$

Proof. Indeed, if $v+U = v'+U$, then $v-v' \in U$ hence $L(v)-L(v') = L(v-v') = 0$ since $U \subset \ker L$. This shows that $\bar{L}$ is a well-defined function. That it is linear follows from that L is linear, using the vector space operations in V/U. □

In this situation we say that $\bar{L}$ is *induced* from L.

Theorem 2.8. (i) (First Isomorphism Theorem) *Let $L : V \to W$ be a linear map, and let $K = \ker L$. Then the induced linear map $\bar{L} : V/K \to \operatorname{im} L$ is an isomorphism.*
(ii) (Second Isomorphism Theorem) *Let $U, W \leq V$. Then $(U+W)/W \cong U/(U \cap W)$.*
(iii) (Third Isomorphism Theorem) *Let $U \leq W \leq V$. Then $(V/U)/(W/U) \cong V/W$.*

3. Tensor Products

3.1. **Definition and Basic Properties.** Let V and W be vector spaces, and let $\mathbb{F}^{(V\times W)}$ denote the free vector space on $V \times W$. It has a basis denoted $\{e_{(v,w)} \mid v \in V, w \in W\}$. Let $\mathcal{U}(V,W)$ denote the subspace of $\mathbb{F}^{(V\times W)}$ spanned by the following set:

$$\begin{aligned}&\{e_{(u+v,w)} - e_{(u,w)} - e_{(v,w)} \mid u, v \in V,\ w \in W\}\\ &\cup \{e_{(v,w+z)} - e_{(v,w)} - e_{(v,z)} \mid v \in V,\ w, z \in W\}\\ &\cup \{e_{(\lambda v,w)} - \lambda e_{(v,w)} \mid v \in V, w \in W, \lambda \in \mathbb{F}\}\\ &\cup \{e_{(v,\lambda w)} - \lambda e_{(v,w)} \mid v \in V, w \in W, \lambda \in \mathbb{F}\}.\end{aligned}$$

Definition 3.1. The quotient space $\mathbb{F}^{(V\times W)}/\mathcal{U}(V,W)$ is called the *tensor product of V and W*, denoted $V \otimes W$. For each $v \in V$ and $w \in W$ we also put

$$v \otimes w = e_{(v,w)} + \mathcal{U}(V,W) \in V \otimes W.$$

Definition 3.2. Let V, W, U be vector spaces. A function $\beta : V \times W \to U$ is called *bilinear* if $\beta(v,\cdot) : W \to U$ and $\beta(\cdot, w) : V \to U$ are linear for all $v \in V$ and $w \in W$. The set of all bilinear maps $\beta : V \times W \to U$ is denoted by $\operatorname{Bil}(V \times W, U)$.

Exercise 3.1. Check that $\operatorname{Bil}(V \times W, U)$ is a subspace of the space $U^{V\times W}$ of all functions from $V \times W$ to U.

Proposition 3.3. *The tensor product of V and W has the following properties:*
(i) *The function $\beta_{V,W} : V \times W \to V \otimes W$, $(v,w) \mapsto v \otimes w$ is bilinear. That is,*

$$(u+v)\otimes w = u\otimes w + v\otimes w, \qquad (\lambda v)\otimes w = \lambda(v\otimes w)$$
$$v\otimes(w+z) = v\otimes w + v\otimes z, \qquad v\otimes(\lambda w) = \lambda(v\otimes w)$$

for all $u, v \in V$, $w, z \in W$, $\lambda \in \mathbb{F}$.

(ii) *If U is a third vector space, the map*

$$\mathrm{Hom}(V \otimes W, U) \to \mathrm{Bil}(V \times W, U), \qquad B \mapsto B \circ \beta_{V,W}$$

is an isomorphism of vector spaces.

(iii) *If $\mathcal{B}$ is a basis for V, and $\mathcal{C}$ is a basis for W then the set $\mathcal{B} \otimes \mathcal{C}$, defined by $\{b \otimes c \mid b \in \mathcal{B}, c \in \mathcal{C}\}$, is a basis for $V \otimes W$.*

In particular,

$$\dim V \otimes W = (\dim V)(\dim W).$$

Example 3.4. $\mathbb{F}^m \otimes \mathbb{F}^n$ has a basis $\{e_i \otimes e_j \mid i \in \underline{m},\ j \in \underline{n}\}$.

3.2. **Tensor products of linear maps; Kronecker products.** If $T : V \to W$ and $S : V' \to W'$ are linear maps, the map $V \times V' \to W \otimes W'$ given by $(v, v') \mapsto T(v) \otimes S(v')$ is bilinear, hence there is an induced map denoted

$$T \otimes S : V \otimes V' \to W \otimes W', \qquad v \otimes v' \mapsto T(v) \otimes T(v').$$

Definition 3.5. Let $A = (a_{ij})_{i \in \underline{m}, j \in \underline{n}} \in \mathbb{F}^{m \times n}$, $B = (b_{kl})_{k \in \underline{s}, l \in \underline{t}} \in \mathbb{F}^{s \times t}$. The *Kronecker tensor product* $A \otimes B \in \mathbb{F}^{ms \times nt}$ is defined by

$$A \otimes B = (a_{ij} b_{kl})_{(i,k) \in \underline{m} \times \underline{s}, (j,l) \in \underline{n} \times \underline{t}}$$

Example 3.6.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \otimes B = \begin{bmatrix} aB & bB \\ cB & dB \end{bmatrix}$$

Proposition 3.7. *We have $[T \otimes T'] = [T] \otimes [T']$. In more detail, if $T : V \to W$ and $T' : V' \to W'$ are linear maps, and $\mathcal{B}, \mathcal{C}, \mathcal{B}', \mathcal{C}'$ are bases for V, W, V', W' respectively, then*

$$[T \otimes T']_{\mathcal{B} \otimes \mathcal{B}', \mathcal{C} \otimes \mathcal{C}'} = [T]_{\mathcal{B}, \mathcal{C}} \otimes [T']_{\mathcal{B}', \mathcal{C}'}$$

where in the left hand side, $\otimes$ is the tensor product of linear maps, and in the right hand side $\otimes$ is the Kronecker tensor product of matrices.

4. Primary Decomposition of a Linear Transformation

4.1. **The Minimal Polynomial of a Linear Transformation.** Let $\mathbb{F}[x]$ be the space of polynomials over $\mathbb{F}$ in an indeterminate x.

Definition 4.1. Let $T : V \to V$ be a linear transformation on a finite-dimensional vector space V.

(i) A polynomial $f(x) \in \mathbb{F}[x]$ is an *annihilating polynomial for T* if $f(T) = 0$.

(ii) A monic polynomial $m(x) \in \mathbb{F}[x]$ is a *minimal polynomial for T* if $\deg m(x) \leq \deg f(x)$ for any annihilating polynomial $f(x)$ for T.

Proposition 4.2. *Every linear transformation $T : V \to V$ on a finite-dimensional vector space V has a unique minimal polynomial.*

Notation 4.3. If $T : V \to V$ is a linear map on a finite-dimensional vector space V, we let $m_T(x)$ denote its minimal polynomial.

4.2. **The Primary Decomposition Theorem.**

Theorem 4.4 (The Primary Decomposition Theorem)**.** *Let $T : V \to V$ be a linear transformation on a finite-dimensional vector space V. Let*

$$m_T(x) = p_1(x)^{m_1} p_2(x)^{m_2} \cdots p_d(x)^{m_d} \tag{4.1}$$

be a factorization of the minimal polynomial $m_T(x)$ into irreducible monic polynomials $p_i(x) \in \mathbb{F}[x]$. Then there is a decomposition of V into a direct sum of subspaces:

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_d \tag{4.2}$$

such that each W_i is T-invariant. If furthermore $T_i = T\big|_{W_i} : W_i \to W_i$ denotes the restriction of T to W_i, then $m_{T_i}(x) = p_i(x)^{m_i}$ for each $i = 1, 2, \ldots, d$.

Corollary 4.5. *Let $T : V \to V$ be a linear transformation on a finite-dimensional vector space V. Suppose that the minimal polynomial of T factors into linear factors in $\mathbb{F}[x]$:*

$$m_T(x) = (x - \lambda_1)^{m_1} (x - \lambda_2)^{m_2} \cdots (x - \lambda_d)^{m_d}$$

where $\lambda_i \in \mathbb{F}$ for $i = 1, 2, \ldots, d$. Then V decomposes into a direct sum of the form (4.2) such that each $T_i = T\big|_{W_i} : W_i \to W_i$ has the property that $T_i - \lambda_i \operatorname{Id}_{W_i}$ is nilpotent.

Proof. Immediate from the Primary Decomposition Theorem since $(T_i - \lambda_i \operatorname{Id}_{W_i})^{m_i} = m_{T_i}(T_i) = 0$. □

5. Normal Forms for Linear Transformations

5.1. **Normal Form for a Nilpotent Linear Transformation.** In this section we state and prove a theorem which characterizes nilpotent linear transformations. It is a special case of the Jordan normal form. On the other hand, this special case is the key step in proving the general case.

Definition 5.1. Let n be a positive integer.

(i) An $n \times n$-matrix A is called *nilpotent*, if $A^r = 0$ for some positive integer r.
(ii) A linear transformation $T : V \to V$ of an n-dimensional vector space V is called *nilpotent* if $T^r = 0$ for some positive integer r.

Exercise 5.1. (a) Show that the direct sum of nilpotent matrices is nilpotent.
(b) Show that any matrix similar to a nilpotent matrix is itself nilpotent.

(c) Show that a linear transformation $T : V \to V$ of a finite-dimensional vector space V is nilpotent if and only $[T]_{\mathcal{B}}$ is nilpotent in some (hence, by (b), in every) basis $\mathcal{B}$.

Now we introduce certain special nilpotent matrices.

Definition 5.2. Let

$$J_1 = \begin{bmatrix} 0 \end{bmatrix}, \qquad J_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad J_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \tag{5.1}$$

and, more generally, for positive integers n,

$$J_n = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}. \tag{5.2}$$

The matrix J_n is the *nilpotent (upper-triangular) Jordan block of size* $n \times n$. The lower-triangular versions are obtained by taking the transpose.

Note that $(J_n)^n = 0$ and $(J_n)^{n-1} \neq 0$. [2] The following theorem says that, in fact, any nilpotent matrix is similar to a direct sum of nilpotent Jordan blocks J_d. We formulate it in terms of linear transformations. (Taking $V = \mathbb{F}^n$ and $T = T_A$ (left multiplication by A), we get the matrix case, since $[T_A]_{\mathcal{B}} = P^{-1}AP$, where the columns of P make up the basis $\mathcal{B}$.)

The following terminology will be convenient in the course of the proof:

Definition 5.3. Let V be a vector space and W be a subspace of V. A subset S of V is *linearly independent over* W if $\bar{S} = \{s + W \mid s \in S\}$ is linearly independent in V/W. Similarly, we say that S *spans* V *over* W or *is a basis for* V *over* W if $\bar{S}$ has the corresponding property in V/W.

Exercise 5.2. Show that S is linearly independent in V over W if and only if whenever $\lambda_1 s_1 + \lambda_2 s_2 + \cdots + \lambda_k s_k \in W$ for some $\lambda_i \in \mathbb{F}$ and distinct $s_i \in S$, we have $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$.

[2] Here we use the convention $[0]^0 = [1]$ when $n = 1$.

Theorem 5.4. *Let n be a positive integer, V an n-dimensional vector space, and $N : V \to V$ a nilpotent linear transformation. Then there is a basis $\mathcal{B}$ for V such that, for some integers $d_1 \geq d_2 \geq \cdots \geq d_s \geq 1$:*

$$[N]_{\mathcal{B}} = J_{d_1} \oplus J_{d_2} \oplus \cdots \oplus J_{d_s}$$

$$= \left[\begin{array}{c|c|c|c|c} \begin{matrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{matrix} & & & & \\ \hline & \begin{matrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{matrix} & & & \\ \hline & & \ddots & & \\ \hline & & & \begin{matrix} 0 & 1 \\ 0 & 0 \end{matrix} & \\ \hline & & & & 0 \end{array}\right] \tag{5.3}$$

Proof. If $N = 0$ the result is clear, so assume $N \neq 0$. Let r be the smallest positive integer such that $N^r = 0$. Thus $r \geq 2$. Let $K_j = \ker N^j$ for $j = 0, 1, \ldots, r$. We have

$$\{0\} = K_0 \subseteq K_1 \subseteq K_2 \subseteq \cdots \subseteq K_r = V.$$

Let $\{v_1, v_2, \ldots, v_{p_1}\}$ be a basis for K_r over K_{r-1}. Note that $Nv_i \in K_{r-1}$ for all i. We claim that $\{Nv_1, Nv_2, \ldots, Nv_{p_1}\}$ is linearly independent over K_{r-2}. To see this, suppose

$$\lambda_1 Nv_1 + \lambda_2 Nv_2 + \cdots + \lambda_{p_1} Nv_{p_1} \in K_{r-2}$$

By definition of K_{r-2} this means that

$$N^{r-2}(\lambda_1 Nv_1 + \lambda_2 Nv_2 + \cdots + \lambda_{p_1} Nv_{p_1}) = 0,$$

Moving the N's to the left, we see that

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_{p_1} v_{p_1} \in K_{r-1}.$$

Since $\{v_1, v_2, \ldots, v_{p_1}\}$ is linearly independent over K_{r-1}, this implies all $\lambda_i = 0$.

It is possible that the set does not span K_{r-1} over K_{r-2}. So extend it by adjoining vectors $v_{p_1+1}, v_{p_1+2}, \ldots, v_{p_2}$ so that

$$\{Nv_1, Nv_2, \ldots, Nv_{p_1},\ v_{p_1+1}, v_{p_1+2}, \ldots, v_{p_2}\}$$

is a basis for K_{r-1} over K_{r-2}. If $r = 2$ we stop here. Otherwise, applying N to these vectors we get a set

$$\{N^2v_1, N^2v_2, \ldots, N^2v_{p_1},\ Nv_{p_1+1}, Nv_{p_1+2}, \ldots, Nv_{p_2}\}$$

of vectors in K_{r-2} which are linearly independent over K_{r-3}. The proof of this is the same as in the previous step. So, we may adjoin vectors $v_{p_2+1}, v_{p_2+2}, \ldots, v_{p_3}$ to form a basis

$$\{N^2v_1, N^2v_2, \ldots, N^2v_{p_1},\ Nv_{p_1+1}, Nv_{p_1+2}, \ldots, Nv_{p_2},\ v_{p_2+1}, v_{p_2+2}, \ldots, v_{p_3}\}$$

for K_{r-2} over K_{r-3}. If $r = 3$, we stop. Otherwise we continue. Eventually, we end up with the following set of vectors in V, where we put $v_i^j = N^j v_i$ for brevity:

$$\mathcal{B} = \Big\{ \begin{array}{cccccccc} v_1, & v_2, \ldots, & v_{p_1}, & & & & & \\ Nv_1, & Nv_2, \ldots, & Nv_{p_1}, & v_{p_1+1}, & v_{p_1+2}, \ldots, & v_{p_2}, & & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \\ v_1^{r-1}, & v_2^{r-1}, \ldots, & v_{p_1}^{r-1}, & v_{p_1+1}^{r-2}, & v_{p_1+2}^{r-2}, \ldots, & v_{p_2}^{r-2}, \ldots, & v_{p_{r-1}+1}, \ldots, & v_{p_r} \end{array} \Big\}.$$

The first row is a basis for $V = K_r$ over K_{r-1}. The second row is a basis for K_{r-1} over K_{r-2} and so on; the last row is a basis for K_1. It follows that $\mathcal{B}$ is a basis for V. Each column spans an N-invariant subspace W_j of V, $j = 1, 2, \ldots, p_r$. Let $\mathcal{B}_1$ be the first column, ordered from the bottom up; $\mathcal{B}_2$ the second column, from the bottom up, and so on. In the ordered basis $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_{p_r}$, the matrix of N has the form (5.3), where $(d_1, d_2, \ldots, d_s) = (\dim W_1, \dim W_2, \ldots, \dim W_{p_r})$.

□

5.2. **Jordan Normal Form of a Linear Transformation.**

Definition 5.5. For $\lambda \in \mathbb{F}$ and positive integer n, we put $J_n(\lambda) = J_n + \lambda I_n$. Thus,

$$J_1(\lambda) = \begin{bmatrix} \lambda \end{bmatrix}, \qquad J_2(\lambda) = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \qquad J_3(\lambda) = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}, \tag{5.4}$$

and, more generally, for positive integers n,

$$J_n(\lambda) \begin{bmatrix} \lambda & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{bmatrix}. \tag{5.5}$$

The matrix $J_n(\lambda)$ is the *(upper-triangular) Jordan block of size $d \times d$ with eigenvalue λ*. The lower-triangular versions are obtained by taking the transpose.

Corollary 5.6 (Jordan Normal Form)**.** *Let V be a finite-dimensional vector space. Suppose $T : V \to V$ is a linear map whose minimal polynomial factors into linear factors:*

$$m_T(x) = (x - \lambda_1)^{m_1}(x - \lambda_2)^{m_2} \cdots (x - \lambda_d)^{m_d}.$$

Then there is a basis $\mathcal{B}$ for V such that

$$[T]_{\mathcal{B}} = A_1(\lambda_1) \oplus A_2(\lambda_2) \oplus \cdots \oplus A_d(\lambda_d), \tag{5.6}$$

where each $A_i(\lambda_i)$ is a direct sum of Jordan blocks with eigenvalue λ_i:

$$A_i(\lambda_i) = J_{n_{i1}}(\lambda_i) \oplus J_{n_{i2}}(\lambda_i) \oplus \cdots \oplus J_{n_{ik_i}}(\lambda_i). \tag{5.7}$$

Here $n_{i1} \geq n_{i2} \geq \cdots \geq n_{ik_i} \geq 1$ and $n_{i1} + n_{i2} + \cdots + n_{ik_i} = m_i$.

Definition 5.7. A field $\mathbb{F}$ is *algebraically closed* if every nonzero polynomial in $\mathbb{F}[x]$ factors into linear factors belonging to $\mathbb{F}[x]$.

Corollary 5.8. *If $\mathbb{F}$ is algebraically closed, then every matrix $A \in \mathbb{F}^{n \times n}$ is similar to a block-diagonal matrix with Jordan blocks on the diagonal.*

A matrix is said to be *in Jordan normal form* if it has the form of the right-hand side of (5.6). For example, the following matrix is in Jordan normal form over $\mathbb{F}$ if $\lambda \neq \mu$:

$$A = \begin{bmatrix} \lambda & 1 & 0 & 0 & & & & & \\ 0 & \lambda & 1 & 0 & & & & & \\ 0 & 0 & \lambda & 1 & & & & & \\ 0 & 0 & 0 & \lambda & & & & & \\ & & & & \lambda & 1 & & & \\ & & & & 0 & \lambda & & & \\ & & & & & & \mu & 1 & 0 \\ & & & & & & 0 & \mu & 1 \\ & & & & & & 0 & 0 & \mu \\ & & & & & & & & & \mu \end{bmatrix}$$

Its minimal polynomial is $m_A(x) = (x - \lambda)^4(x - \mu)^3$.

5.3. **Evaluating Polynomials.**

Theorem 5.9. *Let $f(x) \in \mathbb{F}[x]$ and let $A = J_d(\lambda)$ be an $d \times d$ Jordan block with eigenvalue $\lambda \in \mathbb{F}$. Then*

$$f(A) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2!}f''(\lambda) & \cdots & \frac{1}{(d-1)!}f^{(d-1)}(\lambda) \\ 0 & f(\lambda) & f'(\lambda) & \cdots & \frac{1}{(n-2)!}f^{(d-2)}(\lambda) \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & f'(\lambda) \\ 0 & 0 & \cdots & 0 & f(\lambda) \end{bmatrix}$$

Proof. Since both sides are linear in f it suffices to prove the statement for $f(x) = x^n$. By the binomial theorem (which applies to $(A+B)^n$ whenever A and B are commuting matrices),

$$f(A) = (J_d + \lambda I_d)^n = \sum_{k=0}^{n} \binom{n}{k} J_d^k \lambda^{n-k}$$

Notice that $\binom{n}{k}\lambda^{n-k} = \frac{1}{k!}f^{(k)}(\lambda)$, and $J_d^k = \sum_{j=1}^{n-k} E_{j,j+k}$ is the identity matrix shifted northeast by k steps. □

5.4. **Real Jordan Form.** The *real Jordan block of size* $2a \times 2a$ *associated to* $\sigma + i\tau \in \mathbb{C}$ is

$$J_{2a}^{\mathbb{R}}(\sigma + i\tau) = \left[\begin{array}{cc|cc|c|c|cc} \sigma & \tau & 1 & 0 & & & & \\ -\tau & \sigma & 0 & 1 & & & & \\ \hline & & \sigma & \tau & 1 & 0 & & \\ & & -\tau & \sigma & 0 & 1 & & \\ \hline & & & & \ddots & & \ddots & \\ & & & & & \ddots & & \ddots \\ \hline & & & & & & \ddots & \\ & & & & & & & \ddots \\ \hline & & & & & & & \\ & & & & & & & \end{array}\right. \left.\begin{array}{cc} & \\ & \\ \hline & \\ & \\ \hline & \\ & \\ \hline 1 & 0 \\ 0 & 1 \\ \hline \sigma & \tau \\ -\tau & \sigma \end{array}\right] \tag{5.8}$$

Exercise 5.3. Show that the set of matrices of the form $\begin{bmatrix} \sigma & \tau \\ -\tau & \sigma \end{bmatrix}$ where $\sigma, \tau \in \mathbb{R}$, is actually a field isomorphic to $\mathbb{C}$.

Theorem 5.10. *Let $A \in \mathbb{R}^{n\times n}$. Then there is an invertible $P \in \mathbb{R}^{n\times n}$ such that $P^{-1}AP$ has the form $B \oplus C$ where*

$$\begin{aligned} B &= B_1(\lambda_1) \oplus B_2(\lambda_2) \oplus \cdots \oplus B_r(\lambda_r) \\ C &= C_1(\sigma_1, \tau_1) \oplus C_2(\sigma_2, \tau_2) \oplus \cdots \oplus C_s(\sigma_s, \tau_s) \end{aligned} \tag{5.9}$$

where

$$B_j(\lambda_j) = J_{n_{j1}}(\lambda_j) \oplus J_{n_{jk_j}}(\lambda_j), \tag{5.10}$$

$n_{j1} \geq n_{j2} \geq \cdots \geq n_{jk_j} \geq 1$, *and*

$$C_j(\sigma_j, \tau_j) = J^{\mathbb{R}}_{m_{j1}}(\sigma_j, \tau_j) \oplus J^{\mathbb{R}}_{m_{j2}}(\sigma_j, \tau_j) \oplus \cdots \oplus J^{\mathbb{R}}_{m_{jl_j}}(\sigma_j, \tau_j), \tag{5.11}$$

$m_{j1} \geq m_{j2} \geq \cdots \geq m_{jl_j} \geq 1$.

Proof. It suffices to consider the case when the minimum polynomial $m_A(x)$ for A has the form $p_1(x)^m$ where $p_1(x)$ is monic quadratic without real roots. First suppose the real part of the roots is zero. Then

$$m_A(x) = (x^2 + \tau^2)^m$$

for some $\tau \in \mathbb{R}$. By the Primary Decomposition Theorem, $\mathbb{C}^n = W_+ \oplus W_-$ where the $W_\pm$ are A-invariant subspaces on which A has minimum polynomial $(x \pm i\tau)^m$. Define a function

$$\phi : \mathbb{C}^n \to \mathbb{C}^n$$

by component-wise conjugation. We claim that $\phi(W_\pm) = W_\mp$. Let $v \in W_+$. Then $(A + i\tau)^m v = 0$. Applying ϕ and using that A is real we get $(A - i\tau)^m \phi(v) = 0$. Thus $\phi(v) \in W_-$. Similarly for the other case. In fact we see from this that ϕ restricts to a bijective $\mathbb{R}$-linear map $W_+ \to W_-$ commuting with A. In particular, $\dim_{\mathbb{R}} W_+ = \dim_{\mathbb{R}} W_-$. So $\dim_{\mathbb{C}} W_+ = \dim_{\mathbb{C}} W_-$, therefore n must be even, say $n = 2k$, $k = \dim_{\mathbb{C}} W_\pm$.

Let $(v_1, v_2, \ldots, v_k)$ be a $\mathbb{C}$-basis for W_- in which A restricted to W_- is in Jordan normal form. The real and imaginary parts of the basis vectors v_j are:

$$w_j = \frac{1}{2}\big(v_j + \phi(v_j)\big), \qquad w'_j = \frac{1}{2i}\big(v_j i\phi(v_j)\big),$$

We have $w_j, w'_j \in \mathbb{R}^n$. Now we describe A in the basis $\mathcal{B} = (w_1, w'_1, w_2, w'_2, \ldots, w_k, w'_k)$ for $\mathbb{R}^n$. For this, we apply A to w_j and w'_j, and expand the result in $\mathcal{B}$ and inspect the coefficients. Fix j. There are two cases: Either $Av_j = i\tau v_j$ or $Av_j = i\tau v_j + v_{j-1}$. In the former we have

$$Aw_j = \frac{1}{2}\big(Av_j + \phi(Av_j)\big) = i\tau\frac{1}{2}\big(v_j - \phi(v_j)\big) = -\tau w_j$$

and similarly

$$Aw'_j = \tau w_j$$

which means we have a block $\begin{bmatrix} 0 & \tau \\ -\tau & 0 \end{bmatrix}$ on the diagonal. In the other case we get

$$Aw_j = -\tau w_j + w_{j-1}, \qquad Aw'_j = \tau w_j + w'_{j-1}.$$

This means we get a block as before plus a 2×2 identity matrix above it. □

6. Determinants and Invariant Factors

Let V be n-dimensional and $T : V \to V$ be linear. Every linear map from a one-dimensional space to itself is given by multiplication by a scalar.

By homework, $\wedge^n V$ is one-dimensional, and T induces a linear map $\wedge^n T : \wedge^n V \to \wedge^n V$.

Definition 6.1. The *determinant of* T, denoted $\det T \in \mathbb{F}$ is given by

$$\wedge^n T = (\det T)\,\mathrm{Id}_V\,. \tag{6.1}$$

Let S_n denote the symmetric group, consisting of all permutations (bijections) of the set $\{1, 2, \ldots, n\}$. Since $e_i \wedge e_j = -e_j \wedge e_i$, there is a function $\operatorname{sgn} : S_n \to \{\pm 1\}$, called the *sign function*, such that

$$e_{\sigma(1)} \wedge e_{\sigma(2)} \wedge \cdots \wedge e_{\sigma(n)}.$$

It is not hard to see that $\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)$ for $\sigma, \tau \in S_n$. (In other words, sgn is a group homomrphism.)

Theorem 6.2. *Let $(e_1, e_2, \ldots, e_n)$ be a basis for V. Let $T_{ij} \in \mathbb{F}$ be the entries of $[T]_{\mathcal{B}}$, i.e. $T(e_i) = \sum_j T_{ij} e_j$. Then*

$$\det(T) = \sum_{\sigma \in S_n} (\operatorname{sgn}\sigma) T_{1\sigma(1)} T_{2\sigma(2)} \cdots T_{n\sigma(n)}. \tag{6.2}$$

Proof. We have

$$\begin{aligned}
&\det(T) e_1 \wedge e_2 \wedge \cdots \wedge e_n \\
&\quad = T(e_1 \wedge e_2 \wedge \cdots \wedge e_n) \\
&\quad = T\Big(\big(\sum_{i_1=1}^{n} T_{1 i_1} e_{i_1} \big) \wedge \big(\sum_{i_2=1}^{n} T_{2 i_2} e_{i_2} \big) \wedge \cdots \wedge \big(\sum_{i_n=1}^{n} T_{n i_n} e_{i_n} \big) \Big) \\
&\quad = \sum_{1 \le i_1, i_2, \ldots, i_n \le n} T_{1 i_1} T_{2 i_2} \cdots T_{n i_n} e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_n} \\
&\quad = \sum_{\sigma \in S_n} T_{1\sigma(1)} T_{2\sigma(2)} \cdots T_{n\sigma(n)} e_{\sigma(1)} \wedge e_{\sigma(2)} \wedge \cdots \wedge e_{\sigma(n)} \\
&\quad = \Big(\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) T_{1\sigma(1)} T_{2\sigma(2)} \cdots T_{n\sigma(n)} \Big) e_1 \wedge e_2 \wedge \cdots \wedge e_n
\end{aligned}$$

where we used that $v_1 \wedge v_2 \wedge \cdots \wedge v_n$ is linear in each factor, $e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_n} = 0$ unless the i_j are all distinct hence define a permutation, and the definition of $\operatorname{sgn}(\sigma$. $\square$

Exercise 6.1. Modify the proof to prove that for every $\tau \in S_n$ we have

$$\operatorname{sgn}(\tau) \det(T) = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) T_{\tau(1)\sigma(1)} T_{\tau(2)\sigma(2)} \cdots T_{\tau(n)\sigma(n)}. \tag{6.3}$$

Proposition 6.3. *If $S, T : V \to V$ are linear then*

$$\det(S \circ T) = \det(S) \det(T), \qquad \det(\mathrm{Id}_V) = 1. \tag{6.4}$$

Proof. Let $\mathrm{Id} = \mathrm{Id}_{\wedge^n V}$ for brevity. We have $\det(S \circ T) \mathrm{Id} = \wedge^n (S \circ T) = (\wedge^n S) \circ (\wedge^n T) = (\det S) \mathrm{Id} \circ (\det T) \mathrm{Id} = (\det S)(\det T) \mathrm{Id}$. Also, $\det(\mathrm{Id}_V) \mathrm{Id} = \wedge^n (\mathrm{Id}_V) = \mathrm{Id}$. $\square$

Definition 6.4. For $A \in \mathbb{F}^{n \times n}$ we define $\det(A) = \det(T_A)$, where $T_A : \mathbb{F}^n \to \mathbb{F}^n$ is multiplication by A.

Proposition 6.5. *The function* $\det : \mathbb{F}^{n \times n} \to \mathbb{F}$ *is uniquely characterized by the properties*

(i) $\det$ *is linear in each column*
(ii) $\det$ *is alternating in the columns (i.e. switching two gives a minus sign)*
(iii) $\det I_n = 1$

Proof. That the properties hold follow from the definition. That such a function is unique can be proved using elementary column operations. $\square$