MATH 314 (SPRING 2022)

Basics/Prerequisites: Sets, Maps, Relations,...

- The axiomatic point of view
  - All entities are sets, and for any sets X, A one has: Either $X \in A$, or $X \notin A$. Notation:
    $$A := \{X \mid X \in A\} \quad [\text{read: } A \text{ is the set of all } X \text{ such that } X \in A]$$
  - The Zermelo-Fraenkel (ZF) system of axioms insures, among other things, that for any sets $A, B, X, Y, X', Y' \ldots$ one has:
    - $A = B \overset{\text{def}}{\longleftrightarrow} \forall\, X\colon\ X \in A$ iff $X \in B$; $\ A \subset B \overset{\text{def}}{\longleftrightarrow} \forall\, X\colon\ X \in A \Rightarrow X \in B$.
    - For a property $p(X)$ of elements $X \in A$, one has: $A_{p(X)} := \{X \in A \mid p(X)\ true\}$ is a set. Hence $A_{p(X)} \subset A$, and the empty set $\emptyset := \{\, X \in A \,|\, X \neq X \,\} \subset A$ for every set A.
    - The usual operations $\cup, \cap, \times, \setminus$, the complement $\complement_X(A) := X \setminus A$, the symmetric difference $A \triangle B := (A \setminus B) \cup (B \setminus A)$, etc., are defined.
    - $\{A\} := \{X \,|\, X = A\}$ is a set, hence the successor $s(A) := A \cup \{A\}$ of A is a set.
    - $A \notin A$ for every set A, and $s(A) = s(B)$ iff $A = B$.
    - The power set $\mathcal{P}(A) := \{X \,|\, X \subset A\}$ is a set.
    - $(X, Y) := \{\{X\}, \{X, Y\}\}$ is a set, and $(X, Y) = (X', Y')$ iff $X = X'$, $Y = Y'$. Further, $A \times B := \{(X, Y) \,|\, X \in A, Y \in B\}$ is a set, the (Cartesion) product of A, B.
- Correspondences & Maps
  - Correspondences, functional correspondences
  - Graph of a map, composition of maps, associativity of composition
  - injective, surjective, bijective maps, inverse of a bijective map, etc.
- Equivalence/Order relations
  - Definitions, examples
  - The equivalence classes $\overline{x} := \{x' \in X \mid x \sim x'\} \subset X$ of $\sim$ satisfy: $\overline{x} \cap \overline{y} \neq \emptyset \ \Leftrightarrow\ \overline{x} = \overline{y}$, and the set of equivalence classes $X/\!\sim\, = \overline{X} := \{\overline{x} \mid x \in X\}$ defines a partition of X.
  - **Characterization:** *Giving $\sim$ on X is the same as giving a partition $X = \cup_{\alpha \in I} X_\alpha$ of X.*
- The natural numbers $\mathbb{N}$

  The set of natural numbers $\mathbb{N} = \{\, \emptyset, s(\emptyset), s\big(s(\emptyset)\big), \ldots \}$ is ***intuitively*** & ***axiomatically***, (via the Peano axioms) Google it!, see **Prere-LA**, the smallest set containing $\emptyset$ *as an element,* and is closed under successor, i.e., $X \in \mathbb{N} \Rightarrow s(X) \in \mathbb{N}$. [**Note** that $\emptyset$ is not the successor of any set, because $A \in s(A)$ for every set A, thus $s(A)$ is non-empty for every set A. In particular, $\emptyset$ being the empty set, cannot be a successor set!]
  - Finally identify/denote: $\emptyset \leftrightarrow 0$, $s(\emptyset) \leftrightarrow 1$, $s\big(s(\emptyset)\big) \leftrightarrow 2$, $s\Big(s\big(s(\emptyset)\big)\Big) \leftrightarrow 3$, $\ldots$

A particular but *extremely important instance* of the axiomatic of $\mathbb{N}$ is the so called:

- Induction Principle: *Let $N \subset \mathbb{N}$ be a set with $0 \in N$ and $s(N) \subset N$. Then $N = \mathbb{N}$.*

An important consequence of the above Induction Principle is the following:

Theorem 0.1. *Let $\{\mathcal{P}_n\}_{n\in\mathbb{N}}$ be a set of assertions index by $n \in \mathbb{N}$. In order to prove that all the assertion $\mathcal{P}_n$ are true, it suffices to do the following:*

- *Verification Step: Prove that $\mathcal{P}_0$ is true.*
- *Induction Step: Prove that for all $n \in \mathbb{N}$, one has: $\mathcal{P}_0, \ldots, \mathcal{P}_n \Rightarrow \mathcal{P}_{n+1}$.*

- Recall the definitions of the addition and the multiplication on $\mathbb{N}$. **Note** that $+$ and $\cdot$ are by no means *symmetric in the arguments,* hence one must prove:

Theorem 0.2. *The addition $+$ and multiplication $\cdot$ on $\mathbb{N}$ satisfy: $\forall\, k, m, n \in \mathbb{N}$ one has:*

1) (i) *associativity & commutativity of addition: $k+(m+n) = (k+m)+n$, $m+n = n+m$.*
   (ii) *associativity & commutativity of multiplication: $k\cdot(m\cdot n) = (k\cdot m)\cdot n$, $m\cdot n = n\cdot m$.*
   (iii) *0 is neutral element for addition and 1 is neutral element for multiplication.*
   (iv) *distributivity of multiplication w.r.t. addition: $k\cdot(m+n) = k\cdot m + k\cdot n$.*
4) *Cancellation: $m + k = n + k \Rightarrow m = n$, respectively $m \cdot k = n \cdot k \Rightarrow m = n$ for $k \neq 0$.*

- Define on $\mathbb{N}$ a (binary) relation $\leqslant$ by $m \leqslant n \overset{\text{def}}{\longleftrightarrow} \exists k \in \mathbb{N}$ s.t. $n = m + k$.

Theorem 0.3. *The relation $\leqslant$ on $\mathbb{N}$ is an ordering relation satisfying the following:*

1) *Compatibility with $+$, $\cdot$ and Cancellation, i.e., for all $k, m, n \in \mathbb{N}$ one has the following:*
   $m + k \leqslant n + k \Leftrightarrow m = n$, respectively $m \cdot k \leqslant n \cdot k \Leftrightarrow m \leqslant n$ for $k \neq 0$.
2) *The ordering $\leqslant$ is a total well ordering, i.e., the following hold:*
   a) *For every $n \in \mathbb{N}$ one has: $\mathbb{N} = \{m \in \mathbb{N} \mid m < n\} \cup \{n\} \cup \{m \in \mathbb{N} \mid n < m\}$*
   In particular, for all $m, n \in \mathbb{N}$ one has: Either $m \leqslant n$ or $n \leqslant m$.
   b) *If $\Sigma \subset \mathbb{N}$ is non-empty, there exists a unique $n_0 \in \Sigma$ s.t. $n \in \Sigma \Rightarrow n_0 \leqslant n$.*

An consequence of the above is the (generalized) Induction Principle:

Theorem 0.4. *Let $\{\mathcal{P}_n\}_{n\geqslant n_0}$ be a sequence of assertions. In order to prove that all the assertion $\mathcal{P}_n$, $n \geq n_0$, are true, it suffices to do the following:*

- *Verification Step: Prove that $\mathcal{P}_{n_0}$ is true.*
- *Induction Step: Prove that for all $n \geqslant n_0$, one has: $\mathcal{P}_{n_0}, \ldots, \mathcal{P}_n \Rightarrow \mathcal{P}_{n+1}$.*

- The integer numbers $\mathbb{Z}$

  Review the definition of the integer numbers $\mathbb{Z}$, and the addition $+$, multiplication $\cdot$ and the total ordering $\leqslant$ on $\mathbb{Z}$. Finally recall that $\mathbb{Z}, +, \cdot, \leqslant$ is a totally ordered ring with the cancelation property.

- The field of rational numbers $\mathbb{Q}$

  Review the definition of the rational numbers $\mathbb{Q}$, and the addition $+$, multiplication $\cdot$ and the total ordering $\leqslant$ on $\mathbb{Q}$. Finally recall that $\mathbb{Q}, +, \cdot, \leqslant$ is a totally ordered field.

- The field of real numbers $\mathbb{R}$

  Review the definition of the field of real numbers $\mathbb{R}$, and the addition $+$, multiplication $\cdot$ and the total ordering $\leqslant$ on $\mathbb{R}$. Finally recall that $\mathbb{R}, +, \cdot, \leqslant$ is a totally ordered field.

1. Composition laws & basic algebraic structures

1.1. Basic definitions/Facts.

Definition 1.1. A (binary) composition law on a set $X \neq \emptyset$ is any map $\psi : A \times X \to X$.
Notation. Usually, $\psi(x, y)$ is denoted by $x * y$, or $x \circ y$, or $x \cdot y$, etc. [read "x composed with y"].

Definition 1.2. Let $*, \circ$ be a composition laws on X.

I) We says that $*$ satisfies/has:

- associativity, if $(x * y) * z = x * (y * z)$ $\forall\, x, y, z \in X$.
- commutativity, if $x * y = y * x$ $\forall\, x, y \in X$.
- neutral element $e \in X$, if $x * e = x = e * x$ $\forall\, x \in X$.
- Suppose that $*$ has a neutral element $e \in X$. We say that $x' \in X$ is an inverse of $x \in X$ (w.r.t. $*$), if $x * x' = e = x' * x$; if so, we say that $x \in X$ is invertible.

II) We say that $\circ$ is distributive w.r.t. $*$, if $z \circ (x * y) = (z \circ x) * (z \circ y)$, $(x * y) \circ z = (x \circ z) * (y \circ z)$ $\forall\, x, y, z \in X$.

Proposition 1.3. *Let $*, \circ$ be a composition laws on X. TFH:*

1) *If $e, e' \in X$ are neutral elements for $*$, then $e = e'$.*
2) *If $*$ is associative, and $x', x'' \in X$ are inverse elements of x w.r.t. $*$, then $x' = x''$.*
3) *Let $\circ$ be distributive w.r.t. $*$, $e \neq \epsilon$ be neutral elements for $*$, respectively $\circ$, and ϵ be invertible w.r.t. $*$. Then e cannot be invertible w.r.t. $\circ$.*

Proof. To 1): One has $e' \overset{\text{why}}{=} e' * e \overset{\text{why}}{=} e$. To 2): One has: $x' \overset{\text{why}}{=} x' * e \overset{\text{why}}{=} x' * (x * x'') \overset{\text{why}}{=} (x' * x) * x'' \overset{\text{why}}{=} x''$. To 3): By contradiction, let $e' \in X$ be the inverse of e w.r.t. $\circ$. Then $\epsilon \overset{\text{why}}{=} e \circ e' \overset{\text{why}}{=} (e * e) \circ e' \overset{\text{why}}{=} (e \circ e') * (e \circ e') \overset{\text{why}}{=} \epsilon * \epsilon$, i.e., $\epsilon = \epsilon * \epsilon$. Hence if ϵ' is the inverse of ϵ w.r.t. $*$, then $e = \epsilon * \epsilon' \overset{\text{why}}{=} (\epsilon * \epsilon) * \epsilon' \overset{\text{why}}{=} \epsilon * (\epsilon * \epsilon') \overset{\text{why}}{=} \epsilon$, contradiction! □

Definition 1.4. Consider sets endowed with a composition laws.

1) A (commutative) monoid is any set M endowed with a composition law $\cdot$ which is associative (and commutative), and has a neutral element e_M.
2) An (abelian, or commutative) group is any (commutative) monoid $G, \cdot$ in which every $x \in G$ has an inverse w.r.t. $\cdot$, usually denoted x^{-1}.
3) A subgroup of a group $G, \cdot$ is any subset $G' \subset G$ with $e_G \in G'$ such that $\forall\, x, y \in G'$ one has: $x \cdot y \in G'$ and $x^{-1} \in G'$. Equivalently, $G', \cdot$ is itself a group.

Example 1.5.

a) $+$ and $\cdot$ are composition laws on $\mathbb{N}$, and $\mathbb{N}, +$ and $\mathbb{N}, \cdot$ are commutative monoids (WHY). What are neutral elements and the investible elements in $\mathbb{N}, +$ and $\mathbb{N}, \cdot$?

b) Let $X := \mathcal{P}(A)$ be the power set of a given set A. Then $X, \cap$ and $X, \cup$ are commutative monoids (WHY). What are neutral, rest. invertible elements in these monoids, respectively?

(!) Moreover, X endowed with the symmetric difference $A \triangle B \overset{\text{def}}{:=} (A \backslash B) \cup (B \backslash A)$ is a commutative group (WHY).

c) The difference $a * b \overset{\text{def}}{:=} a - b$ is a composition law on $\mathbb{Z}$, which is not associative, nor commutative (WHY). Does $-$ have a neutral element?

The same question about $a \circ b :\overset{\text{def}}{=} |a-b|$.

d) Let $\leqslant$ be a total ordering on a set X. Then $x*y :\overset{\text{def}}{=} \min(x,y)$ and $x \circ y := \max(x.y)$ are associative and commutative (WHY). Do these composition laws have neutral elements?

e) Let X be a non-empty set. Then $\mathcal{F}(X) :\overset{\text{def}}{=} \{f \mid f : X \to X \text{ map}\}$ endowed with the usual composition of maps $\circ$ is a monoid (WHY), which is commutative iff $|X| = 1$.

Further, $\text{Bij}(X) :\overset{\text{def}}{=} \{f \in \mathcal{F}(X) \mid f \text{ bijective}\}$ consists of precisely the invertible elements in the monoid $\mathcal{F}(X), \circ$ (WHY), and $\text{Bij}(X), \circ$ is a group, which is Abelian iff $|X| \leqslant 2$ (WHY).

(•) **The permutation group S_m.** If $X = \{1, \dots, m\}$, we set $S_m :\overset{\text{def}}{=} \text{Bij}(X), \circ$ and call it the permutation group of n elements. The elements $\sigma \in S_m$ are presented in the form:

$$\sigma := \begin{pmatrix} 1 & \dots & m \\ i_1 & \dots & i_m \end{pmatrix}, \quad i_k = \sigma(k) \text{ for } 1 \leqslant k \leqslant m.$$

Definition 1.6. A (commutative) ring is a set R endowed with two composition laws, the addition and the multiplication, usually denoted $+$, respectively $\cdot$, satisfying the following:

i) $R, +$ is a commutative group, i.e., $+$ is associative, commutative, has a neutral element, denoted 0_R, called the zero (element) of R, and every $x \in R$ has an an inverse w.r.t. $+$, called additive inverse of x, denoted $-x$.

ii) $R, \cdot$ is a (commutative) monoid, with neutral element, denoted 1_R, called the unit of R. The invertible elements w.r.t. $\cdot$ are called units of R, and are denoted $R^\times$.

iii) The multiplication $\cdot$ is distributive w.r.t. addition $+$, i.e., $\forall\, x, y, z \in R$ one has:

$$z \cdot (x+y) = z \cdot x + z \cdot y, \quad (x+y) \cdot z = x \cdot z + y \cdot z$$

1) Let $R, +, \cdot$ be a ring. We define/say that:
- $r \in R$ is a zero divisor, if $\exists\, x, x' \neq 0_R$ in R with $x \cdot r = 0_R = r \cdot x'$.
- $x \in R$ is called nilpotent, if there is $n \in \mathbb{N}_{>0}$ such that $x^n = 0_R$.
- $R, +, \cdot$ is an (integral) domain, if R is comm., $0_R \neq 1_R$ and 0_R is the only zero divisor.

2) A subring of $R, +, \cdot$ is a subset $R' \subset R$ such that $1_R \in R'$ and R' is closed w.r.t. $+$ and $\cdot$ and $R', +, \cdot$ is itself a ring.

3) A ring $R, +, \cdot$ is called a division ring or skew field, if $1 \neq 0_R$, and every $x \in R$, $x \neq 0_R$ is invertible w.r.t. multiplication. Commutative skew fields are called simply fields.

4) A (skew) subfield of a (skew) field $F, +, \cdot$ is a subset $F' \subset F$ which is a subring of F' which is itself a (skew) field.

Remark 1.7. A commutative ring $R, +, \cdot$ is an (integral) domain iff R has the cancelation property w.r.t. multiplication, that is: $\forall\, r, x, y \in R$, $r \neq 0_R$ one has: $r \cdot x = r \cdot y$ iff $x = y$ (WHY).

Example 1.8 (**NOTE:** The rings & (skew) fields below will be "officially" defined later)**.**

a) $\mathbb{Z}, +, \cdot$ is a domain, and $\mathbb{Z}^\times = \{\pm 1\}$ (WHY).

b) $\mathbb{Z}, +, \cdot$ is a subring of the ring of polynomials $R = \mathbb{Z}[t]$ in the variable t over $\mathbb{Z}$ (WHY). What are 0_R, 1_R, and $R^\times$?

c) The set $\mathcal{M}_2(\mathbb{Z}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{Z} \right\}$ of 2×2 matrices over $\mathbb{Z}$ endowed with addition and multiplication of matrices is a *non-commutative* ring (WHY). What is $\mathcal{M}_2(\mathbb{Z})^\times$?
The set of diagonal matrices Δ is a subring of $\mathcal{M}_2(\mathbb{Z})$ (WHY).

b) $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, are fields, and the Hamiltonian quaternions $\mathbb{H}$ is a skew field. Google it!

Definition 1.9. Let $M,\cdot$ be a monoid, $G,\cdot$ be a group, and $R,+,\cdot$ be a ring.

1) For $x\in M$, set $x^0=e_M$, $x^1=x$, and inductively, $x^{n+1}:=x^n\cdot x$ for all $n\in\mathbb{N}$.
2) For $x\in G$ let x^{-1} be its inverse. We denote x^n as usually, and $x^{-n}:\overset{\text{def}}{=}(x^{-1})^n$.

(!) Note: The composition laws of abelian groups id usually denoted $+$.

If $G,+$ is an abelian group with $0_G:\overset{\text{def}}{=}e_G$, the above rules become: $0x:\overset{\text{def}}{=}0_G$, $1x:\overset{\text{def}}{=}x$, $(n+1)x:\overset{\text{def}}{=}nx+x\ \forall\, n\in\mathbb{N}$. Further, $(-n)x:\overset{\text{def}}{=}n(-x)$, $(mn)x=m(nx)$ for all $n\in\mathbb{Z}$.

3) Applying the above rules to the abelian group $R,+$ and the monoid $R,\cdot$ we have:

(i) $0\,r:=0_R$, $(\pm n)r:=n(\pm r)$, $(mn)r=m(nr)$ (WHY) (ii) $r^1:=r$, and $r^{n+1}:\overset{\text{def}}{=}r^n\cdot r$ (WHY).

Proposition 1.10 (Computation rules). *The following hold:*

1) *Let $M,\cdot$ be a monoid. Then $x^m\cdot x^n=x^{m+n}$ and $(x^n)^m=x^{nm}$ for all $m,n\in\mathbb{N}$. Further, if $x\cdot y=y\cdot x$, then $(xy)^n=x^n\cdot y^n$ for all $n\in\mathbb{N}$.*
2) *Let $R,+,\cdot$ be a ring. Then one has the following:*
   a) $0_R\cdot x=0_R=x\cdot 0_R$, $(-1_R)\cdot x=-x=x\cdot(-1_R)$ *for all $x\in R$. Hence $R=\{0_R\}$ iff $0_R=1_R$.*
   b) $(\sum_{i=1}^m a_i)\cdot(\sum_{j=1}^n b_j)=\sum_{i,j}a_i\cdot b_j$ *for all $a_i,b_j\in R$, $m,n\in\mathbb{N}_{>0}$.*
   c) *Let $a\cdot b=b\cdot a$ for some $a,b\in R$. Then one has the **binomial formula**:*

$$(a+b)^1=a+b,\quad (a+b)^n=a^n+\sum_{k=1}^{n-1}\binom{n}{k}a^{n-k}\cdot b^k+b^n\quad \text{for } n\in\mathbb{N}_{>1}.$$

Proof. To 1): **Ex** (**Hint**: double induction on m,n, etc.)
To 2a): $0_R\cdot x\overset{\text{why}}{=}(0_R+0_R)\cdot x\overset{\text{why}}{=}0_R\cdot x+0_R\cdot x$, hence $0_R\cdot x+(-0_R\cdot x)=(0_R\cdot x+0_R\cdot x)+(-0_R\cdot x)$, thus $0_R=0_R\cdot x$ (WHY), etc. Further, $0_R=(1_R-1_R)\cdot x=x+(-1_R)\cdot x$, hence $(-1_R)\cdot x=-x$ (WHY), etc. To 2b): **Ex** (make double induction on m,n, etc.) To 2c): **Ex** (make induction on n, and use the binomial identity $\binom{n}{k}=\binom{n-1}{k}+\binom{n-1}{k-1}$ —which itself can be proved either directly, or by induction (HOW) □

There are important procedures which lead to/produce new algebraic structures involving existing ones. We describe below are three such important constructions.

1.1.1. Monoids/groups/rings of functions.

Let $\cdot$ be a composition law on a non-empty set T, and $\mathrm{Maps}(X,T)$ be the set of maps $f:X\to T$. Define on $\mathrm{Maps}(X,T)$ the composition law:

$$(f\odot g)(x):\overset{\text{def}}{=}\Big(f(x)\Big)\cdot\Big(g(x)\Big),\qquad \forall\, x\in X$$

called the composition law (induced by $\cdot$) *on the T-valued maps.* The following hold:

(i) $\odot$ is associative iff $\cdot$ is so (WHY).
(ii) $\odot$ is commutative iff $\cdot$ is so (WHY).
(iii) $\odot$ has a neutral element $e_\odot$ iff $\cdot$ has neutral element, say $e\in T$, and if so, the constant e-map $f_e:X\to T$, $f_e(x)=e$ is the neutral element of $\odot$ (WHY).
(iv) $f\in\mathrm{Maps}(X,T)$ has an inverse w.r.t. $\odot$ iff all $t\in f(X)\subset T$ have inverse elements in T. If so, then $g\colon X\to T$, $g(x):=\Big(\text{inverse of } t=f(x) \text{ in } T\Big)$ is the inverse of f w.r.t. $\odot$ (WHY).

Proposition 1.11. *In the above notation, for X non-empty, the following hold:*

1) *Let T endowed with a composition law $\cdot$ be given. Then $\mathrm{Maps}(X,T),\odot$ is an (abelian) monoid/group iff $T,\cdot$ is an (abelian) monoid/group.*

If so, $\mathrm{Maps}(X,T),\circ$ *is called the monoid/group of* T*-valued functions on* X*.*

2) *Let* R *endowed with two composition laws* $+,\cdot$ *be given. Then* $R,+,\cdot$ *is a (commutative) [non-trivial] ring iff* $\mathrm{Maps}(X,R),\oplus,\odot$ *is a (commutative) [non-trivial] ring.*
If so, $\mathrm{Maps}(X,T),\oplus,\odot$ *is called the ring of* T*-valued functions on* X*.*

Proof. ... □

Remark 1.12. The following hold:

1) First, $T,\cdot$ is the trivial monoid/group iff $T=\{e_T\}$ iff $\mathrm{Maps}(X,T)$ is trivial (WHY). Similarly, R is the *trivial ring* iff $R=\{0_R\}$ iff $0_R=1_R$ iff $\mathrm{Maps}(X,R)$ is the trivial ring (WHY), respectively $M,+$ is the trivial R-module iff $\mathrm{Maps}(X,M)$ is trivial.
2) Second, if R is non-trivial, and $|X|>1$, then there are $f,g\in\mathrm{Maps}(X,R)$ such that $f,g\neq 0_{\mathrm{Maps}(X,R)}$ and with $f\cdot g=0_{\mathrm{Maps}(X,R)}$ (WHY).

Ex. For a ring $R,+,\cdot$ on has: $\mathrm{Maps}(X,R)$ is a domain/skew field iff R is so and $|X|=1$.

1.1.2. (Direct) Products of Monoids/Groups/Rings.

Let $*_i$ be composition laws on X_i. Define on $X=X_1\times\ldots\times X_n$ the component-wise composition law $*$ on X by $(x_1,\ldots,x_n)*(y_1,\ldots,y_n):\overset{\text{def}}{=}(x_1*_1y_1,\ldots,x_n*_ny_n)$. If M_i are R-modules, define an outer R-multiplication on $M=M_1\times\ldots\times M_n$ by $r\cdot(x_1,\ldots,x_n):\overset{\text{def}}{=}(r\cdot x_1,\ldots,r\cdot x_n)$.

Proposition 1.13. *In the above notation,* $*$ *on* X *is* (i) *associative;* (ii) *commutative;* (iii) *has neutral element iff each* $*_i$ *on* X_i *does so,* $1\leqslant i\leqslant n$*. Further, one has:*

1) $e=(e_1,\ldots,e_n)\in X$ *is the neutral element of* $*$ *iff each* $e_i\in X_i$ *is the neutral element for the composition law* $*_i$*,* $1\leqslant i\leqslant n$*.*
2) $x=(x_1,\ldots,x_n)\in X$ *is invertible w.r.t.* $*$ *iff each* x_i *is invertible w.r.t.* $*_i$*,* $1\leqslant i\leqslant n$*. If so, then* $x^{-1}=(x_1^{-1},\ldots,x_n^{-1})$*.*
3) *If* $\circ_1,\ldots,\circ_n$ *are further composition laws on* $X_1,\ldots,X_n$ *respectively, and* $\circ$ *is the corresponding component-wise composition law on* X*, one has:* $\circ$ *is distributive w.r.t.* $*$ *iff each* $\circ_i$ *is distributive w.r.t.* $*_i$ *on* X_i*,* $1\leqslant i\leqslant n$*.*

In particular, a product of (commutative) monoids/groups/rings endowed with the component-wise composition laws is a (commutative) monoid/group/ring.

Proof. **Ex** ... □

Ex. A product $R:=R_1\times\cdots\times R_n$ of domains/skew fields is a domain/skew field iff $n=1$.

Ex. Show that $x=(x_1,\ldots,x_n)\in R$ is a zero divisor/nilpotent iff each $x_i\in R_i$ is so.

1.1.3. Power Series and Polynomials over a commutative ring R with 1_R.

Definition 1.14. Let R be a commutative ring with 1_R.

1) A symbol of the form $\Sigma_n a_nt^n$ is called a (formal) power series with coefficients $a_n\in R$ in the variable t. Let $R[[t]]$ be the set of formal power series over R.
2) $\Sigma_n a_nt^n\in R[[t]]$ is called a polynomial if $a_n=0_R$ for $n\gg 0$. Let $R[t]\subset R[[t]]$ be the set of polynomials. For $p:=p(t)=\Sigma_n a_nt^n\in R[t]$, define the degree by:

$\deg(p)=-\infty$ if $a_n=0_R$ for all n, $\quad\deg(p)=\max\{n\,|\,a_n\neq 0_R\}$ if $a_n\neq 0_R$ for some n.

Examples 1.15. Let $R = \mathbb{Q}$, or more general $n1_R$ is invertible in R for all $n \in \mathbb{N}_{>0}$.

a) The geometric power series $f(t) = \Sigma_n t^n = 1 + t + \cdots + t^n + \ldots$

b) The formal log around 1 is $\log(1+t) = \Sigma_n (-1)^n \frac{t^{n+1}}{n+1} = t - \frac{t^2}{2} + \frac{t^3}{3} - \cdots + (-1)^{n-1}\frac{t^n}{n} + \ldots$

c) The formal exponential is $\exp(t) = \sum_n \frac{t^n}{n!} = 1 + \frac{t}{1!} + \frac{t^2}{2!} + \cdots + \frac{t^n}{n!} + \ldots$

d) $f(t) = 1 + t - t^2$ is a polynomial, with $\deg(f) = 2$.

e) The power series at a), b), c) are not polynomials.

Definition 1.16. Let R be an arbitrary commutative ring. Define on $R[\![t]\!]$ the addition $+$, multiplication $\cdot$, and an (outer) multiplication by $a \in R$ as follows:

- *Addition*: $\left(\sum_i a_i t^i\right) + \left(\sum_i b_i t^i\right) :\overset{\text{def}}{=} \sum_i (a_i + b_i) t^i$
- *Multiplication*: $\left(\sum_i a_i t^i\right) \cdot \left(\sum_i b_i t^i\right) :\overset{\text{def}}{=} \sum_i c_i t^i$, where $c_i := \sum_{k=0}^{i} a_k b_{i-k}$ for all $i \in \mathbb{N}$.

a) $(\Sigma_n a_n t^n) + (\Sigma_n b_n t^n) :\overset{\text{def}}{=} \Sigma_n c_n t^n$ where $c_n :\overset{\text{def}}{=} a_n + b_n$.

b) $(\Sigma_n a_n t^n) \cdot (\Sigma_n b_n t^n) :\overset{\text{def}}{=} \Sigma_n c_n t^n$, where $c_n :\overset{\text{def}}{=} \sum_{i+j=n} a_i \cdot b_j$.

c) (*outer*) *multiplication by elements* $a \in R$: $a \cdot (\Sigma_n a_n t^n) :\overset{\text{def}}{=} \Sigma_n a \cdot a_n t^n$.

Proposition 1.17. *Let R be a commutative ring. The following hold:*

1) *The set of formal power series $R[\![t]\!]$ with coefficients from R endowed with the addition and multiplication of series is a commutative R-algebra. Precisely one has:*

   (i) *The addition is associative, commutative, $0_{R[\![t]\!]} :\overset{\text{def}}{=} \Sigma_n 0_R t^n$ is neutral element, and $-\Sigma_n a_n t^n :\overset{\text{def}}{=} \Sigma_n (-a_n) t^n$ is the inverse of $\Sigma_n a_n t^n$ w.r.t. addition.*

   (ii) *The multiplication is assoc., comm., and $1_{R[\![t]\!]} :\overset{\text{def}}{=} 1_R + \Sigma_{n>0} 0_R t^n$ is neutral element. Moreover, $\Sigma_n a_n t^n$ is invertible w.r.t. $\cdot$ iff $a_0 \in R$ is invertible w.r.t. $\cdot$ in R.*

   (iii) *The multiplication is distributive w.r.t. addition.*

2) *The set of polynomials $R[t] \subset R[\![t]\!]$ is closed w.r.t. addition, multiplication, multiplication by $a \in R$, and $0_{R[\![t]\!]}, 1_{R[\![t]\!]} \in R[t]$, hence $R[t] \subset R[\![t]\!]$ is an R-subalgebra. Further one has:*

   a) $\deg(p+q) \leqslant \max\left(\deg(p), \deg(q)\right)$.

   b) $\deg(p \cdot q) \leqslant \deg(p) + \deg(q)$, *and equality holds if R is a domain, e.g., a field.*

Proof. To 1): (i): **Ex** . . . (easy direct checking).

To (ii): *Commutativity*: Let $(\Sigma_n a_n t^n) \cdot (\Sigma_n b_n t^n) :\overset{\text{def}}{=} \Sigma_n c_n t^n$ and $(\Sigma_n b_n t^n) \cdot (\Sigma_n a_n t^n) :\overset{\text{def}}{=} \Sigma_n d_n t^n$ with $c_n = \sum_{i+j=n} a_i \cdot b_j = \sum_{i+j=n} b_i \cdot a_j = d_n$ (WHY). Hence $(\Sigma_n a_n t^n) \cdot (\Sigma_n b_n t^n) = (\Sigma_n b_n t^n) \cdot (\Sigma_n a_n t^n)$ (WHY).

Associativity: One has $(\Sigma_n a_n t^n) \cdot (\Sigma_n b_n t^n) :\overset{\text{def}}{=} \Sigma_n e_n t^n$ with $e_n = \Sigma_{i+j=n} a_i \cdot b_j$. Hence $\big((\Sigma_n a_n t^n) \cdot (\Sigma_n b_n t^n)\big) \cdot (\Sigma_n c_n t^n) = (\Sigma_n e_n t^n) \cdot (\Sigma_n c_n t^n) = \Sigma_n d_n t^n$, where $d_n :\overset{\text{def}}{=} \Sigma_{l+k=n} e_l \cdot c_k = (\Sigma_{i+j=l} a_i \cdot b_j) \cdot c_k = \Sigma_{i+j+k=n} a_i \cdot b_j \cdot c_k$. Deduce that the multiplication os series is associative (WHY).

Neutral element: $1_{R[\![t]\!]} = 1 + \Sigma_{n>0} 0_R t^n$ is neutral element (**Ex** . . .)

Invertibility of $\Sigma_n a_n$: To $\Rightarrow$: Let $(\Sigma_n a'_n t^n) \cdot (\Sigma_n a_n t^n) = 1_{R[\![t]\!]}$. Then $a_0 \cdot a'_0 = 1_R$ (WHY), hence $a_0 \in R^\times$.

To $\Leftarrow$: Let $a'_0 \in R$ satisfy $a_0 \cdot a'_0 = 1_R$. We compute $(a'_n)_n$ such that $1_{R[\![t]\!]} = (\Sigma_n a_n t^n) \cdot (\Sigma_n a'_n t^n)$. Equivalently, one must have $a_0 \cdot a'_0 = 1_R$ and $\sum_{i+j=n} a_i \cdot a'_j = 0_R$ for all $n > 0$ (WHY). Computing a'_1: One has $n = 1$, hence must solve the equation $a_1 \cdot a'_0 + a_0 \cdot a'_1 = 0_R$ in the unknown a'_1. One has: $a_0 \cdot a'_1 = -a_1 \cdot a'_0$, hence we get: $a'_1 = -a_1 \cdot a'_0 \cdot a_0^{-1}$ (WHY). By induction, let $a'_1, \ldots, a'_{n-1}$ be computed s.t. $\Sigma_{i+j=l} a_i \cdot a'_j = 0_R$

for $0 \leqslant l < n$. Then $\Sigma_{i+j=n} a_i \cdot a'_j = 0_R$ iff $a_0 \cdot a'_n = -(a_{n-1} \cdot a'_0 + \cdots + a_1 \cdot a'_{n-1})$ (WHY), implying that $a'_n \overset{\text{def}}{:=} (a_n \cdot a'_0 + \cdots + a_1 \cdot a'_{n-1}) \cdot a_0^{-1}$ (WHY).
To (iii): **Ex**...(direct verification using the distributivity of $\cdot$ w.r.t. $+$ in R). **Ex**...

To 2): Let $p(t) = \Sigma_n a_n t^n, q(t) = \Sigma_n b_n t^n \in R[t]$ be given, $N_p \overset{\text{def}}{:=} \deg(p)$, $N_q \overset{\text{def}}{:=} \deg(g)$, hence $a_n = 0_R$ for $n > N_p$, and $b_n = 0_R$ for $n > N_q$ (WHY). First, if $p(t) + q(t) = h(t) = \Sigma_n c_n t^n$, then $c_n = a_n + b_n$. Hence if $n > N_p, N_q$, then $a_n = 0_R = b_n$, thus $c_n = 0_R$ (WHY). Conclude that $h(t) \in R[t]$ (WHY), and $\deg(h) \leqslant \max\big(\deg(p), \deg(q)\big)$. Second, $p(t) \cdot q(t) = h(t) = \Sigma_n c_n t^n$ with $c_n = \sum_{i+j=n} a_i \cdot b_j$. Hence if either $i > N_p$ or $j > N_q$, then $a_i \cdot b_j = 0_R$ (WHY). OTOH, if $n > N_p + N_q$, and $n = i + j$, then $i > N_p$ or $j > N_q$ must hold (WHY). Hence for $n > N_p + N_q$ one has $c_n = 0_R$ (WHY). Conclude that $h(t) \in R[t]$, and $\deg(h) \leqslant N_p + N_q = deg(p) + \deg(q)$. Finally, if R is a domain, one has: Since $a_{N_p} \neq 0_R \neq b_{N_q}$, one has that $c_{N_p+N_q} = a_{N_p} \cdot a_{N_q} \neq 0_R$ (WHY), hence $\deg(p \cdot q) = \deg(p) + \deg(q)$. □

Proposition 1.18. *For a commutative $R, +, \cdot$ the following are equivalent:*
(i) *R is a domain.* (ii) *$R[t]$ is a domain.* (iii) *$R[[t]]$ is a domain.*

Proof. Since $R \subset R[t] \subset R[[t]]$, it follows that (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) (WHY). Hence to complete the proof, it is sufficient to prove that (i) $\Rightarrow$ (iii) (WHY); equivalently, to prove that given $f(t) = \sum_i a_i t^i \neq 0_{R[[t]]} \neq \sum_j b_j t^j = g(t)$, one has $f(t) \cdot g(t) \neq 0_{R[[t]]}$. Now, since $f(t), g(t) \neq 0_{R[[t]]}$, one has by definition: $\Sigma_f := \{i \in \mathbb{N} \,|\, a_i \neq 0_R\}$, $\Sigma_g := \{j \in \mathbb{N} \,|\, b_j \neq 0_R\}$ are non-empty (WHY). Let $m = \min \Sigma_f$, $n = \Sigma_g$, i.e., $a_m \neq 0_R$, $a_k = 0$ for $k < m$, and $b_n \neq 0_R$, $b_l = 0$ for $l < n$. In particular, one has that $f(t) = a_m t^m + a_{m+1} t^{m+1} + \cdots = t^m(a_m + a_{m+1} t + \dots)$, $g(t) = b_n t^n + b_{n+1} t^{n+1} + \cdots = t^n(b_n + b_{n+1} t + \dots)$, and therefore, one has

$$h(t) = f(t)\, g(t) = t^m(a_m + a_{m+1} t + \dots) \cdot t^n(a_n + a_{n+1} t + \dots) = a_m b_n t^{m+n} + c_{m+n+1} t^{m+n+1} + \dots \text{ (WHY)}$$

Since $a_m b_n \neq 0_R$ (WHY), we conclude that $f(t)g(t) \neq 0_{R[[t]]}$. Thus finally, $R[[t]]$ is an integral domain. □

1.2. (Homo)morphisms.

- Given sets $X, *$ and $X', *'$ endowed with composition laws, a map $f : X \to X'$ is called a (homo)morphism if f is compatible with the composition laws, i.e.,
$$f(x * y) = f(x) *' f(y), \quad \forall\, x, y \in X.$$
- We say that $f : X \to X'$ is an isomorphism if there is a (homo)morphisms $g : X' \to X$ such that $f \circ g = \mathrm{id}_X$, $g \circ f = \mathrm{id}_{X'}$.
- **Note:** An isomorphism f must be bijective (WHY), and the morphism g is actually the inverse map of f (WHY).
- In particular, we speak about monoid, group, ring, fields (iso)morphisms.
- **Note:** If $f : X \to X'$ is a morphism involving monoids, it is assumed that $f(e_X) = e_{X'}$ (if not otherwise explicitly stated). Hence if $f : R \to R'$ is a ring morphism, we assume that $f(1_R) = 1_{R'}$ (if not otherwise explicitly stated).

Remark 1.19. In the above notation, if $f : X \to X'$ is a bijective morphism, then f is automatically an isomorphism, i.e., the inverse map $g := f^{-1}$ satisfies $g(x' *' y') = g(x') * g(y')$ and g maps neutral elements / inverse elements to such. Indeed, since f is bijective, one has:

(i) $g(x' *' y') = g(x') * g(y')$ iff $f\big(g(x' *' y')\big) = f\big(g(x') * g(y')\big) \overset{\text{why}}{=} f\big((g(x')) *' f\big(g(y)\big)$ (WHY). OTOH, since $f \circ g = \mathrm{id}_{X'}$, both sides equal $x' *' y'$ (WHY).
(ii) If the neutral element $e \in X$ exists, then $e' := f(e)$ is neutral element of $*'$. Indeed, if $x' = f(x) \in X'$, then $e' *' x' = x' = x' *' e'$ iff $g(e' *' x') = g(x') = g(x' *' e')$ (WHY), etc.
(iii) Finally, $x' * x = e = x * x'$ in X, then $f(x) *' f(x') = e' = f(x') *' f(x)$ in X', etc.

Proposition 1.20. *Let G, H be groups, and $f: G \to H$ be a homomorphism. Then one has $f(e_G) = e_H$ and $f(x^{-1}) = \big(f(x)\big)^{-1}$ for all $x \in G$. Further, the following hold:*

1) $\mathrm{Ker}(f) := \{x \in G \,|\, f(x) = e_H\} \subset G$ *is a subgroup, and* $\mathrm{Ker}(f) = \{e_G\}$ *iff f is injective.*
2) *If $G' \subset G$ and $H' \subset H$ are subgroups, so are $f(G') \subset H$ and $f^{-1}(H') \subset G$.*
3) *If $f: G \to H$ is bijective, then f is an isomorphism.*

Proof. First, let $e' := f(e_G)$. Since $e_G \cdot e_G = e_G$, one has $e' = f(e_G) = f(e_G) \cdot f(e_G) = e' \cdot e'$ in H (WHY). Hence composing with e'^{-1} in H, get: $e_H = e'$ (WHY). Second, since $e_G = x^{-1} \cdot x$, get: $e_H = f(e_G) = f(x^{-1}) \cdot f(x)$. Hence $f(x^{-1})$ is the inverse of $f(x)$ in H (WHY).

To 1): If $x, y \in \mathrm{Ker}(f)$, then $f(x \cdot y) = f(x) \cdot f(y) = e_H \cdot e_H = e_H$, and $f(x^{-1}) = f(x)^{-1} = e_H$ (WHY). Hence $\mathrm{Ker}(f) \subset G$ is a subgroup. We next prove that f is injective iff $\mathrm{Ker}(f) = \{e_G\}$. To "$\Rightarrow$": One has $x \in \mathrm{Ker}(f) \overset{\text{def}}{\longleftrightarrow} e_H = f(x)$. Hence since $e_H = f(e_G)$ too, and f injective, one must have $x = e_G$ (WHY). Thus $\mathrm{Ker}(f) = \{e_G\}$. To "$\Leftarrow$": For $x < 1, x_2 \in G$ one has: $f(x_1) = f(x_2)$ iff $f(x_1) \cdot f(x_2)^{-1} = e_H$ (WHY) iff $f(x_1 x_2^{-1}) = e_H$ (WHY) iff $x_1 x_2^{-1} \in \mathrm{Ker}(f)$ (WHY). Hence $\mathrm{Ker}(f) = \{e_G\}$ implies $x_1 \cdot x_2^{-1} = e_G$, thus $x_1 = x_2$ (WHY).

To 2): First we prove that if $G' \subset G$ is a subgroup, so is $f(G') \subset H$. Indeed, for $f(x), f(y) \in f(G') \subset H$ one has: $f(x) \cdot f(y) = f(x \cdot y) \in f(G')$ (WHY), $f(x)^{-1} = f(x^{-1}) \in f(G')$ (WHY), hence $f('G) \subset H$ is a subgroup. Second, we prove that if $H' \subset H$, then $G' = f^{-1} \subset G$ is a subgroup: Since $e_H \in H'$ (WHY), one has $e_G \in G'$ (WHY). Further, $x, y \in G'$ iff $f(x), f(y) \in H'$ (WHY), and if so, $f(x \cdot y) = f(x) \cdot f(y) \in H'$ and $f(x^{-1}) = f(x)^{-1} \in H'$ (WHY), hence $x \cdot y, x^{-1} \in G'$ (WHY). Conclude that $G' \subset G$ is a subgroup.

To 3): **Ex** . . . □

Proposition 1.21. *Let $R, +, \cdot$ and $S, +, \cdot$ be rings with 1_R and 1_S, and $f: R \to S$ be a ring homomorphism with $f(1_R) = 1_S$. Then $f(0_R) = 0_S$, $f(x^{-1}) = \big(f(x)\big)^{-1}$ if $x \in R^\times$, and TFH:*

1) $\mathrm{Ker}(f) := \{x \in R \,|\, f(x) = 0_S\} \subset R$ *is an ideal, and* $\mathrm{Ker}(f) = \{0_R\}$ *iff f is injective.*
2) *If $R' \subset R$ and $S' \subset S$ are subrings, then so are $f(R) \subset S$ and $R' := f^{-1}(S') \subset R$.*
3) *If $f: R \to S$ is an isomorphisms iff f is bijective.*

Proof. To 1): Since $f: R \to S$ is, in particular, a morphism of the abelian groups $R, +$ to $S, +$, it follows that $\mathrm{Ker}(f) \subset R$ is an (abelian) subgroup of the abelian group $R, +$. Further, for $x \in R$, $r \in \mathrm{Ker}(f)$ one has: $f(x \cdot r) = f(x) \cdot f(r) = f(r) \cdot 0_S = 0_S$ (WHY), hence $x \cdot r \in \mathrm{Ker}(f)$; similarly, $r \cdot x \in \mathrm{Ker}(f)$. Hence $\mathrm{Ker}(f) \subset R$ is an ideal. To 2), 3) **Ex** (proof are very similar to the proofs of the assertions in the case of groups...). □

1.3. **Factor Groups / Rings.**

A) **Modular arithmetic**

Recall the division with remainder by $n \in \mathbb{Z}$ in the ring $\mathbb{Z}$:

For every $x \in \mathbb{Z}$ there exist unique $q \in \mathbb{Z}$, $r \in \mathbb{N}$, $0 \leqslant r < |n|$, such that $x = nq + r$.

[**Hint**: Prove this fact by induction on $|a| \ldots$]

From now on, suppose that $n \in \mathbb{Z}$, $n > 0$ and set $n\mathbb{Z} :\overset{\text{def}}{=} \{nk \,|\, k \in \mathbb{Z}\}$, $\overline{a} :\overset{\text{def}}{=} a + n\mathbb{Z}$.

- The subsets $\overline{x} \subset \mathbb{Z}$ satisfy: $\overline{x} \cap \overline{y} \neq \emptyset$ iff $\overline{x} = \overline{y}$ iff $\exists\, k \in \mathbb{Z}$ s.t. $x - y = nk$ (WHY).
- In particular, $(\overline{x})_{x \in \mathbb{Z}}$ is a partition of $\mathbb{Z}$, thus $x \sim_n y \overset{\text{def}}{\longleftrightarrow} \overline{x} = \overline{y}$ is an equivalence relation on $\mathbb{Z}$ having equivalence classes $\mathbb{Z}/\!\sim_n = \{\overline{x} \,|\, x \in \mathbb{Z}\} = \{\overline{r} \,|\, 0 \leqslant r < n\}$ (WHY).
- Define the addition and multiplication on $\mathbb{Z}/n\mathbb{Z}$ by $\overline{x} + \overline{y} :\overset{\text{def}}{=} \overline{x + y}$ and $\overline{x} \cdot \overline{y} :\overset{\text{def}}{=} \overline{xy}$.

Claim. *$+$ and $\cdot$ are well defined on $\mathbb{Z}/n\mathbb{Z}$,* i.e., if $\overline{x}=\overline{x}'$, $\overline{y}=\overline{y}'$, one has:

$$\overline{x+y}=\overline{x'+y'}, \qquad \overline{xy}=\overline{x'y'}.$$

[**Hint**: By $(*)$ above, one has: $\overline{x}=\overline{x}'$, $\overline{y}=\overline{y}'$ iff $\exists\, k,l\in\mathbb{Z}$ s.t. $x-x'=nk$, $y-y=nl$, implying $(x+y)-(x'+y')=n(k-l)$, $xy-x'y'=xy-(x-nk)(y-nl)=n(xl+yk-nkl)$ (WHY). Thus $\overline{x+y}=\overline{x'+y'}$, $\overline{xy}=\overline{x'y'}$ (WHY), etc.]

Theorem 1.22. *$\mathbb{Z}/n\mathbb{Z},+,\cdot$ is a commutative ring having $0_{\mathbb{Z}/n\mathbb{Z}}=\overline{0}$, $1_{\mathbb{Z}/n\mathbb{Z}}=\overline{1}$, and the map $f:\mathbb{Z}\to\mathbb{Z}/n\mathbb{Z}$, $a\mapsto\overline{a}=a+n\mathbb{Z}$ is a surjective ring homomorphism with $\mathrm{Ker}(f)=n\mathbb{Z}$.*

Terminology. *Computation in $\mathbb{Z}/n\mathbb{Z}$ is called* modular arithmetic $(\mathrm{mod}\, n)$.

Proof. **Ex** . . . □

Example 1.23.

- Consider $\mathbb{Z}/30\mathbb{Z}=\{\overline{0},\dots,\overline{29}\}$. Find all solution of the equation $\overline{6}x+\overline{3}=\overline{5}$ in $\mathbb{Z}/30\mathbb{Z}$.
- $\overline{a}\in\mathbb{Z}/120\mathbb{Z}$ is a zero divisor iff a is divisible by 2, or 3, or 5 (WHY); nilpotent iff $a=2^k$ (WHY).
- $\mathbb{Z}/m\mathbb{Z}$ is a domain iff $m:=p$ is a prime number, and if so, $\mathbb{F}_p:=\mathbb{Z}/p\mathbb{Z}$ is a field (WHY).

Here is a very interesting and important fact:

Theorem 1.24 (Chinese Remainder Thm). *Let $n,m\in\mathbb{N}_{>0}$ be given. Then the map*

$$f:\ \mathbb{Z}/(mn)\mathbb{Z}\ \to\ \mathbb{Z}/m\mathbb{Z}\times\mathbb{Z}/n\mathbb{Z},\quad a+mn\mathbb{Z}\mapsto(a+m\mathbb{Z},a+n\mathbb{Z})$$

is a ring homomorphism, which is an isomorphism iff n,m are relatively prime.

Proof. **Ex**: verify that f is a ring homomorphism. Further, $|\mathbb{Z}/(mn)\mathbb{Z}|=mn, |\mathbb{Z}/m\mathbb{Z}|=m, |\mathbb{Z}/n\mathbb{Z}|=n$ (WHY), hence the rings $\mathbb{Z}/(mn)\mathbb{Z}$ and $\mathbb{Z}/m\mathbb{Z}\times\mathbb{Z}/n\mathbb{Z}$ are finite and have the same cardinality (WHY). Thus f is an isomorphism iff f is a bijection (WHY) iff f is injective (WHY) iff $\mathrm{Ker}(f)=\{\overline{0}\}$ (WHY).

Claim. $\mathrm{Ker}(f)=\{\overline{0}\}$ iff m,n are relatively prime.

Proof of Claim. To "⇒": By contradiction, suppose that $d:=\gcd(m,n)>1$, and set $m=m'd$, $n=n'd$ with m',n' relatively prime. Then $k=mn'=nm'=dm'n'<d^2m'n'=mn$, hence $\overline{k}\neq\overline{0}$ in $\mathbb{Z}/(nm)\mathbb{Z}$ (WHY). Further, $f(\overline{k})=(k+m\mathbb{Z},k+n\mathbb{Z})=(mn'+m\mathbb{Z},nm'+n\mathbb{Z})=(m\mathbb{Z},n\mathbb{Z})$, that is, $\overline{k}\in\mathrm{Ker}(f)$. Since $\overline{k}\neq\overline{0}$ and $\mathrm{Ker}(f)=\{\overline{0}\}$, we get a contradiction!

To "⇐": Let $\overline{a}\in\mathrm{Ker}(f)$. Then $a\in m\mathbb{Z}$ and $a\in n\mathbb{Z}$ (WHY), hence $a=mk=nl$ for some $k,l\in\mathbb{Z}$ (WHY). Since since m,n are relatively prime and $mk=nl$, one must have $n|k$ (WHY), hence $k=nk'$ for some $k'\in\mathbb{Z}$. Conclude that $a=mk=mnk'$ in $\mathbb{Z}$, hence $\overline{a}=\overline{0}$ in $Z/(nm)\mathbb{Z}$ (WHY). □

B) **Factor groups**

Let $G,\cdot$ be a group, and $G'\subset G$ be a subgroup.

- The sets $xG':=\{x\cdot g'\,|\,g'\in G'\}$ for $x,y\in G$ satisfy: $xG'\cap yG'\neq\emptyset$ iff $xG'=yG'$ (WHY).
  The sets $\overline{x}:=xG'$ are called the G'-cosets of G, and we denote $G/G':=\{\overline{x}\,|\,x\in G\}$.
- In particular, $(xG')_{x\in G}$ is a partition of G, thus $x\sim_{G'}y\overset{\text{def}}{\longleftrightarrow}xG'=yG'$ is a equivalence relation in G, whose equivalence classes are the cosets $\overline{x}=xG'$ (WHY).
- Similarly, $G'x:=\{g'\cdot x\,|\,g'\in G'\}$ for $x,y\in G$ satisfy: $G'x\cap G'y\neq\emptyset$ iff $G'x=G'y$ (WHY), and the same holds correspondingly for $x\,{}_{G'}\!\sim y\overset{\text{def}}{\longleftrightarrow}G'x=G'y$.

Definition 1.25. A subgroup $\Delta\subset G$ is called normal subgroup, if $x\Delta=\Delta x$ for all $x\in G$.

Remark 1.26. If G is abelian, i.e., $x\cdot y = y\cdot x$ for all $x, y \in G$, then $x\Delta = \Delta x$ for every subgroup $\Delta \subset G$ (WHY). Therefore, *in an abelian group every subgroup is normal.*

The composition law of abelian groups is usually denoted by $+$ and therefore: If $G, +$ is an abelian group, and $\Delta \subset G$ is a subgroup, the Δ-cosets are $x + \Delta = \Delta + x$ for all $x \in G$.

- If $\Delta \subset G$ is normal, define on G/Δ a composition law $*$ by:

$$\overline{x} * \overline{y} := \overline{x\cdot y}$$

Claim. $*$ *is well defined on* G/Δ, i.e., if $\overline{x} = \overline{x}'$, $\overline{y} = \overline{y}'$ in G/Δ, one has:

$$\overline{x\cdot y} = \overline{x'\cdot y'}$$

[**Hint**: One has $\overline{x} = \overline{x}'$, $\overline{y} = \overline{y}'$ iff $\exists\, g, h \in \Delta$ s.t. $x' = x\cdot g$, $y' = y\cdot h$, implying $x'y' = x\cdot g\cdot y\cdot h$. OTOH, since $h\Delta = \Delta h$, $\exists\, g' \in \Delta$ s.t. $g\cdot y = y\cdot g'$ (WHY). Hence setting $h' := g'\cdot h$, one gets $x'\cdot y' = x\cdot g\cdot y\cdot h = x\cdot y\cdot g'\cdot h = (x\cdot y)\cdot h' \in (x\cdot y)\Delta$ (WHY), etc.]

Proposition 1.27. *In the above notations, let* $\Delta \subset G$ *be a normal subgroup. TFH:*

1) *The set of cosets* G/Δ *endowed with* $*$ *is a group with* $e_{G/\Delta} = \overline{e}_G$.
2) *The map* $\varphi : G \to G/\Delta$ *defined by* $\varphi(x) := \overline{x}$ *is a surjective group homomorphism.*

Terminology: *G/Δ is called the* factor group *of G by its (normal) subgroup Δ, and the surjective homomorphism $\varphi : G \to G/\Delta$, $\varphi(x) := \overline{x}$ is the called the canonical projection.*

Proof. To 1): (i) $*$ is associative: $(\overline{x} * \overline{y}) * \overline{z} = \overline{(x\cdot y)} * \overline{z} = \overline{(x\cdot y)\cdot z} = \overline{x\cdot(y\cdot z)} = \overline{x} * \overline{(y\cdot z)} = \overline{x} * (\overline{y} * \overline{z})$ (WHY). (ii) $\overline{e}_G$ is neutral element: $\overline{x} * \overline{e}_G = \overline{x\cdot e_G} = \overline{x} = \overline{e_G\cdot x} = \overline{e}_G * \overline{x}$ (WHY). (iii) Finally, if x' is the inverse of x, then $\overline{x}' * \overline{x} = \overline{x'\cdot x} = \overline{e}_G = \overline{x\cdot x'} = \overline{x} * \overline{x}'$ (WHY), hence $\overline{x}' = \overline{x}^{-1}$ (WHY). Thus $G/\Delta, *$ is a group with $e_{G/\Delta} = \overline{e}_G$.

To 2): By definition one has $\varphi(x\cdot y) = \overline{x\cdot y} = \overline{x} * \overline{y} = \varphi(x) * \varphi(y)$ (WHY). Further, since $\overline{x} = \varphi(x)$, the morphism φ is surjective. □

Proposition 1.28. *Let* $\varphi : G \to H$ *be a group homomorphism. The following hold:*

1) $\mathrm{Ker}(\varphi) := \{x \in G \,|\, f(x) = e_H\} \subset G$ *is normal subgroup, and* $\overline{\varphi} : G/\mathrm{Ker}(f) \to f(G)$ *defined by* $\overline{\varphi}(\overline{x}) :\overset{\text{def}}{=} \varphi(x)$ *is a group isomorphism.*
2) *Moreover, if* $H' \subset \varphi(G)$ *is a subgroup and* $G' := \varphi^{-1}(H')$, *then* $\mathrm{Ker}(\varphi) \subset G'$ *is a normal subgroup, and* $\overline{\varphi}$ *maps* $G'/\mathrm{Ker}(\varphi) \subset G/\mathrm{Ker}(\varphi)$ *isomorphically onto* H'.

Proof. To 1): We first prove that $\mathrm{Ker}(\varphi) \subset G$ is a normal subgroup: Given $x \in G$, let $y \in G$ satisfy $\varphi(x) = \varphi(y)$. Then $\varphi(x^{-1}\cdot y) = \varphi(x)^{-1}\cdot\varphi(y) = e_H$ (WHY), and similarly $\varphi(y\cdot x^{-1}) = \varphi(y)\cdot\varphi(x)^{-1} = e_H$ (WHY). Hence $y\cdot x^{-1}, x^{-1}\cdot y \in \mathrm{Ker}(\varphi)$, and therefore, $y \in x\,\mathrm{Ker}(\varphi)$ and $y \in \mathrm{Ker}(\varphi)x$ (WHY). Conclude that $\varphi(y) = \varphi(x)$ iff $y \in x\,\mathrm{Ker}(\varphi)$ iff $y \in \mathrm{Ker}(\varphi)x$ (WHY). Thus finally $x\,\mathrm{Ker}(\varphi) = \mathrm{Ker}(\varphi)x$ for all$x \in G$. Second, we notice that the map $\overline{\varphi} : G/\mathrm{Ker}(\varphi) \to H$, $\overline{x} \mapsto \varphi(x)$ is well defined and injective. Indeed, $\overline{x} = \overline{x}'$ iff $x' \in x\,\mathrm{Ker}(\varphi)$ (WHY) iff $\varphi(x) = \varphi(x')$ (WHY) $\overset{\text{def}}{\longleftrightarrow}$ $\overline{\varphi}(x) = \overline{\varphi}(x')$. Obviously, $\overline{\varphi}\big(G/\mathrm{Ker}(\varphi)\big) = \varphi(G)$, hence $\overline{\varphi} : G/\mathrm{Ker}(\varphi) \to \varphi(G)$ is a bijective, hence an isomorphism (WHY).

To 2): Since $\mathrm{Ker}(\varphi) \subset G'$ (WHY), it follows that $G'/\mathrm{Ker}(\varphi) \subset G/\mathrm{Ker}(\varphi)$, and $\overline{\varphi}\big(G'/Ker(\varphi)\big) \overset{\text{why}}{=} \varphi(G') \overset{\text{why}}{=} H'$. Hence the restriction of $\overline{\varphi}$ to $G'/\mathrm{Ker}(\varphi)$ maps $G'/\mathrm{Ker}(\varphi) \subset G/\mathrm{Ker}(\varphi)$ isomorphically onto $H' \subset \varphi(G)$. □

Example 1.29.

- $(0,\infty), \cdot$ is a subgroup of the multiplicative group $\mathbb{R}^\times$ and $\mathbb{R}/(0,\infty) = \{\overline{1}, \overline{-1}\}$ (WHY).
- $\mathbb{R}, +$ is a subgroup of $\mathbb{C}, +$ and $\mathbb{C}/\mathbb{R} = \{\,\overline{bi} \,|\, b \in \mathbb{R}\,\}$ (WHY), where $i^2 = -1$.
- $2\pi\mathbb{Z}, +$ is a subgroup of $\mathbb{R}, +$, and $\mathbb{R}/2\pi\mathbb{Z} = \{\,\overline{a} \mid 0 \leqslant a < 2\pi\,\}$ is the group of rotations around the origin O in the xy-plane (WHY).

- $G := \{ax + b \,|\, a, b \in \mathbb{R}, a \neq 0\}$ is a non-abelian group w.r.t the composition of maps $\circ$, and $H := \{t + b \,|\, b \in \mathbb{R}\} \subset G$ is a normal subgroup s.t. $G/H = \{\,\overline{at} \,|\, a \in \mathbb{R}^\times\}$ (WHY).

C) **Factor rings**

Let $R, +, \cdot$ be ring with 1_R. A subset $I \subset R$ is an ideal of R, if it satisfies:

- $x + y \in I$ for all $x, y \in I$.
- $r \cdot x, x \cdot r \in I$ for all $r \in R$, $x \in I$.

Note that in particular, the set of cosets $R/I = \{\overline{x} = x + I \,|\, x \in R\}$ endowed with the addition of cosets $\overline{x} + \overline{y} = \overline{x + y}$ is an abelian group, by the discussion at B) above.

- Define on R/I a *multiplication* by $\overline{x} \cdot \overline{y} :\overset{\text{def}}{=} \overline{x \cdot y}$

Proposition 1.30. *In the above notations, the following hold:*

1) *The multiplication $\cdot$ on R/I is well defined, and $R/I, +, \cdot$ is a commutative ring having the zero element $0_{R/I} = \overline{0}_R$ and the unit $1_{R/I} = \overline{1}_R$.*
2) *The map $\varphi : R \to R/I$ defined by $\varphi(x) := \overline{x}$ is a surjective ring homomorphism.*

Terminology: *R/I is called the factor ring of R by its ideal I, and the surjective ring homomorphism $\varphi : R \to R/I$, $\varphi(x) := \overline{x}$ is called the canonical projection.*

Proof. ... □

Proposition 1.31. *Let $\varphi : R \to S$ be a ring homomorphism. The following hold:*

1) $\operatorname{Ker}(\varphi) := \{x \in R \,|\, \varphi(x) = 0_S\} \subset R$ *is an ideal, and the map $\overline{\varphi} : R/\operatorname{Ker}(\varphi) \to \varphi(R)$ defined by $\overline{\varphi}(\overline{x}) :\overset{\text{def}}{=} \varphi(x)$ is a ring isomorphism.*
2) *Moreover, if $S' \subset \varphi(R)$ is a subring and $R' := \varphi^{-1}(S')$, then $\operatorname{Ker}(\varphi) \subset R'$, and $\overline{\varphi}$ maps $R'/\operatorname{Ker}(\varphi) \subset R/\operatorname{Ker}(\varphi)$ isomorphically onto S'.*

Proof. ... □

Example 1.32.

- If $r \in R$, then $I := rR := \{\,rs \,|\, s \in R\,\} \subset R$ is an ideal.
- The ideals in $R = \mathbb{Z}$ are of the form $I = nR$ with $n \geqslant 0$ (WHY).
- The ideals in $R = \mathbb{Q}[t]$ are of the form $I = p(t)\mathbb{Q}[t]$ with $p(t) \in \mathbb{Q}[t]$ (WHY).
- $R = \mathbb{R}[t]$ the polynomial ring, and $I = (t^2 + 1)\,\mathbb{R}[t]$. Then $R/I = \{\,a + b\overline{t} \mid a, b \in \mathbb{R}\,\}$ can be identified with $\mathbb{C}$ via the $+$ and $\cdot$ compatible map $a + b\overline{t} \mapsto a + bi$ (WHY).
- $R = \mathbb{R}[t]$ the polynomial ring, and $I = (t^2 - 1)\mathbb{R}[t]$. Then $R/I = \{\,a + b\overline{t} \mid a, b \in \mathbb{R}\,\}$ can be identified with $\mathbb{R} \times \mathbb{R}$ via the $+$ and $\cdot$ compatible map $a + b\overline{t} \mapsto (a + b, a - b)$ (WHY).
- Let $\mathcal{C}(I, \mathbb{R})$ be the ring of continuous functions on an open nonempty interval $I := (a, b)$.
  a) Let $X \subset (a, b)$. Then $I_X := \{f \in \mathcal{C}(I, \mathbb{R}) \,|\, f(x) = 0 \,\forall\, x \in X\,\} \subset \mathcal{C}(I, \mathbb{R})$ is an ideal (WHY).
  What is $\mathcal{C}(I, \mathbb{R})/I_X$? What is $\mathcal{C}(I, \mathbb{R})/I_X$ if X is dense in I, resp. $|X| = 1$?
  b) If $f \in \mathcal{C}(I, \mathbb{R})$ is a polynomial function, then f is not a zero divisor (WHY).
  c) $f \in \mathcal{C}(I, \mathbb{R})$ is a zero divisor iff $\exists\, a \leqslant c < d \leqslant b$ having $f(x) = 0$ for $c < x < d$ (WHY).

1.4. **Fields of fractions.**

Let R be an integral domain.

- For $a, s \in R$, $s \neq 0_R$, set $\frac{a}{s} := \{(a', s') \mid a', s' \in R, s' \neq 0_R, as' = a's\}$.
- $\frac{a}{s}$ are precisely the equivalence classes of $(a, s) \sim (a', s') \overset{\text{def}}{\longleftrightarrow} as' = a's$.
- On the set of equivalence classes $F := \{\frac{a}{s} \mid a, s \in R, s \neq 0_R\}$, define:
  - *Addition* $+$ defined by $\frac{a}{s} + \frac{b}{t} :\overset{\text{def}}{=} \frac{at+bs}{st}$
  - *Multiplication* $\cdot$ defined by $\frac{a}{s} \cdot \frac{b}{t} :\overset{\text{def}}{=} \frac{ab}{st}$

Proposition 1.33. *Let R be an integral domain. TFH:*

1) *$+$ and $\cdot$ on F are well defined, and $F, +, \cdot$ is a field with $0_F = \frac{0_R}{1_R}$ and $1_F = \frac{1_R}{1_R}$.*
2) *The map $\imath : R \to F$ by $r \mapsto \frac{r}{1_R}$ is injective ring morphism, and via this map we identify R with the subring $\imath(R) \subset F$ of F.*

Terminology*: $F, +, \cdot$ is called the field of fractions of R. Notation: $K := \mathrm{Quot}(R)$.*

Proof. ... □

Remark 1.34. The field of fractions $F = \mathrm{Quot}(R)$ is the *minimal field* containing R in the following sense: Let $F' \supset R$ be a field containing R as a subring. Then the map $\jmath : F \to F'$ by $\jmath(\frac{a}{s}) :\overset{\text{def}}{=} a \cdot s^{-1}$ is a field morphism (via which one can identify F with a subfield of F').

Example 1.35.

- The field of fractions of $R := \mathbb{Z}$ is the field of rational numbers $F := \mathbb{Q}$.
- The field of fraction of the polynomial ring $R = \mathbb{R}[t]$ is the rational function field $F = \mathbb{R}(t)$.
- The field of fractions of the polynomial ring $R = \mathbb{Z}[t]$ is the rational function field $\mathbb{Q}(t)$.
- In general, if R is a domain with $K = \mathrm{Quot}(R)$, then $R[t]$ has $K(t)$ as a field of fractions.
- If F is a field, then the fraction field of $F[[t]]$ is the so called Laurent power series field
$$F((t)) := \left\{ \sum_{i \geqslant -n} a_i t^i \mid n \in \mathbb{N},\ a_i \in F \right\} = \cup_{n>0} \tfrac{1}{t^n} R[[t]]$$
- If R is a ring, and $F = \mathrm{Quot}(R) \neq R$, then $\mathrm{Quot}\Big(R[[t]]\Big)$ is strictly contain in $F((t))$ (WHY)

Next two items I) and II) are a Supplement

I) The group attached to a commutative monoid

Definition 1.36. Let $X, *$ be a set endowed with a composition law. One says that $*$ has left cancellation, or the left cancellation property, if for all $x, y, z \in X$ one has: $z*x = z*y \Rightarrow x = y$. Define correspondingly the right cancellation, and notice that if $*$ is commutative, then left/right cancellations are equivalent (WHY).

Example 1.37. The following hold:

- $\mathbb{N}, +$ and $\mathbb{N}_{>0}, \cdot$ have cancellation (WHY).
- Which among the composition laws $\cup$, $\cap$, Δ on $X := \mathcal{P}(A)$ have cancellation?
- Is there cancellation in the monoid $\mathrm{Maps}(X), \circ$?

Proposition 1.38. *Let $M,*$ be a commutative monoid. On the set $M\times M$ consider the relation $(a,b)\sim(a',b') \overset{\text{def}}{\longleftrightarrow} \exists\, x\in M$ s.t. $a*b'*x=a'*b*x$. Then the following hold:*

1) *The relation $\sim$ is an equivalence relation on the set $M\times M$. For $(a,b)\in M\times M$, let $\overline{(a,b)}$ be its equivalence class, and set $G :\overset{\text{def}}{=} M\times M/\sim$ be the set of equivalence classes.*
2) *Define on G the composition law: $\overline{(a,b)}*\overline{(c,d)} :\overset{\text{def}}{=} \overline{(a*c,b*d)}$. Then $*$ is well defined, and $G,*$ is a group, with $e_G := \overline{(e_M,e_M)} = \overline{(a,a)}$, and $\overline{(a,b)}^{-1} = \overline{(b,a)}$ for all $a,b\in M$.*
3) *Moreover, suppose that $M,*$ has cancellation. Then $(a,b)\sim(a',b')$ iff $a*b'=a'*b$, and the map $\imath : M\to G$ by $a\mapsto\overline{(a,e)}$ is injective, and satisfies $\imath(a*b)=\imath(a)*\imath(b)$.*

Proof. **Ex** . . . (direct checking) □

Example 1.39. The additive group of integer numbers $\mathbb{Z},+$

Let $M,*$ be $\mathbb{N},+$. Then one has $(k,l)\sim(k',l') \overset{\text{def}}{\longleftrightarrow} k+l'=k'+l$. Therefore, the equivalence relation $\sim$ is the previously defined equivalence relation on $\mathcal{Z}:=\mathbb{N}\times\mathbb{N}$, and the above abstract construction for $M,*$ delivers the additive group $\mathbb{Z},+$ of the integer numbers with the usual addition of such numbers.

Example 1.40. The multiplicative group of positive rational numbers $\mathbb{Q}_{>0}$

Let $M,*$ be $\mathbb{N}_{>0}$ endowed with the multiplication $\cdot$. Then $(n,m)\sim(n',m')$ iff $n\cdot m'=n'\cdot m$ (WHY). In particular, this is the previously equivalence relation on $\mathcal{N}\subset\mathcal{Z}$ used to define the rational numbers. The resulting group attached to $M,*$ is the group of positive rational numbers w.r.t. multiplication (WHY).

II) The ring/field attached to a semiring/semifield

Definition 1.41.

1) A commutative semiring is a set $\mathcal{R}$ endowed with two composition laws: addition $+$ and multiplication $\cdot$ such that $\mathcal{R},+$ and $\mathcal{R},\cdot$ are monoids, and $\cdot$ is distributive w.r.t. $+$. One denotes the neutral elements of $+$ and $\cdot$ by $0_{\mathcal{R}}$, respectively $1_{\mathcal{R}}$, and called them the zero element, respectively the unit element of $\mathcal{R}$.
2) A semifield is a commutative semiring $\mathcal{F},+,\cdot$ such that every $x\in\mathcal{F}$, $x\neq 0_{\mathcal{F}}$ is invertible w.r.t. the multiplication $\cdot$, i.e., $\mathcal{F}^\times := \mathcal{F}\setminus\{0_{\mathcal{F}}\}$ endowed with $\cdot$ is a commutative group.

Example 1.42. On has the following:

a) $\mathbb{N},+,\cdot$ is a commutative semiring (WHY).
b) $\mathbb{Q}_{\geqslant 0},+,\cdot$ is a semifield (WHY).

Proposition 1.43. *Let $\mathcal{R},+,\cdot$ be a commutative semiring such that $+$ has cancellation, and let $R,\oplus$ be the monoid attached to $\mathcal{R},+$. Further, denote the equivalence class $\overline{(a,b)}$ by $(a\text{-}b) :\overset{\text{def}}{=} \overline{(a,b)}$ for $a,b\in\mathcal{R}$, and set $1_R := (1\text{-}0_{\mathcal{R}})$. Define on R a multiplication $\odot$ by the rule:*

$$(a\text{-}b)\odot(c\text{-}d) \;:\overset{\text{def}}{=}\; ((a\cdot c+b\cdot d)\text{-}(a\cdot d+b\cdot c))\,.$$

Then the multiplication $\odot$ is well defined, and the following hold:

1) *Then $R,\oplus,\odot$ is a commutative ring with 1_R as above, and $\imath:\mathcal{R}\to R$ by $\imath(a)=(a\text{-}0_{\mathcal{R}})$ is injective and satisfies: $\imath(a+b)=\imath(a)\oplus\imath(b)$, $\imath(a\cdot b)=\imath(a)\odot\imath(b)$,*
2) *Moreover, if $\mathcal{F},+,\cdot$ is a semifield, then the corresponding $F,\oplus,\odot$ is a field.*

Proof. **Ex** (the proof is virtually identical with the one constructing $\mathbb{Z},+,\cdot$ from the semiring $\mathbb{N},+,\cdot$, etc.) □

Terminology/Convention. In the above context, $R,\oplus,\odot$ and $F,\oplus,\odot$ are called the ring, respectively field, attached to $\mathcal{R}$, respectively $\mathcal{F}$. Via the embedding $\imath:\mathcal{R}\to R$, one identifies $a\in\mathcal{R}$ with $\imath(a)\in R$, thus views $\mathcal{R}$ as a subset of R. One gets identifications:

$$0_{\mathcal{R}}=(0_{\mathcal{R}}\text{-}0_{\mathcal{R}})=0_R,\quad 1_{\mathcal{R}}=(1_{\mathcal{R}}\text{-}0_{\mathcal{R}}),\quad a=\imath(a)=(a\text{-}0_{\mathcal{R}})$$

Notice that under these identifications one has: $\imath(a)-\imath(b)=(a\text{-}b)=a-b$ for all $a,b\in\mathcal{R}$ (WHY).

Example 1.44. On has the following:

a) The ring attached to the semi-ring $\mathbb{N},+,\cdot$ is $\mathbb{Z},+,\cdot$ (WHY).

b) The field attached to the division semi-field $\mathbb{Q}_{\geqslant 0},+,\cdot$ is $\mathbb{Q},+,\cdot$ (WHY).

---

2. Modules and Vector spaces

2.1. Basic definitions/Facts.

Definition 2.1 (Modules/Vector spaces). Let $R,+,\cdot$ be a ring, $M,+$ be an abelian group.

1) An outer multiplication/action of R on M is any map $\psi:R\times M\to M$, $(r,x)\mapsto\psi(r,x)$, usually denoted $\psi(r,x)=r\cdot x$, such that $\forall\, r,s\in R$ and $x,y\in M$ one has:
   (i) $r\cdot(x+y)=r\cdot x+r\cdot y$
   (ii) $(r+s)\cdot x=r\cdot x+s\cdot x$
   (iii) $r\cdot(s\cdot x)=(r\cdot s)\cdot x$
   (iv) $1_R\cdot x=x$
2) An R-module is any abelian group endowed with an outer multiplication of R.
   An F-vector space is a module V over a (skew) field F. The elements $a\in F$ are called scalars, and the elements $v\in V$ are called vectors.
3) An R-submodule of an R-module M is any abelian subgroup $M'\subset M$ of M such that $\forall\, r\in R, x\in M'$ one has: $r\cdot x\in M'$. In particular, M' is itself an R-module (WHY).
   An F-vector subspace of an F-vector space V is any F-submodule V' of V. In particular, V' is itself an F-vector space (WHY).

Note *that virtually everything we say about modules over rings holds in the corresponding form for vector spaces.* **But** there are many *specific facts* facts hold **only** for vector spaces(!)

Example 2.2. Let $G,+$ be an abelian group, $R,+,\cdot$ be a ring, $F,+,\cdot$ be a (skew) field.

- $G,+$ is a $\mathbb{Z},+,\cdot$ module via the outer multiplication $(n,x)\mapsto nx$ (WHY).
  Every subgroup $G'\subset G$ is automatically a $\mathbb{Z}$-submodule (WHY).
- $R,+$ is an $R,+,\cdot$ module via the "outer" multiplication $(r,x)\mapsto r\cdot x$ (WHY).
- $F,+$ is an $F,+,\cdot$ vector space via the "outer" multiplication $(r,x)\mapsto r\cdot x$ (WHY).
- $R[t],+$ is an R-module via the outer multiplication $r\cdot\sum_i a_it^i:\overset{\text{def}}{=}\sum_i r\cdot a_it^i$ (WHY).
  Similarly, $F[t]$ is an F-vector space (WHY).
- $\mathsf{Pol}_n:=\{p(t)\in R[t]\,|\,\deg p(t)\leqslant n\}$ is a subgroup of $R[t],+$ and an R-submodule (WHY).
  Similarly, $\mathsf{Pol}_n\subset F[t]$ is an F-vector subspace of $F[t]$ (WHY).

Proposition 2.3. *Let M be an R-module. TFH:*

1) Computation rules*: For all $r \in R$, $s \in R^\times$, and $x \in M$ one has:*
$$r\cdot 0_M = 0_M = 0_R\cdot x,\ (-1_R)x = -x,\ \text{and } s\cdot x = 0_M \text{ iff } x = 0_M.$$
$(*)$ *In particular, if V is an F-vector space, then $r\cdot x = 0_V$ iff $r = 0_F$ or $x = 0_V$.*

2) *Let $r, r_i, s_i \in R$, $x_i, y_i \in M$ for $1 \leqslant i \leqslant n$. Then $\sum_i r_i\cdot x_i \in M$, and further:*
$$r\cdot(\textstyle\sum_i r\cdot x_i) = \sum_i (r\cdot r_i)\cdot x_j, \qquad (\sum_i r_i\cdot x_i) + (\sum_i s_i\cdot y_i) = \sum_i (r_i\cdot x_i + s_i\cdot y_i).$$

3) *Let $M_1,\dots,M_n \subset M$ be a set of R-submodules. One has:*
   a) *$\cap_i M_i \subset M$ is an R-submodule of M.*
   b) *$\sum_i M_i := M_1 + \dots + M_n \overset{\text{def}}{=} \{\sum_i x_i \,|\, x_i \in M_i, 1 \leqslant i \leqslant n\} \subset M$ is an R-submodule.*

Terminology. *$\sum_i M_i \subset M$ is called the sum of the R-submodules $M_1,\dots,M_n$.*
We say that the sum $\sum_i M_i \subset M$ is direct, if $M_i \cap \sum_{i'\neq i} M_{i'} = \{0_M\}$ for all $i = 1,\dots,n$.

Proof. To 1): Let $x' := 0_R\cdot x$. Then $x' = (0_R + 0_R)\cdot x = x' + x'$ (WHY). Hence adding $-x'$ on both sides, get $0_M = x'$ (WHY). Further, $0_R\cdot x = (1_R + (-1_R))\cdot x = x + (-1_R)\cdot x$, hence $(-1_R)\cdot x = -x$ in M (WHY). Similarly, $r\cdot 0_M = 0_M$ (HOW). Finally, let $s \in R^\times$ and $s' := s^{-1}$, i.e., $s\cdot s' = 1_R$. Then $x \overset{\text{why}}{=} 1_R\cdot x \overset{\text{why}}{=} (s'\cdot s)\cdot x = s'\cdot(r\cdot x)$. Hence $r\cdot x = 0_M$ implies $x = s'\cdot 0_M = 0_M$ (WHY). Conversely, if $r\cdot x \neq 0_M$, then $x \neq 0_M$ (WHY).

To 2): **Ex** (Make induction on n).

To 3): We first prove that $M_0 := \cap_i M_i$ is an R-submodule. Indeed: (i) If $r \in R, x, y \in M_0$, then $0_M, x, y \in M_i$ and $r\cdot x \in M_i$ for all M_i (WHY). Hence $x+y, r\cdot x \in M_i$ for all M_i (WHY). Thus $0_M, x+y, r\cdot x \in M_0$, concluding that M_0 is an R-submodule. Second, we prove that $\sum_i M_i$ is an R-submodule of M. Indeed, let $r \in R$, $x_i, y_i \in M_i$ for $1 \leqslant i \leqslant n$ be be given, and set $x := \sum_i x_i$, $y = \sum_i y_i$. Then by assertion 2) above one has: $r\cdot x = \sum_i r\cdot x_i$, and $x + y = \sum_i (x_i + y_i)$ (WHY). Finally, since $r\cdot x_i, x_i + y_i \in M_i$ (WHY), it follows that $r\cdot x, x + y \in \sum_i M_i$, concluding that $\sum_i M_i$ is an R-submodule of M. □

Example 2.4.

- Let $M_1 = \{(r, 0_R) \,|\, r \in R\}, M_2 = \{(r, r) \,|\, r \in R\} \subset M$. Then M_1, M_2 are R-submodules of $M := R^2$ and $M = M_1 + M_2$ as direct sum (WHY).
- Let $M_1 = \{(r, 0_R, 0_R) \,|\, r \in R\}, M_2 = \{(r, r, s) \,|\, r, s \in R\}, M_3 = \{(0_R, 0_R, r) \,|\, r \in R\}$. Then M_1, M_2, M_3 are R-submodules of $M = R^3$, $M = M_1 + M_2 + M_3$, but not direct (WHY). What about $M_1 + M_2$, $M_1 + M_3$, $M_2 + M_3$?
- Let $\mathsf{Pol}^0, \mathsf{Pol}^1 \subset R[t]$ be the sets of even, respectively odd polynomials. Then $\mathsf{Pol}^0, \mathsf{Pol}^1$ are R-submodules of $R[t]$, and $R[t] = \mathsf{Pol}^0 + \mathsf{Pol}^1$, the sum being direct (WHY).

Remark 2.5. Note that the intersection of any set of submonoids/subgroups/subrings of a given monoid/group/ring is again a substructure of the given structure. OTOH, to define the notion corresponding to the above notion of "sum of R-subodules" is more involved.

The previously introduced constructions/procedures with monoids/groups/rings have corresponding analogs or R-modules and F-vector spaces as follows.

Proposition 2.6. *In the above notation, the following hold:*

1) ***R-Modules of functions**. Let M be an R-module, X a non-empty set. Then the outer multiplication of $r \in R$ on function $f : X \to M$ defined by $(r\cdot f)(x) \overset{\text{def}}{=} r\cdot f(x)$ makes the additive group $\mathrm{Maps}(X, M), +$ into an R-module.*
Correspondingly, if V is an F-vector space, $\mathrm{Maps}(X, F)$ is an F-vector space.

2) *Let $M_1,\dots,M_n$ be R-modules. Then the abelian group $M = M_1\times\cdots\times M_n$ endowed with the outer multiplication $r\cdot(x_1,\dots,x_n) :\overset{\text{def}}{=} (r\cdot x_1,\dots,r\cdot x_n)$ is an R-module, called the (direct) product of the R-modules $M_1,\dots,M_n$.*

Proof. **Ex** . . . □

Example 2.7. Let $R,+,\cdot$ be a ring, $F,+,\cdot$ be a (skew) field.

- $R^n :\overset{\text{def}}{=} \{(x_1,\dots,x_n)\,|\,x_1,\dots,x_n\in R\}$ with the component-wise $+$ is an R module.
- $F^n :\overset{\text{def}}{=} \{(x_1,\dots,x_n)\,|\,x_1,\dots,x_n\in F\}$ with the component-wise $+$ is an F vector space.

2.2. Morphisms of R-modules / R-linear maps.

Recall *that virtually everything we say about modules over rings holds in the corresponding form for vector spaces.* **But** there are many *specific facts* facts hold **only** for vector spaces(!)

Recall the facts about morphisms of groups/rings and factor groups/rings.

Definition 2.8. Let N,M be R-modules. A map $f: N\to M$ is called a morphism of R-modules, or R-module homomorphism, or a R-linear map, if f satisfies:

(i) $f(x+y)=f(x)+f(y)\ \forall\, x,y\in N$, i.e., $f: N\to M$ is a group homomorphism.
(ii) $f(r\cdot x)=r\cdot f(x)\ \forall\, r\in R, x\in N$, i.e., f is compatible with the outer multiplication.

Proposition 2.9 (Morphisms). *Let $f: N\to M$ be an R-linear map of R-modules. Then $f(0_N)=0_M$ and $f(-x)=-f(x)$ for all $x\in N$. Further, the following hold:*

1) $\mathrm{Ker}(f):=\{x\in N\,|\,f(x)=0_M\}$ *is an R-submodule, and* $\mathrm{Ker}(f)=\{0_N\}$ *iff f is injective.*
2) *If $N'\subset N$ and $M'\subset M$ are R-submodules, so are $f(N')\subset M$ and $f^{-1}(M')\subset N$.*
3) *If $f: N\to M$ is bijective, then f is an isomorphism of R-modules.*

Proof. **Ex** (The proof is word-by-word the same as the proof of the assertions for group morphisms . . .) □

Ex 2.10. Let $(M_i)_{1\leqslant i\leqslant n}$ be R-modules of M. Prove that the map $f: M_1\times\ldots\times M_n\to\sum_i M_i$, $(x_1,\dots,x_n)\mapsto\sum_i x_i$ is an R-morphism. Moreover, f is an isomorphism iff $\sum_i M_i$ is direct.

2.3. The R-module $\mathrm{Hom}_R(N,M)$ and the R-algebra $\mathrm{End}_R(M)$.

Recall: Let M be an R-module, and X,T be abstract non-empty sets.

- The set of maps $\mathrm{Maps}(X,M)$ endowed with the usual addition $f+g$ of maps f, g and outer multiplication $r\cdot f$ of maps f by elements $r\in R$ is an R-module (WHY).
- The set of all the maps $\mathrm{Maps}(T) :\overset{\text{def}}{=} \mathrm{Maps}(T,T)$ endowed with the composition $f\circ g$ of maps f, g is a monoid (WHY).

Proposition 2.11. *Let R be a* ***commutative*** *ring, and P,N,M be R-modules. TFH:*

1) $\mathrm{Hom}_R(N,M) := \{f: N\to M\,|\,f$ *is an R-morphism*$\}\subset\mathrm{Maps}(N,M)$ *is an R-submodule.*
2) *The composition of maps* $\mathrm{Hom}_R(P,N)\times\mathrm{Hom}_R(N,M)\to\mathrm{Hom}_R(P,M)$, $(g,f)\mapsto f\circ g$ *is an R-morphism s.t. $\forall\, f,f_1,f_2\in\mathrm{Hom}_R(N,M)$ and $g,g_1,g_2\in\mathrm{Hom}_P(P,N)$ one has:*

$$f\circ(g_1+g_2)=f\circ g_1+f\circ g_2,\qquad (f_1+f_2)\circ g=f_1\circ g+f_2\circ g$$

3) $\mathrm{End}_R(M) := \mathrm{Hom}_R(M,M)$ *endowed with the addition of maps* $+$ *and the composition of maps* $\circ$ *as multiplication is a ring having* $0_{\mathrm{End}} = f_{0_M}$ *the zero map and* $1_{\mathrm{End}} = \mathrm{id}_M$. *Moreover,* $\imath : R \to \mathrm{End}_R(M)$, $r \mapsto r\cdot \mathrm{id}_M$ *is a ring homomorphism which satisfies:*

$$\imath(r)\circ f = f\circ \imath(r)\ \forall\, r\in R, f\in \mathrm{End}_R(M).$$

Proof. To 1): Let $f,g\in \mathrm{Hom}_R(N,M)$, $r\in R$. To show: (a) $f+g\in \mathrm{Hom}_R(N,M)$, (b) $r\cdot f\in \mathrm{Hom}_R(N,M)$.
To (a): (i) $(f+g)(x+y) \overset{\text{def}}{:=} f(x+y)+g(x+y) \overset{\text{why}}{=} f(x)+f(y)+g(x)+g(y) \overset{\text{why}}{=} (f+g)(x)+(f+g)(y)$.
(ii) If $s\in R$, then $(f+g)(s\cdot x) \overset{\text{why}}{=} f(s\cdot x)+g(s\cdot x) \overset{\text{why}}{=} s\cdot f(x)+s\cdot g(x) \overset{\text{why}}{=} s\cdot(f(x)+g(x)) \overset{\text{why}}{=} s\cdot(f+g)(x)$.
To (b): (i) $(r\cdot f)(x+y) \overset{\text{def}}{:=} r\cdot(f(x+y)) = r\cdot(f(x)+f(y)) = r\cdot f(x)+r\cdot f(y) \overset{\text{def}}{:=} (r\cdot f)(x)+(r\cdot f)(y)$.
(ii) If $s\in R$, then $(r\cdot f)(s\cdot x) \overset{\text{def}}{:=} r\cdot f(s\cdot x) \overset{\text{why}}{=} r\cdot(s\cdot f(x)) \overset{\text{why}}{=} (r\cdot s)\cdot f(x) \overset{\text{why}}{=} s\cdot(r\cdot f(x)) \overset{\text{why}}{=} s\cdot((r\cdot f)(x))$.
To 2): We first show that $f\circ g$ is an R-module homomorphism:
(i) $(f\circ g)(x+y) \overset{\text{def}}{:=} f(g(x+y)) \overset{\text{why}}{=} f(g(x)+g(y)) \overset{\text{why}}{=} f(g(x)+f(g(y)) \overset{\text{def}}{:=} (f\circ g)(x)+(f\circ g)(y)$.
(ii) For $s\in R$ one has: $(f\circ g)(s\cdot x) \overset{\text{def}}{:=} f(g(s\cdot x)) \overset{\text{why}}{=} f(s\cdot g(x)) \overset{\text{why}}{=} s\cdot(f(g(x)) \overset{\text{def}}{:=} s\cdot((f\circ g)(x))$.
Second we prove the distributivity of $\circ$ w.r.t. $+$
$(f\circ(g_1+g_2))(x) \overset{\text{def}}{:=} f((g_1+g_2)(x)) \overset{\text{why}}{=} f(g_1(x))+f(g_2(x)) \overset{\text{def}}{:=} (f\circ g_1)(x)+(f\circ g_2)(x)\ \forall\, x\in M$.
$((f_1+f_2)\circ g)(x) \overset{\text{def}}{:=} (f_1+f_2)(g(x)) \overset{\text{def}}{:=} f_1(g(x))+f_2(g(x)) \overset{\text{def}}{:=} (f_1\circ g)(x)+(f_2\circ g)(x)\ \forall\, x\in M$.
To 3): **Ex** (apply assertions 1) and 2) above, etc.) □

Remark 2.12. In the notation form Proposition above, one has:
1) $\mathrm{End}_R(M)$ and $\mathrm{End}_R(N)$ are (usually non-commutative) rings with $0_{\mathrm{End}_R(\bullet)} \neq 1_{\mathrm{End}_R(\bullet)}$.
2) With $\phi\in \mathrm{Hom}_R(N,M)$, $f\in \mathrm{End}_R(M)$, $g\in \mathrm{End}_R(N)$, the composition of maps defines:
   a) A ***left*** *outer multiplication* of the R-algebra $\mathrm{End}_R(M)$ on the R-module $\mathrm{Hom}_R(N,M)$ by $f\bullet\phi \overset{\text{def}}{:=} f\circ\phi$, which makes $\mathrm{Hom}_R(N,M)$ into a (left) $\mathrm{End}_R(M)$-module (WHY).
   b) A ***right*** *outer multiplication* of the R-algebra $\mathrm{End}_R(N)$ on the R-module $\mathrm{Hom}_R(N,M)$ by $\phi\bullet g \overset{\text{def}}{:=} \phi\circ g$, which makes $\mathrm{Hom}_R(N,M)$ into a (right) $\mathrm{End}_R(N)$-module (WHY).

We will see later that these outer multiplications have a concrete realization as left / right actions a rings of $m\times m$ / $n\times n$ matrices on $m\times n$ matrices.

2.4. **Factor R-modules.**

Let M be an R-module, and $M'\subset M$ be an R-submodule. In particular, the factor group $\overline{M} = M/M'$ of the abelian group M by $M'\subset M$ is an abelian group.

Define on $\overline{M} \overset{\text{def}}{:=} M/M'$ an outer multiplication of R by the rule $r\bullet\overline{x} \overset{\text{def}}{:=} \overline{r\bullet x}$.

Claim. *The outer multiplication* $\bullet$ *of* R *on* $\overline{M}$ *is well defined,* i.e., $\forall\, x\in M, r\in R$ one has that if $\overline{x}=\overline{x}'$ in $\overline{M}$, then $\overline{r\bullet x} = \overline{r\bullet x}'$ in $\overline{M}$.

Proof. One has: $\overline{x}=\overline{x}'$ iff $x-x'\in M' \Rightarrow r\bullet(x-x')\in M'$ (WHY) iff $r\bullet x - r\bullet x'\in M'$ iff $\overline{r\bullet x} = \overline{r\bullet x}'$ (WHY).] □

Proposition 2.13. *In the above notations, let* $M'\subset M$ *be an* R*-submodule. TFH:*
1) *The outer multiplication* $\bullet$ *of* R *on* $\overline{M} = M/M'$ *makes* $\overline{M}$ *into an* R*-module.*
2) *The map* $\varphi : M\to\overline{M}$ *defined by* $\varphi(x) := \overline{x}$ *is a surjective morphism of* R*-modules.*

Terminology: *The* R*-module* $\overline{M} = M/M'$ *is the* factor, *or* quotient *of* M *by its* R*-submodule* M', *and the* R*-morphism* $\varphi : M\to M/M'$, $\varphi(x) := \overline{x}$ *is the called the canonical projection.*

Proof. **Ex** (The proof is word-by-word the same as the proof of the assertions for factor groups . . .) □

Proposition 2.14. *Let $\varphi : N \to M$ be an R-linear map of R-modules. The following hold:*

1) *The map $\overline{\varphi} : N/\operatorname{Ker}(\varphi) \to M$ define by $\overline{\varphi}(\overline{x}) = \varphi(x)$ is an isomorphism of R-modules.*
2) *Moreover, for every R-submodule $M' \subset f(N)$, letting $N' := \varphi^{-1} \subset N$, one has: The map $\overline{\varphi} : N/\operatorname{Ker}(\varphi) \to M$ maps $N'/\operatorname{Ker}(\varphi)$ isomorphically onto M'.*

Proof. **Ex** (The proof is word-by-word the same as in the case of group homomorphisms). □

Example 2.15.

Let $X' \subset X$ be a non-empty subset, $R, +$ and $R^m, +$ be the "usual" R-modules.

- Define φ : $\operatorname{Maps}(X, R) \to \operatorname{Maps}(X', R)$, $f \mapsto f' := f|_{X'}$. Then φ is a morphism of R-modules (WHY). What are $\operatorname{Ker}(\varphi), \operatorname{Im}(\psi)$?
- Define $\phi : \operatorname{Maps}(X', R) \to \operatorname{Maps}(X, R)$, $f' \mapsto f$ by $f(x) = f'(x)$ if $x \in X'$ and $f(x) = 0_R$ for $x \notin X'$. Then φ is a morphism of R-modules (WHY). What are $\operatorname{Ker}(\varphi), \operatorname{Im}(\psi)$?
- Let $pr_i : R^m \to R$ be the *projection on the i^{th} coordinate,* i.e., $pr_i(x_1,\dots,x_m) = x_i$. Then pr_i is a morphism of R-modules (WHY). What is $\operatorname{Ker}(pr_i)$?
- Let $p = p(t) \in R[t]$ be fixed. Define $f_p : R[t] \to R$ by $p(t) \mapsto p(0_R)$. Then f_p is a morphism of R-modules (WHY). What is $\operatorname{Ker}(f_p)$?
- Suppose that $|X| = m$. Are he R-modules $\operatorname{Maps}(X, R)$ and R^m isomorphic? Is the same true about $\operatorname{Maps}(X, R)$ and $\operatorname{Maps}(X', R)$ iff $|X| = |X'|$?

2.5. Linear combinations / Span / Linear independence / Bases.

Definition 2.16. Given an R-module M, $X \subset M$ non-empty subset, define:

1) Linear combinations: $\sum_i r_i \cdot x_i = r_1x_1 + \cdots + r_nx_n$, $n \geqslant 1$, $r_1,\dots,r_n \in R$, $x_1,\dots,x_n \in M$.
   (i) If $x = \sum_i r_i \cdot x_i$, we say that $\sum_i r_i \cdot x_i$ represents x.
   (ii) $\sum_i r_i \cdot x_i = r_1x_1 + \cdots + r_nx_n$ is called trivial, if $r_i = 0_R \ \forall\, i$. If so, $\sum_i r_i \cdot x_i = 0_M$ (WHY).
2) The span of X is $\langle X \rangle_R := \{r_1x_1 + \cdots + r_nx_n \,|\, n \geqslant 1, r_1,\dots,r_n \in R,\ x_1,\dots,x_n \in M\}$.
3) We say that X generates / spans the R-module M if $M = \langle X \rangle_R$ equals the span of X.

Definition/Remark 2.17. Recall that a system $(x_i)_{i\in I}$ of elements of a set X is nothing but a short notation for a map $f : I \to X$, $i \mapsto f(i) =: x_i$, where $I \neq \emptyset$. A subsystem of a given system $(x_i)_{i\in I}$ is any $(x_i)_{i\in J}$ for $J \subset I$ non-empty. In particular, if $(x_i)_{i\in I}$ is defined by $f : I \to X$, then $(x_i)_{i\in J}$ defined by the restriction $f|_J : J \to X$ of f to J.

- **Note** that it is **essential to not confuse** $(x_i)_{i\in I}$ with $\{x_i \,|\, i \in I\} = f(I) \subset X$.
- **Finally**, one (often) writes simply $(x_i)_i$ for $(x_i)_{i\in I}$ when that does not lead to confusions.

Definition 2.18. Let $\mathcal{X} = (x_i)_{i\in I}$ and $(x_1,\dots,x_n)$ be systems of elements of an R-module M.

1) $(x,\dots,x_n)$ is called free, or linearly independent (over R), if the following holds:
   $$\forall\, r_1,\dots,r_n \in R \text{ one has: } \sum_i r_i \cdot x_i = 0_M \Rightarrow r_i = 0_R \ \forall\, i.$$
   $\mathcal{X} = (x_i)_i$ is called free, if every *finite subsystem* $(x_{i_1},\dots,x_{i_n})$, $n \geqslant 1$ of $\mathcal{X} = (x_i)_i$ is free.
2) We say that $\mathcal{X} := (x_i)_{i\in I}$ is a basis of M, if $(x_i)_i$ is free and generates the R-module M. We say that M is a free R-module, it M has a basis.

Example 2.19.

- $R,+$ is a free module with basis $\mathcal{A}=(1_R)$ (WHY).
  Similarly, $F,+$ is a free F-vector space with basis $\mathcal{A}=(1_F)$.
- $\mathcal{E}:=(e_1,e_2)$ with $e_1=(1_R,0_R), e_2=(0_R,1_R)$ is a basis of $R^2,+$ (WHY), the standard basis.
  What about the R-module R^m for arbitrary $m\in\mathbb{N}_{>0}$?
- Let $\delta_x\in \mathrm{Maps}(X,R)$ be defined by $\delta_x(x')=1_R$ if $x=x'$ and $\delta_x(x')=0_R$ otherwise. Then $\mathcal{D}:=(\delta_x)_{x\in X}$ is free (WHY). Is $\mathcal{D}$ a basis of $\mathrm{Maps}(X,R)$?
- The set $\mathsf{Pol}_n\subset R[t]$ of polynomials of degree $\leqslant n$ is an R-submodule of $R[t],+$ have as basis $\mathcal{T}=(t^i)_{0\leqslant i\leqslant n}$ (WHY). What about the set of polynomials of odd degree in $R[t]$?

Proposition 2.20. *Let M be an R-module. TFH:*

1) *Let $M_1,\dots,M_n\subset M$ be submodules. Then setting $X=M_1\cup\dots\cup M_n$, one has:*
$$\langle X\rangle_R=\textstyle\sum_i M_i:=\{x_1+\dots+x_n\,|\,x_i\in M_i, 1\leqslant i\leqslant n\}$$
2) *Let $X\subset M$ be a non-empty subset. The span $\langle X\rangle_R\subset M$ is the **smallest** R-submodule of M containing X, i.e., if $N\subset M$ is a submodule with $X\subset N$, then $\langle X\rangle_R\subset N$.*

Proof. To 1): We prove that $\langle X\rangle_R\subset\sum_i M_i$ and $\langle X\rangle_R\supset\sum_i M_i$.

To "$\supset$": Let $x:=x_1+\dots+x_n\in\sum_i M_i$ with $x_i\in M_i$ be given. Then $x_1,\dots,x_n\in X$ (WHY), and x is the linear combination of $x_1,\dots,x_n$ with coefficients $r_1=\dots=r_n=1_R$ (WHY). Hence $x\in\langle X\rangle_R$.

To "$\subset$": Prove by induction on $m\in\mathbb{N}_{m>0}$ that every linear combination $y:=\sum_{j=1}^m r_jy_j\in\langle X\rangle_R$ lies in $\sum_i M_i$. Indeed, if $m=1$, one has: Since $y_1\in X=M_1\cup\dots\cup M_n$, one has $y_1\in M_k$ for some $k\leqslant n$, hence $y=r_1{\cdot}y_1\in M_k\subset\sum_i M_i$ (WHY). The induction step: For $y=r_1{\cdot}y_1+\dots+r_m{\cdot}y_m+r_{m+1}{\cdot}y_{m+1}\in\langle X\rangle_R$, one has: By the induction hypothesis, $y':=r_1{\cdot}y_1+\dots+r_m{\cdot}y_m\in\sum_i M_i$. Hence if $y_{m+1}\in M_k$, then $y'':=r_{m+1}{\cdot}y_{m+1}\in M_k$ (WHY), thus finally one has $y=y'+y''\in\sum_i M_i+M_k=\sum_i M_i$ (WHY).

To 2): First, if $X\subset N$, then $\langle X\rangle_R\subset N$ (WHY). Hence if M_α, $\alpha\in I$ is the set of all the R-submodules of M with $X\subset M_\alpha$, and $M_0:=\cap_\alpha M_\alpha$, one has: First, $X\subset M_0$ (WHY), hence $\langle X\rangle_R\subset M_0$ (WHY). On the other hand, since $X\subset\langle X\rangle_R$, and $\langle X\rangle_R$ is an R-submodule, it follows that $\langle X\rangle_R$ is one of the R-submodules M_α. Hence $M_0\subset\langle X\rangle_R$. Conclude that $M_0=\langle X\rangle_R$ (WHY). □

Proposition 2.21. *Let M be an R-module. Then the following hold:*

1) *If $\mathcal{X}=(x_i)_{i\in I}$, $x_i\in M$, has $x_i=x_j$ or $x_i=0_M$ for some $i\neq j$, then $\mathcal{X}$ is not free.*
2) *Let $\mathcal{X}=(x_i)_{i\in I}$ be free, $I=I_1\cup I_2$ with I_1,I_2 disjoint and non-empty and denote/consider its subsystems $\mathcal{X}_1=(x_i)_{i\in I_1}$, $\mathcal{X}_2=(x_i)_{i\in I_2}$, i.e., $\mathcal{X}=(\mathcal{X}_1,\mathcal{X}_2)$. Then one has:*
$$\langle\mathcal{X}\rangle_R=\langle\mathcal{X}_1\rangle_R+\langle\mathcal{X}_2\rangle_R,\qquad \langle\mathcal{X}_1\rangle_R\cap\langle\mathcal{X}_2\rangle_R=\{0_M\}.$$
In particular, $\langle\mathcal{X}\rangle_R$ is the direct sum of $\langle\mathcal{X}_1\rangle_R$ and $\langle\mathcal{X}_2\rangle_R$ in M.

Proof. To 1): **Ex**... To 2): Step 1: $\langle\mathcal{X}\rangle_R=\langle\mathcal{X}_1\rangle_R+\langle\mathcal{X}_2\rangle_R$.
To "$\subset$": If $x:=r_1x_{i_1}+\dots+r_nx_{i_n}\in\langle\mathcal{X}\rangle_R$, then $y_1:=\sum_{i_k\in I_1}r_k{\cdot}x_{i_k}\in\langle\mathcal{X}_1\rangle_R$, $y_2:=\sum_{i_k\in I_2}r_k{\cdot}x_{i_k}\in\langle\mathcal{X}_2\rangle_R$ (WHY), and further $x=y_1+y_2\in\langle\mathcal{X}_1\rangle_R+\langle\mathcal{X}_2\rangle_R$ (WHY). Hence $\langle\mathcal{X}\rangle_R\subset\langle\mathcal{X}_1\rangle_R+\langle\mathcal{X}_2\rangle_R$.
To "$\supset$": Since $\mathcal{X}_1$ is a subsystem of $\mathcal{X}$, every linear combination of elements $x_{i_1},\dots,x_{i_n}$ with $i_1,\dots,i_n\in I_1$ lies in $\langle\mathcal{X}_1\rangle_R$ (WHY), hence $\langle\mathcal{X}_1\rangle_R\subset\langle\mathcal{X}\rangle_R$ (WHY). Similarly, $\langle\mathcal{X}_2\rangle_R\subset\langle\mathcal{X}\rangle_R$. Hence finally $\langle\mathcal{X}_1\rangle_R+\langle\mathcal{X}_2\rangle_R\subset\langle\mathcal{X}\rangle_R$.

Step 2: $\langle\mathcal{X}_1\rangle_R\cap\langle\mathcal{X}_2\rangle_R=\{0_M\}$. Let namely $x=\sum_k r_k{\cdot}x_{i_k}=\sum_l s_l{\cdot}x_{i_l}$ with $i_k\in I_1$, $i_l\in I_2$ lie in $\langle\mathcal{X}_1\rangle_R\cap\langle\mathcal{X}_2\rangle_R$. Then $0_M=x-x=\sum_k r_k{\cdot}x_{i_k}-\sum_l s_l{\cdot}x_{i_l}$. On the other hand, $\mathcal{Y}:=\big((x_{i_k})_k,(x_{i_l})_l\big)$ is a (finite) subsystem of $\mathcal{X}$ (WHY), hence $\mathcal{Y}$ is free (WHY). Hence $r_k=0_R=s_l$ for all k,l (WHY), thus $x=0_M$. □

Proposition 2.22. *Let $f:N\to M$ be a morphism of R-modules. The following hold:*

1) *Let $X \subset N$ be a non-empty set, hence $f(X) \subset M$. Then $f\big(\langle X\rangle_R\big) = \langle f(X)\rangle_R$.*
2) *Let $\mathcal{Y} = (y_j)_j$ be a system of linearly independent elements in $f(N)$. Then any preimage $\mathcal{X}_1 = (x_j)_j$ of $\mathcal{Y}$ under f is linearly independent inside N, and $f\big(\langle \mathcal{X}_1\rangle_R\big) = \langle \mathcal{Y}\rangle_R$.*
3) *Let $\mathcal{X}_0 = (x_i)_i$ be linearly independent in $\mathrm{Ker}(f)$ and $\mathcal{X}_1 = (x_j)_j$ be as above. Then $\mathcal{X} = \big((x_i)_i, (x_j)_j\big)$ is linearly independent in N, and therefore one has:*

$$\langle \mathcal{X}\rangle_R = \langle \mathcal{X}_0\rangle_R + \langle \mathcal{X}_1\rangle_R, \qquad \langle \mathcal{X}_0\rangle_R \cap \langle \mathcal{X}_1\rangle_R = \{0_N\}.$$

$(*)$ *In particular, if $\mathrm{Ker}(f)$ and $f(N)$ are free R-modules, so is N, and $N = \mathrm{Ker}(f) \oplus \mathrm{Im}(f)$.*

Proof. To 1): **Ex** (if $x = \sum_{i=1}^n r_i \cdot x_i \in \langle X\rangle_R$, then $f(x) = f(\sum_i r_i \cdot x_i) \overset{\text{why}}{=} \sum_i r_i \cdot f(x)i) \in \langle f(X)\rangle_R$, etc.

To 2): Let $x = \sum_{k=1}^n r_k \cdot x_{i_k} \in \langle \mathcal{X}_1\rangle_R$ be given. Then $f(x) \overset{\text{why}}{=} \sum_k r_k \cdot f(x_{i_k}) \overset{\text{why}}{=} \sum_k r_k \cdot y_{i_k}$. Hence if $x = 0_N$, then $0_M \overset{\text{why}}{=} f(x) = \sum_k r_k \cdot y_{i_k}$, thus $r_k = 0_R \ \forall k$ (by the fact that $\mathcal{Y}$ is free). Conclude $\mathcal{X}_1$ is free (WHY).

To 3): Let $(x_{i_k})_k$, $(x_{j_l})_l$ be finite subsystems of $\mathcal{X}_0$, $\mathcal{X}_1$, $r_k, s_l \in R$, say with $1 \leqslant k \leqslant m$, $1 \leqslant l \leqslant n$. Setting $x_0 := \sum_k r_k \cdot x_{i_k}$, $x_1 := \sum_l s_l \cdot x_{j_l}$, one has $x_0 \in \langle \mathcal{X}_0\rangle_R \subset \mathrm{Ker}(f)$ (WHY), $x_1 \in \langle \mathcal{X}_1\rangle_R$ (WHY), and $x = x_0 + x_1 \in \langle \mathcal{X}\rangle_R$ (WHY). Moreover, every $x \in \langle \mathcal{X}\rangle_R$ is of the form $x = x_0 + x_1$ for properly chosen $m, n, r_k, s_l, x_{i_k}, x_{j_l}$ (WHY). Next suppose that $x = x_0 + x_1 = 0_N$. Since $x_0 \in \mathrm{Ker}(f)$, one has $0_M = f(x) = f(x_0) + f(x_1) = f(x_1) = \sum_l s_l \cdot y_{j_l}$ (WHY), and therefore, $s_l = 0_R$ (by the fact that $\mathcal{Y}$ is free). Hence $0_N = x = x_1 = \sum_k r_k \cdot x_{i_k}$, thus $r_k = 0_R$ (by the fact that $\mathcal{X}_0$ is free). Thus finally, $r_k, s_l = 0_R \ \forall k, l$. Conclude that $\mathcal{X}$ is free (WHY). □

Theorem 2.23 (Rank / Dimension). *The following hold:*

1) *All the bases of a finitely generated free R-module M have same cardinality, which we call the rank of M, denoted $\mathrm{rk}_R(M)$.*
2) *Every F-vector space V is free, and all its bases have the same cardinality, called the dimension of V, denoted $\dim_F(V)$.*

Proof. Later, see Section 4, Thm 4.20, etc. . . . □

3. Matrices and Morphisms / Elementary matrices

3.1. Definitions / Basic facts.

For integers $m, n > 0$, set $I := \{1, \dots, m\}$, $J := \{1, \dots, n\}$. For any set X, define the $m \times n$ matrices with coefficients in X as being the $m \times n$ tables of elements of X, as follows:

$$X^{m\times n} = \{(a_{ij})_{i,j} \,|\, 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n,\ a_{ij} \in X\} = \{\varphi : I \times J \to X \mid \varphi \text{ map}\}.$$

Notations: One usually sets $X^n := X^{1\times n}$ and ${}^mX := X^{m\times 1}$.
Given $A = (a_{ij})_{i,j} \in X^{m\times n}$, one defines

- The rows $\boldsymbol{\mathcal{R}}_1, \dots, \boldsymbol{\mathcal{R}}_m \in X^{1\times n} = X^n$ of A, where $\boldsymbol{\mathcal{R}}_i := (a_{ij})_j$
- The columns $\boldsymbol{\mathcal{C}}_1, \dots, \boldsymbol{\mathcal{C}}_n \in X^{m\times 1} = {}^mX$ of A, where $\boldsymbol{\mathcal{C}}_j := (a_{ij})_i$.
- $A^\tau := (a_{ji})_{j,i} \in X^{n\times m}$ is called the transpose (matrix) of A.
  Hence the rows of A are the columns of A^τ, and vice-versa.

$(*)$ In particular: If $\boldsymbol{\mathcal{R}} \in X^n$ iff $\boldsymbol{\mathcal{R}}^\tau \in {}^nX$ and $\boldsymbol{\mathcal{C}} \in {}^nX$ iff $\boldsymbol{\mathcal{C}}^\tau \in X^n$ (WHY).

- Therefore, one has:

$$A = \big(\boldsymbol{\mathcal{C}}_1, \dots, \boldsymbol{\mathcal{C}}_n\big) = \begin{pmatrix} \boldsymbol{\mathcal{R}}_1 \\ \vdots \\ \boldsymbol{\mathcal{R}}_m \end{pmatrix}, \quad A^\tau = \big(\boldsymbol{\mathcal{R}}_1^\tau, \dots, \boldsymbol{\mathcal{R}}_m^\tau\big) = \begin{pmatrix} \boldsymbol{\mathcal{C}}_1^\tau \\ \vdots \\ \boldsymbol{\mathcal{C}}_n^\tau \end{pmatrix}, \quad \text{hence } (A^\tau)^\tau = A \text{ (WHY)}.$$

Definition 3.1 (Kronecker symbol). Let $R, +, \cdot$ be a ring with $1_R \neq 0_R$, and X an arbitrary non-empty set. For $x, y \in X$ we define Kronecker symbol δ_{xy} with values in R by the rule:

$$\delta_{xy} = 1_R \text{ if } x = y \text{ and } \delta_{xy} = 0_R \text{ for } x \neq y.$$

Note that $\delta_{xy}\delta_{uv} = 1_R$ iff $(x, u) = (y, v)$, respectively $\delta_{xy}\delta_{uv} = 0_R$ iff $(x, u) \neq (y, v)$ (WHY).

Proposition 3.2. *Let R be a **commutative** ring with $1_R \neq 0_R$, and M be an R-module. TFH:*

1) *Both $R^{m\times n} = \mathrm{Maps}(I \times J, R)$ and $M^{m\times n} = \mathrm{Maps}(I \times J, M)$ endowed with the usual addition of maps are R-modules.*
   - *Precisely, if $A = (a_{ij})_{i,j}$, $B = (b_{ij})_{i,j}$, then $A + B = (a_{ij} + b_{ij})_{i,j}$, $r \cdot A = (ra_{ij})_{i,j}$*
2) *Moreover, $R^{m\times n}$ is a free R-module having as standard basis*

$$\mathcal{E} := (\boldsymbol{e}_{kl})_{k,l}, \quad k \in I, \ j \in J, \quad \text{where} \quad \boldsymbol{e}_{kl} :\overset{\text{def}}{=} (a_{ij})_{i,j} \quad \text{with} \quad a_{ij} :\overset{\text{def}}{=} \delta_{ki}\delta_{lj}\,.$$

3) *For $X = R, M$, the transpose map $X^{m\times n} \to X^{n\times m}$ is an isomorphisms of R-modules.*

Proof. ... □

Multiplication of matrices:

Let X denote either the ring R, or an R-module M.

I) For every $\boldsymbol{\mathcal{R}} = (x_j)_j \in X^n$ and $\boldsymbol{\mathcal{C}} = (y_i)_i \in {}^nR$, or $\boldsymbol{\mathcal{R}} \in R^n$ and $\boldsymbol{\mathcal{C}} \in {}^nX$, one defines the row-column multiplication $\boldsymbol{\mathcal{R}} \cdot \boldsymbol{\mathcal{C}}$ by the recipe:

$$x := \boldsymbol{\mathcal{R}} \cdot \boldsymbol{\mathcal{C}} = (x_1, \ldots, x_n)\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1y_1 + \cdots + x_ny_n \in X$$

The row-column multiplication $\boldsymbol{\mathcal{R}} \cdot \boldsymbol{\mathcal{C}}$ defined above satisfies:

- *Distributivity w.r.t. addition* in both variables, i.e.:

$$(\boldsymbol{\mathcal{R}} + \boldsymbol{\mathcal{R}}') \cdot \boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{R}} \cdot \boldsymbol{\mathcal{C}} + \boldsymbol{\mathcal{R}}' \cdot \boldsymbol{\mathcal{C}}, \quad \boldsymbol{\mathcal{R}} \cdot (\boldsymbol{\mathcal{C}} + \boldsymbol{\mathcal{C}}') = \boldsymbol{\mathcal{R}} \cdot \boldsymbol{\mathcal{C}} + \boldsymbol{\mathcal{R}} \cdot \boldsymbol{\mathcal{C}}'$$

- *Compatibility w.r.t. outer (scalar) multiplication* by elements $r \in R$, i.e.:

$$(r \cdot \boldsymbol{\mathcal{R}}) \cdot \boldsymbol{\mathcal{C}} = r \cdot (\boldsymbol{\mathcal{R}} \cdot \boldsymbol{\mathcal{C}}) = \boldsymbol{\mathcal{R}} \cdot (r \cdot \boldsymbol{\mathcal{C}})$$

- *Anticommutative w.r.t. transpose* as follows:

$$\boldsymbol{\mathcal{R}} \cdot \boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{C}}^\tau \cdot \boldsymbol{\mathcal{R}}^\tau$$

II) Let $A = (a_{ij})_{i,j} = (\boldsymbol{\mathcal{R}}_i)_i \in X^{m\times n}$, $B = (b_{jk})_{j,k} = (\boldsymbol{\mathcal{C}}_k)_k \in R^{n\times p}$, or $A = (a_{ij})_{i,j} = (\boldsymbol{\mathcal{R}}_i)_i \in R^{m\times n}$, $B = (b_{jk})_{j,k} = (\boldsymbol{\mathcal{C}}_k)_k \in X^{n\times p}$. One defines the matrix multiplication

$$A \cdot B := \big(\boldsymbol{\mathcal{R}}_i \cdot \boldsymbol{\mathcal{C}}_k\big)_{i,k} \in X^{m\times p}$$

The properties of the above row-column multiplication $\boldsymbol{\mathcal{R}} \cdot \boldsymbol{\mathcal{C}}$ imply that the matrix multiplication has the following properties:

- *Distributivity w.r.t. addition* in both variables, i.e.:

$$(A + A') \cdot B = A \cdot B + A' \cdot B, \quad A \cdot (B + B') = A \cdot B + A \cdot B'$$

- *Compatibility w.r.t. outer (scalar) multiplication* by elements $r \in R$, i.e.:

$$(r \cdot A) \cdot B = r \cdot (A \cdot B) = A \cdot (r \cdot B)$$

- *Associativity of multiplication,* i.e., $(A \cdot B) \cdot C = A \cdot (B \cdot C)$, when defined.
- *Anticommutative w.r.t. transpose,* i.e., $(A \cdot B)^\tau = B^\tau \cdot A^\tau$

Proposition 3.3. *Let R be a **commutative** ring with $0_R \neq 1_R$, and M be an R-module. TFH:*

1) *$R^{m\times m}$ endowed with $+$ and $\cdot$ of matrices is a ring, non-commutative if $m > 0$, having $0_{R^{m\times m}} := 0_{m\times m}$ the zero matrix, and $1_{R^{m\times m}} = (\delta_{ij})_{i,j} =: \boldsymbol{I}_m$ the unit matrix.*
2) *Let X denote either R or M. Then $X^{m\times n}$ is a left module w.r.t. the left multiplication by $R^{m\times m}$, and a right module w.r.t the right multiplication by $R^{n\times n}$.*

Proof. ... □

3.2. **Morphisms and Matrices.**

Coordinate vectors: Let M be finitely generated free R-module, $\mathcal{A} = (\alpha_1, \dots, \alpha_m)$ be a basis.
- For $x \in M$ there exist a unique $m \times 1$ matrix $[x]_\mathcal{A} \in {}^mR$ such that $x = \mathcal{A} \cdot [x]_\mathcal{A}$ (WHY).
- One has: $[x + y]_\mathcal{A} = [x]_\mathcal{A} + [y]_\mathcal{A}$, and $r \cdot [x]_\mathcal{A} = [r \cdot x]_\mathcal{A}$ for $r \in R$ (WHY).
- The map $[\ ]_\mathcal{A} : M \to {}^mR$ defined by $x \mapsto [x]_\mathcal{A}$, is an isomorphism of R-modules (WHY).

Definition 3.4. In the above notations, $[x]_\mathcal{A}$ is called the coordinate (vector) of x in the basis $\mathcal{A}$. Further, $[\ ]_\mathcal{A} : M \to {}^mR$ is called the coordinate isomorphism in basis $\mathcal{A}$

Let N, M be finitely generated free R-modules, with bases $\mathcal{B} = (\beta_1, \dots, \beta_n)$, $\mathcal{A} = (\alpha_1, \dots, \alpha_m)$. Let $f : N \to M$ be an R-module. Then for every $\beta \in N$ one has $f(\beta) = \mathcal{A} \cdot [f(\beta)]_\mathcal{A}$. Hence setting $\boldsymbol{c}_j := [f(\beta_j)]_\mathcal{A} \in {}^mR$, $j = 1, \dots, n$, one gets: The matrix $[f]_{\mathcal{A}\mathcal{B}} := (\boldsymbol{c}_j)_j \in R^{m\times n}$ is the unique $m \times n$ matrix over R satisfying:

$$f(\mathcal{B}) := \big(f(\beta_1), \dots, f(\beta_n)\big) = \mathcal{A} \cdot \big(\boldsymbol{c}_1, \dots, \boldsymbol{c}_n\big) = \mathcal{A} \cdot [f]_{\mathcal{A}\mathcal{B}}$$

In particular, for $y = \mathcal{B} \cdot [y]_\mathcal{B} \in N$ one has:

$$\mathcal{A} \cdot [f(y)]_\mathcal{A} = f(y) = f\big(\mathcal{B} \cdot [y]_\mathcal{B}\big) \overset{\text{why}}{=} f(\mathcal{B}) \cdot [y]_\mathcal{B} = \big(\mathcal{A} \cdot [f]_{\mathcal{A}\mathcal{B}}\big) \cdot [y]_\mathcal{B} = \mathcal{A} \cdot \big([f]_{\mathcal{A}\mathcal{B}}\, [y]_\mathcal{B}\big)$$

thus concluding the coordinate formula

$$[f(y)]_\mathcal{A} = [f]_{\mathcal{A}\mathcal{B}}\, [y]_\mathcal{B}$$

Proposition 3.5. *Let P, N, M be finitely generated free R-modules with bases $\mathcal{C} = (\gamma_1, \dots, \gamma_p)$, $\mathcal{B} = (\beta_1, \dots, \beta_n)$, respectively $\mathcal{A} = (\alpha_1, \dots, \alpha_m)$.*

1) *The canonical map $\Psi_{\mathcal{A}\mathcal{B}}$ below is an isomorphism of R-modules:*
$$\Psi_{\mathcal{A}\mathcal{B}} : \mathrm{Hom}_R(N, M) \to R^{m\times n}, \quad f \mapsto [f]_{\mathcal{A}\mathcal{B}}$$
i.e., $[f + g]_{\mathcal{A}\mathcal{B}} = [f]_{\mathcal{A}\mathcal{B}} + [f]_{\mathcal{A}\mathcal{B}}$, and $[r \cdot f]_{\mathcal{A}\mathcal{B}} = r \cdot [f]_{\mathcal{A}\mathcal{B}}$.
2) *Given morphisms $g : P \to N$, $f : N \to M$, one has: $[f \circ g]_{\mathcal{A}\mathcal{C}} = [f]_{\mathcal{A}\mathcal{B}}\, [g]_{\mathcal{B}\mathcal{C}}$.*
3) *Suppose that $M = N$, $\mathcal{A} = \mathcal{B}$, and set $[f]_\mathcal{A} := [f]_{\mathcal{A}\mathcal{A}}$. Then the canonical map*
$$\Psi_\mathcal{A} : \mathrm{End}_R(M) \to R^{m\times m}, \quad f \mapsto [f]_\mathcal{A}$$
is an isomorphism of rings.

Proof. ... □

3.3. **Change of basis formulas.**

Let M be a finitely free R-module, and $\mathcal{A} = (\alpha_1, \ldots, \alpha_m)$, $\mathcal{A}' = (\alpha'_1, \ldots, \alpha'_m)$ be bases of M. Applying the above Proposition for the identity morphism $\mathrm{id}_M : M \to M$ defined by $\mathrm{id}_M(x) = x$ for all $x \in M$, in the bases $\mathcal{A}'$ and $\mathcal{A}$, one gets: There exists unique matrices $S_{\mathcal{A}\mathcal{A}'} := [\mathrm{id}_M]_{\mathcal{A}\mathcal{A}'}$ and $S_{\mathcal{A}'\mathcal{A}} := [\mathrm{id}_M]_{\mathcal{A}'\mathcal{A}}$ in $R^{m\times m}$ satisfying:

$$\mathcal{A}' = \mathrm{id}_M(\mathcal{A}') = \mathcal{A} \cdot S_{\mathcal{A}\mathcal{A}'}, \quad \mathcal{A} = \mathrm{id}_M(\mathcal{A}) = \mathcal{A}' \cdot S_{\mathcal{A}'\mathcal{A}}$$

Hence $\mathcal{A} = \mathcal{A}' \cdot S_{\mathcal{A}'\mathcal{A}} = \left(\mathcal{A} \cdot S_{\mathcal{A}\mathcal{A}'}\right) S_{\mathcal{A}'\mathcal{A}} = \mathcal{A} \cdot \left(S_{\mathcal{A}\mathcal{A}'} S_{\mathcal{A}'\mathcal{A}}\right)$, thus $\boldsymbol{I}_m = S_{\mathcal{A}\mathcal{A}'} S_{\mathcal{A}'\mathcal{A}}$ (WHY). Similarly, $\boldsymbol{I}_m = S_{\mathcal{A}'\mathcal{A}} S_{\mathcal{A}\mathcal{A}'}$ (WHY). Therefore, $S_{\mathcal{A}\mathcal{A}'}$ and $S_{\mathcal{A}'\mathcal{A}}$ are *inverse to each other* in the ring $R^{m\times m}$. Further, concerning the changing of coordinate vectors, one has:

$$[x]_{\mathcal{A}'} = S_{\mathcal{A}'\mathcal{A}}\, [x]_{\mathcal{A}}$$

This justifies the following:

Definition 3.6. The matrix $S_{\mathcal{A}'\mathcal{A}}$ is called the matrix of change of basis from $\mathcal{A}$ to $\mathcal{A}'$.

Proposition 3.7. *Let N, M be free R-modules with bases $\mathcal{B}$, $\mathcal{B}'$, respectively $\mathcal{A}$, $\mathcal{A}'$.*

1) *For every morphism $f : N \to M$, one has:*

$$[f]_{\mathcal{A}'\mathcal{B}'} = S_{\mathcal{A}'\mathcal{A}}\, [f]_{\mathcal{A}\mathcal{B}}\, S_{\mathcal{B}\mathcal{B}'}$$

2) *Let $M = N$, $\mathcal{A} = \mathcal{B}$, $\mathcal{A}' = \mathcal{B}'$. Then*

$$[f]_{\mathcal{A}'} = S_{\mathcal{A}'\mathcal{A}}\, [f]_{\mathcal{A}}\, S_{\mathcal{A}\mathcal{A}'} = S_{\mathcal{A}'\mathcal{A}}\, [f]_{\mathcal{A}}\, S_{\mathcal{A}'\mathcal{A}}^{-1}$$

Proof. ... □

3.4. **Elementary matrices & row/column transformations.**

Let R be a commutative ring with $0_R \neq 1_R$. Recall that the R-module $R^{m\times n}$ has as standard basis $(\boldsymbol{e}_{kl})_{k,l}$. Hence if $(\boldsymbol{e}'_{k'l'})_{k',l'}$ is the standard basis of $R^{n\times p}$, then the product $\boldsymbol{e}_{kl} \cdot \boldsymbol{e}'_{k'l'} \in R^{m\times p}$ is defined (WHY). Let $(\boldsymbol{e}''_{k''l''})_{k'',l''}$ be the standard basis of $R^{m\times p}$.

Key Lemma 3.8 (Multiplication of standard basis matrices). *In the above notation, one has the following rules for the multiplication of standard basis matrices:*

(i) $\boldsymbol{e}_{kl} \cdot \boldsymbol{e}'_{k'l'} = \boldsymbol{e}''_{kl'} \in R^{m\times p}$ *if* $l = k'$.

(ii) $\boldsymbol{e}_{kl} \cdot \boldsymbol{e}'_{k'l'} = \boldsymbol{0}_{m\times p} \in R^{m\times p}$ *if* $l \neq k'$.

In particular, the product of any two standard basis matrices, if defined, is either the zero matrix or a standard basis matrix.

Proof. Indeed, recall that $\boldsymbol{e}_{kl} = (\delta_{ki}\delta_{lj})_{i,j} \in R^{m\times n}$ and $\boldsymbol{e}'_{k'l'} = (\delta_{k'i}\delta_{l'j})_{i,j} \in R^{n\times p}$. Hence by the rules of multiplication of matrices, the coefficients a_{ij} of the matrix $\boldsymbol{e}_{kl} \cdot \boldsymbol{e}'_{k'l'} =: (a_{ij})_{i,j} \in R^{m\times p}$ are given by:

$$a_{ij} \overset{\text{why}}{=} \sum_{\mu} (\delta_{ki}\delta_{l\mu}) \cdot (\delta_{k'\mu}\delta_{l'j} \overset{\text{why}}{=} \delta_{ki} \cdot \delta_{lk'} \cdot \delta_{l'j}$$

Hence if $l = k'$, then $a_{ij} = \delta_{ki} \cdot \delta_{l'j}$ for all $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant p$, hence $(a_{ij})_{i,j} = \boldsymbol{e}''_{kl'}$. If $l \neq k$ then $a_{ij} = 0_R$ for all i, j (WHY), thus $(a_{ij})_{i,j} = \boldsymbol{0}_{m\times p}$ is the zero matrix. □

Corollary 3.9. *Let $\boldsymbol{e}_{kl} \in R^{m\times m}$ or $\boldsymbol{e}_{kl} \in R^{n\times n}$ be given, and $A = (a_{ij})_{i,j} \in R^{m\times n}$ have rows $\boldsymbol{\mathcal{R}}_i$, $i = 1,\ldots,m$ and columns $\boldsymbol{\mathcal{C}}_j$, $j = 1,\ldots,n$. Then one has:*

1) *$A' := \boldsymbol{e}_{kl}A$ has rows $\boldsymbol{\mathcal{R}}'_i$ satisfying $\boldsymbol{\mathcal{R}}'_i = \delta_{ik}\boldsymbol{\mathcal{R}}_l$, i.e., $\boldsymbol{\mathcal{R}}'_k = \boldsymbol{\mathcal{R}}_l$ and $\boldsymbol{\mathcal{R}}'_i = \mathbf{0}_{1\times n}$ for $i \neq k$.*
2) *$A' := A\boldsymbol{e}_{kl}$ has columns $\boldsymbol{\mathcal{C}}'_j$ satisfying $\boldsymbol{\mathcal{C}}'_j = \delta_{jl}\,\boldsymbol{\mathcal{C}}_k$, i.e., $\boldsymbol{\mathcal{C}}'_l = \boldsymbol{\mathcal{C}}_k$ and $\boldsymbol{\mathcal{C}}'_j = \mathbf{0}_{m\times 1}$ for $j \neq l$.*

Proof. **Ex** (apply the definition of multiplication of matrices together with Key Lemma above, etc. . . .) □

Definition 3.10. The $m \times m$ elementary matrices over R are the matrices defined below:

- $E_{kl}(a) := \boldsymbol{I}_m + a \cdot \boldsymbol{e}_{kl}$ defined for $1 \leqslant k, l \leqslant m$, $k \neq l$, and $a \in R$.
  That is, if $E_{kl}(a) = (a_{ij})_{i,j}$, then: $a_{kl} = a$, $a_{ii} = 1_R$ for all $1 \leqslant i \leqslant m$, $a_{ij} = 0_R$ else.
- $E(k,l) = \boldsymbol{I}_m - \boldsymbol{e}_{kk} - \boldsymbol{e}_{ll} + \boldsymbol{e}_{kl} + \boldsymbol{e}_{lk}$ defined for $1 \leqslant k, l \leqslant m$.
  That is, if $E(k,l) = (a_{ij})_{i,j}$, then: $a_{ii} = 1_R$ for $i \neq k, l$, $a_{kl} = 1_R = a_{lk}$, and $a_{ij} = 0_R$ else.
- $E_k(a) = \boldsymbol{I}_m + (a - 1_R) \cdot \boldsymbol{e}_{kk}$ defined for $1 \leqslant k \leqslant m$, and $a \in R$, $a \neq 0_R$.
  That is, if $E_k(a) = (a_{ij})_{i,j}$, then: $a_{ii} = 1_R$ for $i \neq k$, $a_{kk} = a$, and $a_{ij} = 0_R$ else.

Definition 3.11. An elementary row operation on matrices $A \in R^{m\times m}$ is of one of the following operations performed on A, having $A' \in R^{m\times m}$ as a result:

a) Replace the row $\boldsymbol{\mathcal{R}}_k$ of A by $\boldsymbol{\mathcal{R}}_k + a \cdot \boldsymbol{\mathcal{R}}_l$, for any $1 \leqslant k, l \leqslant m$, $k \neq l$, $a \in R$
b) Interchange the k^{th} row $\boldsymbol{\mathcal{R}}_k$ of A with the l^{th} row $\boldsymbol{\mathcal{R}}_l$ of A.
c) Replace the k^{th} row $\boldsymbol{\mathcal{R}}_k$ of A by its multiple $a \cdot \boldsymbol{\mathcal{R}}_k$ for $a \in R$, $a \neq 0_R$.

Define correspondingly the elementary column operation on matrices $A \in R^{m\times m}$.

Proposition 3.12. *Let $A \in R^{m\times m}$ be given. TFH:*

1) *The elementary matrices satisfy the following:*
   - *$E_{kl}(a) \in R^{m\times m}$ is invertible, and $E_{kl}(a)^{-1} = E_{kl}(-a)$.*
   - *$E(k,l) = E(l,k)$ and $E(k,l) \in R^{m\times m}$ is invertible, having $E(k,l)^{-1} = E(k,l)$.*
   - *$E_k(a) \in R^{m\times m}$ is invertible iff $a \in R^\times$ and if so, then $E_k(a)^{-1} = E(a^{-1})$.*
2) *The row operations on $A \in R^{m\times m}$ relate to elementary matrices as follows:*
   - *The result of the row operation* a) *on A is the matrix $E_{kl}(a)A$.*
   - *The result of the row operation* b) *on A is the matrix $E(k,l)A$.*
   - *The result of the row operation* c) *on A is the matrix $E_k(a)A$.*

Proof. **Ex** (use the rule of multiplication of standard basis elements . . .) □

Definition 3.13. Let $A = (a_{ij})_{i,j} \in R^{m\times n}$ be an $m \times n$ matrix over R.

I) Row reduced (echelon) form
   a) We say that A is (in) row reduced (form), if either $A = 0_{m\times n}$, or there exist $1 \leqslant r \leqslant m$ and $1 \leqslant j_1 \leqslant \cdots \leqslant j_r \leqslant n$ strictly increasing, such that for $1 \leqslant \alpha \leqslant r$, the a_{ij} satisfy:
      - $a_{\alpha j_\alpha} \neq 0_R$ and $a_{\alpha' j_\alpha} = 0_R$ for $\alpha' \neq \alpha$.
      - $a_{ij} = 0_R$ for $i \geqslant \alpha$ and $j < j_\alpha$.
   b) If A is in row reduced form as above, the entries $a_{1j_1}, \ldots, a_{rj_r}$ are the pivots of A.
   c) A row reduced matrix having all pivots equal to 1_R is called row reduced echelon form.
II) Define correspondingly the column reduced (echelon) form of a matrix.

Proposition 3.14. *Let R an integral domain. For $A \in R^{m\times n}$ the following hold:*

1) *There exist* **effectively computable** *elementary matrices $E_1, \dots, E_s \in R^{m\times m}$ such that $E_s \dots E_1 A$ is in row reduced (echelon) form (if $F := R$ is a field).*
2 *There exist* **effectively computable** *elementary matrices $E'_1, \dots, E'_{s'} \in R^{n\times n}$ such that $AE'_1 \dots E'_{s'}$ is in column reduced (echelon) form (if $F := R$ is a field).*

Proof. To 1): Make induction on m, etc. To 2): ibidem... □

- Application: Effective method to compute the inverse of matrix $A \in F^{m\times m}$

Let F be a field, $A \in F^{m\times m}$ be a matrix. The following hold:

- Let $E_1, \dots, E_s \in R^{m\times m}$ be (effectively computable) elementary matrices such that:

$$A' := E_s \dots E_1 A \quad \textit{is in row reduced echelon form.}$$

- Let $a_{1j_1} = \cdots = a_{rj_r} = 1_F$ be the pivots of A'.
- If $r < m$, the rows $\boldsymbol{\mathcal{R}}'_i$ of A' equal $0_m := (0_F, \dots, 0_F)$ for $r < i \leqslant m$ (WHY). Therefore, for every $B \in F^{m\times m}$ one has: The rows $\boldsymbol{\mathcal{R}}''_i$ of $A'' := A'B$ equal 0_m for $r < i \leqslant m$ (WHY). Hence $A'B \neq \boldsymbol{I}_m$ for all $B \in F^{m\times m}$, hence A' is not invertible (WHY).
- Since $E_1, \dots, E_s$ are invertible, for the matrix $A' = E_s \dots E_1 A$, one has:

$$A \text{ is invertible iff } A' \text{ is invertible (WHY).}$$

- Conclude: *A is invertible iff $r = m$ iff $A' = \boldsymbol{I}_m$.*

Procedure: Given $A \in F^{m\times m}$ any $m \times m$ matrix, let $E_1, \dots, E_s \in \mathrm{GL}_m(F)$ be elementary matrices such that $A' := E_s \dots E_1 A$ is in row reduced echelon form.

- If $A' \neq \boldsymbol{I}_m$, then A is not invertible. **STOP**
- If $A' = \boldsymbol{I}_m$, then set:

$$\widetilde{A} := \left(A \;\middle|\; \boldsymbol{I}_m\right) = \left(A \;\middle|\; \boldsymbol{0}_{m\times m}\right) + \left(\boldsymbol{0}_{m\times m} \;\middle|\; \boldsymbol{I}_m\right) \in F^{m\times 2m}$$

  Then one gets:

$$E_s \dots E_1 \widetilde{A} = \left(E_s \dots E_1 \widetilde{A} \;\middle|\; E_s \dots E_1 \boldsymbol{I}_m\right) = \left(\boldsymbol{I}_m \;\middle|\; A^{-1}\right) \quad \text{(WHY)}$$

- **Conclude:** *The row reduced echelon form $\widetilde{A}'$ of $\widetilde{A}$ is nothing but*

$$\widetilde{A}' = \left(\boldsymbol{I}_m \;\middle|\; A^{-1}\right)$$

Example 3.15. ...

- The row/column R-modules of a matrix

Let $A = (a_{ij})_{i,j} \in R^{m\times n}$ be given. Define:

- $\boldsymbol{\mathcal{C}}_A := \langle \boldsymbol{\mathcal{C}}_1, \dots, \boldsymbol{\mathcal{C}}_n \rangle_R \subset {}^mR$ the span of the set of columns of A.
- $\boldsymbol{\mathcal{R}}_A := \langle \boldsymbol{\mathcal{R}}_1, \dots, \boldsymbol{\mathcal{R}}_m \rangle_R \subset R^n$ the span of the set of rows of A.

Note that A gives rise canonically to morphisms of R-modules

$$ {}_A\varphi : {}^nR \to {}^mR, \quad {}_A\varphi(\boldsymbol{x}) := A \cdot \boldsymbol{x} \ \text{ for } \ \boldsymbol{x} \in {}^nR \quad \text{(WHY)} $$

$$ \varphi_A : R^m \to R^n, \quad \varphi_A(\boldsymbol{x}) := \boldsymbol{x} \cdot A \ \text{ for } \ \boldsymbol{x} \in R^m \quad \text{(WHY)} $$

Further, if $\mathcal{E}_n$ and $\mathcal{E}_m$ are the standard basis of ${}^nR$, respectively ${}^mR$, one has:

$$ {}_A\varphi(\mathcal{E}_n) = \mathcal{E}_m A, \ \text{ hence } \ [{}_A\varphi]_{\mathcal{E}_m\mathcal{E}_n} = A \quad \text{(WHY)} $$

Proposition 3.16. *Let $A \in R^{m\times n}$ be given. In the above notations, one has:*
Concerning $\boldsymbol{\mathcal{R}}_A$.

1) *For every $E \in R^{m\times m}$, one has $\boldsymbol{\mathcal{R}}_{EA} \subset \boldsymbol{\mathcal{R}}_A$, and if $E \in \mathrm{GL}_m(R)$, then $\boldsymbol{\mathcal{R}}_{EA} = \boldsymbol{\mathcal{R}}_A$.*
   In particular, if E is a product of invertible elementary $m\times m$ matrices, then $\boldsymbol{\mathcal{R}}_{EA} = \boldsymbol{\mathcal{R}}_A$.
2) *$\boldsymbol{\mathcal{C}}_{EA} = {}_E\varphi(\boldsymbol{\mathcal{C}}_A)$, and if $E \in \mathrm{GL}_m(R)$, then ${}_E\varphi : \boldsymbol{\mathcal{C}}_A \to \boldsymbol{\mathcal{C}}_{EA}$ is an isomorphism.*
   In particular, if E is a product of invertible elementary $m\times m$ matrices, then $\boldsymbol{\mathcal{C}}_A \cong \boldsymbol{\mathcal{C}}_{EA}$.

Concerning $\boldsymbol{\mathcal{C}}_A$: *The same holds correspondingly.*

Proof. **Ex** . . . □

Definition 3.17. If $\boldsymbol{\mathcal{R}}_A$ and/or $\boldsymbol{\mathcal{C}}_A$ are **free** R-modules, their ranks $\mathrm{rk}(\boldsymbol{\mathcal{R}}_A)$, $\mathrm{rk}(\boldsymbol{\mathcal{C}}_A)$ are called the row/column ranks of A.

Proposition 3.18. *Let $A \in R^{m\times n}$ be such that $\boldsymbol{\mathcal{R}}_A$ and/or $\boldsymbol{\mathcal{C}}_A$ are **free** R-modules, hence $\mathrm{rk}(\boldsymbol{\mathcal{R}}_A)$, $\mathrm{rk}(\boldsymbol{\mathcal{C}}_A)$ are defined. TFH:*

1) *$\mathrm{rk}(\boldsymbol{\mathcal{R}}_A)$ and/or $\boldsymbol{\mathcal{C}}_A$ are invariant under elementary invertible row/column operations.*
2) *If $R = F$ is a field, then $\boldsymbol{\mathcal{R}}_A \subset F^n$, $\boldsymbol{\mathcal{C}}_A \subset {}^mF$ are vector subspaces, and one has:*
   a) *$\boldsymbol{\mathcal{R}}_A \subset F^n$ and $\boldsymbol{\mathcal{C}}_A \subset {}^mF$ are free F-modules, hence $\mathrm{rk}(\mathcal{R}_A)$ and $\mathrm{rk}(\mathcal{C}_A)$ are defined.*
   b) *$\mathrm{rk}(\boldsymbol{\mathcal{R}}_A)$, $\mathrm{rk}(\boldsymbol{\mathcal{C}}_A)$ are invariant under elementary row/column operations.*
   c) *$\mathrm{rk}(\boldsymbol{\mathcal{R}}_A)$, $\mathrm{rk}(\boldsymbol{\mathcal{C}}_A)$ equal the number of pivots in the row/column reduced forms of A.*
   $(*)$ *Finally, $\mathrm{rk}(\boldsymbol{\mathcal{R}}_A) =: \mathrm{rk}(A) := \mathrm{rk}(\boldsymbol{\mathcal{C}}_A)$, and $\mathrm{rk}(A)$ is called the rank of A.*

Proof. To 1), 2a), 2b), 2c): Apply the above Proposition.
To 2d): Let $E \in \mathrm{GL}_m(F)$ be a product of elementary matrices such that EA is in row reduced echelon form. Then by the previous Proposition one has: $T_E : \boldsymbol{\mathcal{C}}_A \to \boldsymbol{\mathcal{C}}_{EA}$ is an isomorphism, hence $\mathrm{rk}(\boldsymbol{\mathcal{C}}_A) = \mathrm{rk}(\boldsymbol{\mathcal{C}}_{EA})$. On the other hand, if $A' := EA \in F^{m\times n}$ is in row reduced echelon form, and easily gets the *column reduced echelon form* of A by performing the obvious $n(n-1)/2$ elementary row column operations. □

4. Systems of Linear Equations

4.1. Definitions, Basic facts.

Let R be a ring with $1_R \neq 0_R$. A linear system of equations over a R is a symbol of the form:

$$ \mathcal{S}\colon A\,\boldsymbol{x} = \boldsymbol{b}, \quad A \in R^{m\times n} \ \text{ and } \ \boldsymbol{b} \in {}^mR. $$

Definition/Remark 4.1. Let a linear systems of equations $\mathcal{S} \ : \ A \cdot \boldsymbol{a} = \boldsymbol{b}$ be given.

1) The solutions of $\mathcal{S}$ are the elements $\boldsymbol{a} \in {}^nR$ such that $A \cdot \boldsymbol{a} = \boldsymbol{b}$. Equivalently, if the matrix A has the columns $A = \big(\boldsymbol{c}_j\big)_{j=1,\dots,n}$, then $\boldsymbol{a} \in {}^nR$ is a solution of $\mathcal{S}$ iff $\boldsymbol{b}$ is a linear combination of the columns $(\boldsymbol{c}_j)_j$ as follows:

$$\boldsymbol{b} = \boldsymbol{c}_1 a_1 + \cdots + \boldsymbol{c}_n a_n = \big(\boldsymbol{c}_1, \dots, \boldsymbol{c}_n\big) \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

2) *Solving* $\mathcal{S}$ means to give a complete description of the set $Sol(\mathcal{S}) \subset {}^nR$ of all solutions of $\mathcal{S}$, which might be the empt set.

3) Let $A, A' \in R^{m\times n}$ and $\boldsymbol{b}, \boldsymbol{b}' \in R^{m\times 1}$ be given. We say that the systems of equations $\mathcal{S}$: $A\boldsymbol{x} = \boldsymbol{b}$ and $\mathcal{S}'$: $A'\boldsymbol{x} = \boldsymbol{b}'$ are equivalent, if $Sol(\mathcal{S}) = Sol(\mathcal{S}')$.

(*) Obviously, solving $\mathcal{S}: A\boldsymbol{x} = \boldsymbol{b}$ is equivalent to solving some $\mathcal{S}: A'\boldsymbol{x} = \boldsymbol{b}'$ which is equivalent to the given $\mathcal{S}$: $A\boldsymbol{x} = \boldsymbol{b}$.

Definition/Remark 4.2. Given $\mathcal{S}$: $A\,\boldsymbol{x} = \boldsymbol{b}$, the homogeneous system attached to $\mathcal{S}$ is

$$\mathcal{S}_0 : A\,\boldsymbol{x} = 0_{m\times 1}$$

Note that $0_{n\times 1} \in Sol(\mathcal{S}_0)$, hence $Sol(\mathcal{S}_0)$ is always non-empty.

Proposition 4.3. *In the above notations, TFH:*

1) *If $\boldsymbol{a}, \boldsymbol{a}' \in Sol(\mathcal{S})$, then $\boldsymbol{a}' - \boldsymbol{a} \in Sol(\mathcal{S}_0)$.*
2) *If $\boldsymbol{a} \in Sol(\mathcal{S})$ and $\boldsymbol{a}_0 \in Sol(\mathcal{S}_0)$, then $\boldsymbol{a} + \boldsymbol{a}_0 \in Sol(\mathcal{S})$.*

Therefore, if $\boldsymbol{a} \in Sol(\mathcal{S})$, one has: $Sol(\mathcal{S}) = \boldsymbol{a} + Sol(\mathcal{S}_0)$.

Proof. ... □

4.2. **Geometric interpretation.**

Let $A \in R^{m\times n}$ be given. Then A gives rise canonically to morphism of R-modules

$${}_A\varphi : {}^nR \to {}^mR, \quad {}_A\varphi(\boldsymbol{x}) := A \cdot \boldsymbol{x} \quad \text{for} \quad \boldsymbol{x} \in {}^nR \quad \text{(WHY)}.$$

Further, if $\mathcal{E}_n$ and $\mathcal{E}_m$ are the standard basis of ${}^nR$, respectively ${}^mR$, one has:

$${}_A\varphi(\mathcal{E}_n) = \mathcal{E}_m A, \quad \text{hence} \quad [{}_A\varphi]_{\mathcal{E}_m\mathcal{E}_n} = A \quad \text{(WHY)}$$

Proposition 4.4. *Let $\mathcal{S}$: $A\boldsymbol{x} = \boldsymbol{b}$ be a system of linear equations. TFH:*

1) $\mathrm{Ker}({}_A\varphi) = Sol(\mathcal{S}_0)$ *and* $\mathrm{Im}({}_A\varphi) = {}_A\varphi({}^nR) = \boldsymbol{C}_A \subset {}^mR$.
b) $Sol(\mathcal{S}) = {}_A\varphi^{-1}(\boldsymbol{b})$, *and in particular, $Sol(\mathcal{S})$ is non-empty iff $\boldsymbol{b} \in \mathrm{Im}({}_A\varphi)$.*

Proof. ... □

4.3. **Solving linear systems of equations: The Gauss Method.**

Let $\mathcal{S}$: $A\boldsymbol{x} = \boldsymbol{b}$ be a liner system of equations.

- For $E \in R^{m\times m}$ arbitrary, set $A' := EA$, $\boldsymbol{b}' := E\boldsymbol{b}$, and get $\mathcal{S}'$: $A'\boldsymbol{x} = \boldsymbol{b}'$.
- If $\boldsymbol{a} \in Sol(\mathcal{S})$, then $\boldsymbol{a} \in Sol(\mathcal{S}')$ (WHY), hence $Sol(\mathcal{S}) \subset Sol(\mathcal{S}')$.
- In particular, if $E \in \mathrm{GL}_m(R)$, and $E^{-1} =: E'$, then setting $A'' := E'A'$ and $\boldsymbol{b}'' := E'\boldsymbol{b}'$, one has: $A'' = A$, $\boldsymbol{b}'' = \boldsymbol{b}$ (WHY).
- Hence $\mathcal{S}''$: $A''\boldsymbol{x} = \boldsymbol{b}''$ is actually identical with $\mathcal{S}$: $A\boldsymbol{x} = \boldsymbol{b}$, and therefore:

$$Sol(\mathcal{S}) \subset Sol(\mathcal{S}') \subset Sol(\mathcal{S}'') = Sol(\mathcal{S}), \quad \text{thus} \quad Sol(\mathcal{S}) = Sol(\mathcal{S}')$$

Now suppose that $E := E_1 \dots E_s \in R^{m\times m}$ is a product of elementary matrices such that $A' := EA$ is in *row reduced form.* Then in the above notations,

- $Sol(\mathcal{S}) \subset Sol(\mathcal{S}')$, and $Sol(\mathcal{S}) = Sol(\mathcal{S}')$ if E is invertible.
- In particular, if $R = F$ is a field, one the following:

The Gauss Method

Given $\mathcal{S}$: $A\boldsymbol{x} = \boldsymbol{b}$, choose elementary matrices $E_1, \dots, E_s \in R^{m\times m}$ such that the matrix $A' = (a'_{ij})_{i,j} = E_r \dots E_1 A$ is in ***row reduced echelon form***. The following hold:

a) Let $a_{1j_1} = \cdots = a_{rj_r} = 1$ be the pivots of $A' = (a'_{ij})_{i,j}$, and recall the standard basis $(\boldsymbol{e}_j)_j$ of ${}^nF$. Then $Sol(\mathcal{S}_0)$ has as standard (or canonical) basis the system $(\alpha_j)_{j\neq j_1,\dots,j_r}$, which setting $j_{r+1} := n+1$, is defined as follows:

$$\alpha_j := \boldsymbol{e}_j, j < j_1, \text{ and } \alpha_j := a'_{1j}\boldsymbol{e}_{j_1} + \cdots + a'_{\nu j}\boldsymbol{e}_{j_\nu} - \boldsymbol{e}_j, 1 \leqslant \nu \leqslant r,\ j_\nu < j < j_{\nu+1} \text{ (WHY)}$$

b) $Sol(\mathcal{S}) \neq \emptyset$ iff $b'_i = 0_F$ for $i > r$ (WHY). If so, then $\boldsymbol{a} := (a_j)_j^\tau \in Sol(\mathcal{S})$, where: $a_{j_\nu} = b'_\nu$ for $\nu = 1, \dots, r$, and $a_j = 0_F$ else. Hence $Sol(\mathcal{S}) = \boldsymbol{a} + \langle\!\langle (\alpha_j)_{j\neq j_1,\dots,j_r} \rangle\!\rangle_F$.

Example 4.5. Let the row reduced echelon form of $\mathcal{S}$: $A\boldsymbol{x} = \boldsymbol{a}$ be the following:

$$\left(\begin{array}{ccccccc|c} 0 & 1 & a'_{13} & a'_{14} & 0 & 0 & a'_{17} & b'_1 \\ 0 & 0 & 0 & 0 & 1 & 0 & a'_{27} & b'_2 \\ 0 & 0 & 0 & 0 & 0 & 1 & a'_{37} & b'_3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & b'_4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & b'_5 \end{array}\right)$$

The second indices of the pivots are $j_1 = 2, j_2 = 5, j_3 = 6$ (WHY) and therefore, $j \neq j_1, j_2, j_3$ if and only if $j = 1, 3, 4, 7$. Conclude that $Sol(\mathcal{S}_0)$ has as canonical (or standard) basis $(\alpha_1, \alpha_3, \alpha_4, \alpha_7)$, which is defined as follows:

- $\alpha_1 = (1, 0, 0, 0, 0, 0, 0)^\tau$
- $\alpha_3 = a'_{13}\boldsymbol{e}_2 - \boldsymbol{e}_3 = (0, a'_{13}, -1, 0, 0, 0, 0)^\tau$
- $\alpha_4 = a'_{14}\boldsymbol{e}_2 - \boldsymbol{e}_4 = (0, a'_{14}, 0, -1, 0, 0, 0)^\tau$
- $\alpha_7 = a'_{17}\boldsymbol{e}_2 + a'_{27}\boldsymbol{e}_5 + a'_{37}\boldsymbol{e}_6 - \boldsymbol{e}_7 = (0, a'_{17}, 0, 0, a'_{27}, a'_{37}, -1)^\tau$

- If $b'_4 \neq 0$ or $b'_5 \neq 0$, then $Sol(\mathcal{S}) = \emptyset$. Otherwise, $\boldsymbol{a} = (0, b'_1, 0, 0, b'_2, b'_3, 0)^\tau \in Sol(\mathcal{S})$ (WHY).

Remark 4.6. Renumbering unknowns: Let $\mathcal{S}$: $A\boldsymbol{x} = \boldsymbol{b}$ be a system of linear equations, and w.l.o.g., suppose that A is already in reduced echelon form. Another trick/method to simplify the description of the $Sol(\mathcal{S})$ is to *renumber the unknowns* in such a way that the second index

of the pivots are $j_1 = 1, \ldots, j_r = r$. Precisely, given the columns $\boldsymbol{c}_1 x_1, \ldots, \boldsymbol{c}_n x_x$ of $A\boldsymbol{x}$, one swaps/permutes inductively: $\boldsymbol{c}_1 x_1$ with $\boldsymbol{c}_{j_1} x_{j_1}$ and denotes $\boldsymbol{c}_{j_1} x_{j_1}$ by $\boldsymbol{c}'_1 x'_1$, respectively $\boldsymbol{c}_1 x_1$ by $\boldsymbol{c}'_{j_1} x'_{j_1}$. On proceedes inductively (note that in this process, the new x'_{j_1}, as well as other unknowns might need to be renumbered more than one time, and columns correspondingly permuted). After at most r steps of renumbering unknowns and swapping columns, one gets a system of m linear of equations in the (new) unknowns $\boldsymbol{x}' = (x'_1, \ldots, x'_n)^\tau$ the form:

$$\mathcal{S}' \colon A'\boldsymbol{x}' = \boldsymbol{b}, \quad A' = \begin{pmatrix} \boldsymbol{I}_r & (a'_{ij})_{i,j} \\ \boldsymbol{0} & \boldsymbol{0}' \end{pmatrix},$$

where $\boldsymbol{0}, \boldsymbol{0}'$ are zero-matrices of the right size (which might be empty), $1 \leqslant i \leqslant r$, and $r < j \leqslant n$, provided $r < n$. Then the pivots of A' have indices $j_1 = 1, \ldots, j_r = r$. And if $r < n$, the standard basis of $Sol(\mathcal{S}'_0)$ is given by $(\alpha_j)_{r<j\leqslant n}$ with $\alpha'_j = \sum_i a'_{ij}\boldsymbol{e}_i - \boldsymbol{e}_j$ (WHY).

Finally, the description of $Sol(\mathcal{S}_0)$ for a system $\mathcal{S} \colon A\boldsymbol{x} = \boldsymbol{b}$ given in the Gauss method explained before the Example, implies the following:

Theorem 4.7. *Let $A \in R^{m\times m}$ be given. TFH:*

1) *The row reduced echelon form of A is unique, provided it exists.*
(†) *In general, if R has no zero divisors $\neq 0_R$, e.g., if R is a field, then given row reduced forms $B, C \in R^{m\times m}$ of A, say having pivots $b_{1j_1}, \ldots, b_{rj_r}$ and $c_{1k_1}, \ldots, c_{sk_s}$, one has:*
   - $r = s$ and $j_\alpha = k_\alpha$ for $1 \leqslant \alpha \leqslant r$.
   - If B and C has equal corresponding pivots $b_{\alpha j_\alpha} = c_{\alpha j_\alpha}$, $1 \leqslant \alpha \leqslant r$, then $B = C$.
2) *Correspondingly, the same holds for the column reduced (echelon) form.*

Proof. To 1): Let A', A'' be a row reduced echelon forms of A. Then $\mathcal{R}_{A'} = \mathcal{R}_A = \mathcal{R}_{A''}$ (WHY), and the rows $\mathcal{A}' := (\boldsymbol{\mathcal{R}}'_i)_{i=1,\ldots,r'}$, $\mathcal{A}'' := (\boldsymbol{\mathcal{R}}''_i)_{i=1,\ldots,r''}$ are bases of $\mathcal{R}_A$. Hence $r' =: r := r''$ and $\boldsymbol{\mathcal{R}}'_i = \boldsymbol{\mathcal{R}}''_i$, $1 \leqslant i \leqslant r$ (WHY). □

5. Multilinear Maps / Multilinear Forms

5.1. Dual of a module / vector space.

Let R be a commutative ring with $0_R \neq 1_R$, and M be an R-module.

Definition/Remark 5.1. The R-module $M^\vee := \mathrm{Hom}_R(M, R)$ is called the dual module of M. If V is an F-vector space, $V^\vee$ is called the dual (vector) space of V.

Remark 5.2. The following hold:

1) Let $f : N \to M$ be a morphism. Then for every $\varphi \in M^\vee$, the resulting map
$$\varphi \circ f : N \to M \to R$$
is a morphism (WHY), thus $\varphi \circ f \in N^\vee$.
2) Hence $f : N \to M$ gives rise to a map $f^\vee : M^\vee \to N^\vee$ by $f^\vee(\varphi) := f \circ \varphi$, called the dual of f, which is a morphism of modules (WHY).
3) Moreover, for morphisms $P \xrightarrow{g} N \xrightarrow{f} M$, one has $(f \circ g)^\vee = g^\vee \circ f^\vee$ (WHY).

Proposition 5.3. *Let $f : N \to M$ be a morphism of free R-modules of finite rank, with bases $\mathcal{B} = (\beta_1, \dots, \beta_n)$ and $\mathcal{A} = (\alpha_1, \dots, \alpha_m)$. The following hold:*

1) *$M^\vee$ is a free R-module with basis $\mathcal{A}^\vee := (\alpha_1^\vee, \dots, \alpha_m^\vee)$, called the dual basis to $\mathcal{A}$, where:*

$$\alpha_i^\vee(\alpha_j) = \delta_{ij} := \begin{cases} 1_R & \text{if } i = j \\ 0_R & \text{if } i \neq j \end{cases}$$

2) *The matrix of the dual morphism $f^\vee : N^\vee \to M^\vee$ in the bases $\mathcal{A}^\vee$, $\mathcal{B}^\vee$ is the transpose*

$$[f^\vee]_{\mathcal{B}^\vee\mathcal{A}^\vee} = [f]_{\mathcal{A}\mathcal{B}}^\tau$$

of the matrix $[f]_{\mathcal{A}\mathcal{B}}$ of f in the bases $\mathcal{A}$, $\mathcal{B}$.

Proof. To 1): Since M is free and $\mathcal{A}$ is a basis, and $R, +$ is free with basis $\mathcal{E}_1 := (1)$, one has: The map

$$M^\vee = \mathrm{Hom}_R(M, R) \to R^{1\times m}, \quad \varphi \mapsto [\varphi]_{\mathcal{E}_1\mathcal{A}}$$

is a isomorphisms of R-modules. Let $\mathcal{E}_m = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_m)$ be the standard basis of $R^{1\times m} = R^m$, hence $\boldsymbol{e}_i =: (\delta_{ij})_{j=1,\dots,m}$. Then for every $\boldsymbol{e}_i$ there exists $\varphi_i \in M^\vee$ such that $[\varphi]_{\mathcal{E}_1\mathcal{A}} = \boldsymbol{e}_i$. Equivalently,

$$\varphi(\mathcal{A}) = \big(\varphi(\alpha_1), \dots, \varphi(\alpha_m)\big) = \mathcal{E}_1\boldsymbol{e}_i = (1)\boldsymbol{e}_i = \boldsymbol{e}_i = (\delta_{ij})_{j=1,\dots,m}, \text{ hence } \varphi(\alpha_j) = \delta_{ij}, \text{ thus } \varphi = \alpha_i^\vee .$$

To 2): The column $(a_{ij})_{1\leqslant i\leqslant m}$ of $[f]_{\mathcal{A}\mathcal{B}}$ and the column $(c_{ji})_{1\leqslant j\leqslant n}$ of $[f^\vee]_{\mathcal{B}^\vee\mathcal{A}^\vee}$ are defined by

$$f(\beta_j) = a_{1j}\alpha_1 + \cdots + a_{mj}\alpha_m, \qquad f^\vee(\alpha_i^\vee) = \beta_1^\vee c_{1i} + \cdots + \beta_n^\vee c_{ni}.$$

Evaluating both sides at each β_k, we get: First, since $f^\vee(\alpha_i^\vee) \overset{\text{def}}{:=} \alpha_i^\vee \circ f$, the left hand side is:

$$\big(f^\vee(\alpha_i^\vee)\big)(\beta_k) = (\alpha_i^\vee \circ f)(\beta_k) = \alpha_i^\vee\big(f(\beta_k)\big) = \alpha_i^\vee(a_{1k}\alpha_1 + \cdots + a_{mk}\alpha_m) = a_{ik}.$$

Second, the right hand side is $(\beta_1^\vee c_{1i} + \cdots + \beta_n^\vee c_{ni})(\beta_k) = c_{ki}$. Hence $c_{ki} = a_{ik}$ for all $i = 1, \dots, m$ and $k = 1, \dots, n$, concluding that $[f^\vee]_{\mathcal{B}^\vee\mathcal{A}^\vee} = [f]_{\mathcal{A}\mathcal{B}}^\tau$, as claimed. □

Definition/Remark 5.4.

1) Let M be an R-module. The R-module $M^{\vee\vee} = \mathrm{Hom}_R(M^\vee, R)$ is called the bidual of M. The canonical map

$$\imath : M \to M^{\vee\vee} \text{ defined by } \imath(x) = \varphi(x) \text{ for all } \varphi \in M^\vee$$

is a morphism of R-modules.

2) Let $f : N \to M$ be a morphism. Then one has a commutative diagram of the form

$$\begin{array}{ccc} N & \xrightarrow{f} & M \\ \downarrow{\scriptstyle \imath_N} & & \downarrow{\scriptstyle \imath_M} \\ N^{\vee\vee} & \xrightarrow{f^{\vee\vee}} & N^{\vee\vee} \end{array}$$

3) If M is finitely a generated free R-module, then $\imath : M \to M^{\vee\vee}$ is an *isomorphism.*

5.2. **Bilinear maps / Bilinear forms.**

Definition 5.5. Let M_1, M_2, N be R-modules, and $f \in \mathrm{Maps}(M_1 \times M_2, N)$ be a fixed map.

1) $f : M_1 \times M_2 \to N$ is called R-bilinear, if for all $x, x' \in M_1$, $y, y \in M_2$, $a, a' \in R$ one has:

$$f(ax + a'x', y) = af(x, y) + a'f(x', y), \quad f(x, ay + a'y') = af(x, y) + a'f(x, y').$$

Notation: $\mathcal{L}(M_1 \times M_2, N) := \{f \in \mathrm{Maps}(M_1 \times M_2, N) \mid f \text{ is } R\text{-bilinear}\}$.
If $M_1 = M_2 = M$ we denote: $\mathcal{L}^2(M, N) := \mathcal{L}(M \times M, N)$.

2) $f \in \mathcal{L}^2(M, N)$ is called symmetric, if $f(x_1, x_2) = f(x_2, x_1)$ for all $x_1, x_2 \in M$.

Notation: $\mathcal{L}^2_{\mathrm{sym}}(M, N) = \{ f \in \mathcal{L}^n(M, R) \mid f \text{ symmetric} \}$.

3) $f \in \mathcal{L}^2(M, N)$ is called anti-symmetric, or alternating, if $f(x, x) = 0_N$ for all $x \in M$, and $f(x_2, x_1) = -f(x_1, x_2)$ for all $x_1, x_2 \in M$.

Notation: $\mathcal{L}^2_{\mathrm{alt}}(M, N) = \{ f \in \mathcal{L}^2(M, R) \mid f \text{ alternating} \}$.

Terminology: If $N = R, +$ we speak about bilinear forms on $M_1 \times M_2$, and about symmetric, respectively alternating bilinear forms on M.

Remark 5.6. In the above context, let $f : M_1 \times M_2 \to N$ be an R-bilinear map. Then one has:

1) $\mathrm{Ker}_l(f) := \{x \in M_1 \mid f(x, y) = 0 \, \forall y \in M_2\}$, $\mathrm{Ker}_r(f) := \{y \in M_2 \mid f(x, y) = 0 \, \forall \, x \in M_1\}$ are called the left kernel, respectively right kernel of f.

$*$ Moreover, $\mathrm{Ker}_l(f) \subset M_1$ and $\mathrm{Ker}_r(f) \subset M_2$ are R-submodules (WHY).

Terminology: f is called non-degenerate if $\mathrm{Ker}_l(f) = \{0_{M_1}\}$ and $\mathrm{Ker}_r(f) = \{0_{M_2}\}$.

2) The R-bilinear map $f : M_1 \times M_2 \to N$ gives rise to morphisms of R-modules:

$$f_1 : M_1 \to \mathrm{Hom}_R(M_2, N) \text{ by } f_1(x_1) =: \varphi_{x_1} : M_2 \to N \text{ by } \varphi_{x_1}(y) := f(x_1, y)$$
$$f_2 : M_2 \to \mathrm{Hom}_R(M_1, N) \text{ by } f_2(x_2) =: \varphi_{x_2} : M_1 \to N \text{ by } \varphi_{x_2}(x) := f(x, x_2)$$

And one has: $\mathrm{Ker}_l(f) = \mathrm{Ker}(f_1)$ and $\mathrm{Ker}_r(f) = \mathrm{Ker}(f_2)$ (WHY).

3) If $N = R$, the R-bilinear map $f : M_1 \times M_2 \to R$ gives rise to R-morphisms:

$$f_1 : M_1 \to {M_2}^\vee \text{ by } f_1(x_1) =: \varphi_{x_1} \in {M_2}^\vee, \text{ where } \varphi_{x_1} : M_2 \to R, \; y \mapsto f(x_1, y)$$
$$f_2 : M_2 \to {M_1}^\vee \text{ by } f_2(x_2) =: \varphi_{x_2} \in {M_1}^\vee, \text{ where } \varphi_{x_2} : M_1 \to R, x \mapsto f(x, x_2)$$

4) In particular, f_1 is injective iff $\mathrm{Ker}_l(f) = \{0_{M_1}\}$, and f_2 is injective iff $\mathrm{Ker}_r(f) = \{0_{M_2}\}$. Hence f_1 and f_2 are injective iff f is non-degenerate.

Example 5.7.

1) Let $M_1 := M$ be an R-module, and $M_2 := M^\vee = \mathrm{Hom}_R(M, R)$ be its dual module. Define

$$f : M_1 \times M_2 \to R \quad \text{by} \quad f(x, \varphi) := \varphi(x)$$

Then f is a bilinear form (WHY), and the above morphisms are:

$$f_1 : M = M_1 \to {M_2}^\vee = M^{\vee\vee}, \qquad f_2 : M^\vee = M_2 \to {M_1}^\vee = M^\vee$$

2) Let $M_1 = R^m = R^{1 \times m}$ and $M_2 = {}^mR := R^{m \times 1}$. Then the multiplication of mattrices

$$f : R^m \times {}^mR \to R, \quad f(\boldsymbol{a}, \boldsymbol{b}^\tau) = \boldsymbol{a} \cdot \boldsymbol{b}^\tau$$

is a non-degenerate bilinear form, and the resulting morphisms

$$f_1 : R^m \to ({}^mR)^\vee, \quad f_2 : {}^mR \to (R^m)^\vee$$

are both isomorphisms (WHY).

3) In particular, we can view R^m and ${}^mR$ as dual to each other, and identify ${}^mR$ with $(R^m)^\vee$ and R^m with $({}^mR)^\vee$.

Proposition 5.8. (Properties) *In the above notations one has the following:*

1) $\mathcal{L}(M_1 \times M_2, N) \subset \mathrm{Maps}(M_1 \times M_2, N)$ *is an* R*-submodule.*
Furthermore, $\mathcal{L}^2_{\mathrm{sym}}(M,N), \mathcal{L}^2_{\mathrm{alt}}(M,N) \subseteq \mathcal{L}^2(M,N)$ *are* R*-submodules for all* M, N*.*
2) *Let* $\phi : N \to P$ *and* $\psi_1 \times \psi_2 : L_1 \times L_2 \to M_1 \times M_2$ *be* R*-morphisms.*
Then $\phi \circ f \circ (\psi_1 \times \psi_2) \in \mathcal{L}(L_1 \times L_2, P)$ *for all* $f \in \mathcal{L}(M_1 \times M_2, N)$*.*
3) *Furthermore, if* $\phi : N \to P$ *and* $\psi : L \to M$ *are morphisms, then* $\phi \circ f \circ \psi^2 \in \mathcal{L}^2_{\mathrm{sym}}(L,P)$ *for* $f \in \mathcal{L}^2_{\mathrm{sym}}(M,N)$*, and* $\phi \circ f \circ \psi^2 \in \mathcal{L}^2_{\mathrm{alt}}(L,P)$ *for* $f \in \mathcal{L}^2_{\mathrm{alt}}(M,N)$*.*

Proof. Exercise: direct computations... □

Special case: M_1, M_2 *are free of finite rank.*

Let $\mathcal{A} = (\alpha_1, \dots, \alpha_{m_1})$, $\mathcal{B} = (\beta_1, \dots, \beta_{m_2})$ be bases of M_1, respectively M_2. Then $x \in M_1$ is of the form $x = \mathcal{A} \cdot [x]_{\mathcal{A}}$, and $y \in M_2$ is of the form $y = \mathcal{B} \cdot [y]_{\mathcal{B}}$ (WHY).

Then given an R-bilinear map $f : M_1 \times M_2 \to N$, one has:

$$f(x,y) = f\big(\sum_i x_i\alpha_i, \sum_j y_j\beta_j\big) = \sum_{i,j} x_i y_j f(\alpha_i,\beta_j) \quad \text{(WHY)}$$

In particular, setting $A_f := \big(f(\alpha_i,\beta_j)\big)_{i,j} \in N^{m_1 \times m_2}$, one has the following:

$$f(x,y) = [x]^\tau_{\mathcal{A}} \cdot A_f \cdot [y]_{\mathcal{B}} \quad \text{(WHY)}$$

Let $\mathcal{A}' = (\alpha'_1, \dots, \alpha'_{m_1})$ and $\mathcal{B}' = (\beta'_1, \dots, \beta'_{m_2})$ be further bases of M_1, respectively M_2, and consider the matrix $A'_f = \big(f(\alpha'_i, \beta'_j)\big)_{i,j} \in N^{m_1 \times m_2}$ satisfying

$$f(x,y) = [x]^\tau_{\mathcal{A}'} \cdot A'_f \cdot [y]_{\mathcal{B}'}.$$

Finally, recall the change of basis matrices $S_{\mathcal{A}'\mathcal{A}} \in \mathrm{GL}_{m_1}(R)$, $S_{\mathcal{B}'\mathcal{B}} \in \mathrm{GL}_{m_2}(R)$, hence:

$$[x]_{\mathcal{A}'} = S_{\mathcal{A}'\mathcal{A}}[x]_{\mathcal{A}} \ \ \forall\, x \in M_1, \qquad [y]_{\mathcal{B}'} = S_{\mathcal{B}'\mathcal{B}}[y]_{\mathcal{B}} \ \ \forall\, y \in M_2.$$

Proposition 5.9. *In the above notations, let* $\Psi_{\mathcal{A}\mathcal{B}}, \Psi_{\mathcal{A}'\mathcal{B}'} : \mathcal{L}(M_1 \times M_2, N) \to N^{m_1 \times m_2}$ *be defined* $\Psi_{\mathcal{A}\mathcal{B}}(f) := A_f$ *and* $\Psi_{\mathcal{A}'\mathcal{B}'}(f) := A'_f$*. Then setting* $S^{-\tau} := (S^\tau)^{-1}$ *the following hold:*

1) *The map* $\Psi_{\mathcal{A}\mathcal{B}}$ *is a isomorphism of* R*-modules. Further, if* $\mathcal{A}', \mathcal{B}'$ *are as above, one has*

$$\Psi_{\mathcal{A}'\mathcal{B}'}(f) = S^{-\tau}_{\mathcal{A}'\mathcal{A}} \Psi_{\mathcal{A}\mathcal{B}}(f) S_{\mathcal{B}\mathcal{B}'}, \ \text{i.e.,} \ \ A'_f = S^{-\tau}_{\mathcal{A}'\mathcal{A}} A_f S_{\mathcal{B}\mathcal{B}'} \ \forall\ f \in \mathcal{L}(M_1 \times M_2, N).$$

2) *Let* $M_1 = M = M_2$, $m_1 = m = m_2$, $\mathcal{A} = \mathcal{B}$, $\mathcal{A}' = \mathcal{B}'$, *and set* $\Psi_{\mathcal{A}} := \Psi_{\mathcal{A}\mathcal{A}}$, $\Psi_{\mathcal{A}'} := \Psi_{\mathcal{A}'\mathcal{A}'}$*.*
Then $\Psi_{\mathcal{A}'} = S^{-\tau}_{\mathcal{A}'\mathcal{A}} \Psi_{\mathcal{A}} S_{\mathcal{A}\mathcal{A}'}$ *and for* $A_f = \Psi_{\mathcal{A}}(f) = (x_{ij})_{i,j} \in N^{m \times m}$ *TFH:*
a) $f \in \mathcal{L}^2_{\mathrm{sym}}(M,N)$ *iff* $A_f \in N^{m \times m}$ *is symmetric, i.e.,* $x_{ij} = x_{ji}$ *for all* i, j*.*
b) $f \in \mathcal{L}^2_{\mathrm{alt}}(M,N)$ *iff* $A_f \in N^{m \times m}$ *is anti-symmetric, i.e.,* $x_{ij} = -x_{ji}$*,* $x_{ii} = 0_N$ *for all* i, j*.*

Proof. **Ex** (direct verifications...) □

Special subcase: *M_1, M_2 are free of finite rank, $N = R, +$*

Then in the above notations, given an R-bilinear map $f : M_1 \times M_2 \to R$, one has that $\Psi(f) = A_f := \big(f(\alpha_i, \beta_j)\big)_{i,j} \in R^{m_1 \times m_2}$, and

$$(*) \qquad f(x, y) = [x]_{\mathcal{A}}^{\tau} \cdot A_f \cdot [y]_{\mathcal{B}}$$

On the other hand, given any $A = (a_{ij})_{i,j} \in R^{m_1 \times m_2}$, the recipe $(*)$ given above, is R-linear in $\boldsymbol{x}, \boldsymbol{y}$. Indeed, if $x, x' \in M_1$, then $[x + x']_{\mathcal{A}}^{\tau} = [x]_{\mathcal{A}}^{\tau} + [x']_{\mathcal{A}}^{\tau}$, hence:

$$f(x + x', y) = [x + x']_{\mathcal{A}}^{\tau} \cdot A_f \cdot [y]_{\mathcal{B}} = [x]_{\mathcal{A}}^{\tau} \cdot A_f \cdot [y]_{\mathcal{B}} + [x']_{\mathcal{A}}^{\tau} \cdot A_f \cdot [y]_{\mathcal{B}} = f(x, y) + f(x', y)$$

and $[ax]_{\mathcal{A}}^{\tau} = a[x]_{\mathcal{A}}^{\tau}$ implies: $f(ax, y) = [ax]_{\mathcal{A}}^{\tau} A_f [y]_{\mathcal{A}}^{\tau} = a\,[x]_{\mathcal{A}}^{\tau} A_f [y]_{\mathcal{A}}^{\tau} = a\, f(x, y)$. Similarly,

$$f(x, y + y') = f(x, y) + f(x, y'), \quad f(x, by) = b\, f(x, y), \quad \forall\ y, y' \in M_2,\ b \in R.$$

Definition/Remark 5.10. Let R be a commutative ring with $1_r \neq 0_R$. One defines:

a) $A = (a_{ij})_{i,j} \in R^{m \times m}$ is symmetric, if $a_{ij} = a_{ji}\ \forall\, i, j$.

b) $A = (a_{ij})_{i,j} \in R^{m \times m}$ is anti-symmetric, or alternating, if $a_{ii} = 0_R$ and $a_{ij} = -a_{ji}\ \forall\, i, j$.

Recalling the standard basis $\mathcal{E} = (\boldsymbol{e}_{ij})_{i,j}$ of the free R-module $R^{m \times m}$, one has:

1) The set of symmetric matrices $R_{\mathrm{sym}}^{m \times m} \subset R^{m \times m}$ is a free R-submodule of rank $\frac{1}{2}m(m+1)$ and standard basis $\mathcal{E}_{\mathrm{sym}} = (\boldsymbol{e}_{ii}, \boldsymbol{e}_{ij} + \boldsymbol{e}_{ji})_{i.i<j}$. Further, $A \in R_{\mathrm{sym}}^{m \times m}$ iff $A^{\tau} \in R_{\mathrm{sym}}^{m \times m}$.

2) The set of annti-symmetric matrices $R_{\mathrm{alt}}^{m \times m} \subset R^{m \times m}$ is a free R-submodule of rank $\frac{1}{2}m(m-1)$ and standard basis $\mathcal{E}_{\mathrm{alt}} = (\boldsymbol{e}_{ij} - \boldsymbol{e}_{ji})_{i<j}$. Further, $A \in R_{\mathrm{alt}}^{m \times m}$ iff $A^{\tau} \in R_{\mathrm{alt}}^{m \times m}$.

3) Suppose that $\frac{1}{2} \in R$, i.e., $2 \cdot 1_R \in R^{\times}$. Then $\mathbf{0}_{m \times m}$ is the only matrix which is both symmetric and alternating. Further, for all $A \in R^{m \times m}$ one has:

   (i) $A_{\mathrm{sym}} := \frac{1}{2}(A + A^{\tau})$ is symmetric (WHY).

   (ii) $A_{\mathrm{alt}} := \frac{1}{2}(A - A^{\tau})$ is anti-symmetric (WHY).

   (iii) $A = A_{\mathrm{sym}} + A_{\mathrm{alt}}$ (WHY).

   Therefore, $R^{m \times m} = R_{\mathrm{sym}}^{m \times m} + R_{\mathrm{alt}}^{m \times m}$ is a direct sum of R-modules.

Proposition 5.11. *In the above notations, let $\Psi : \mathcal{L}(M_1 \times M_2) \to R^{m_1 \times m_2}$ by $\Psi(f) = A_f$, and set $A_f = (a_{ij})_{i,j} \in R^{m_1 \times m_2}$. Then the following hold:*

1) *$\mathcal{L}(M_1 \times M_2, R)$ is a free R-module having $(f_{kl})_{k,l}$ as a basis, where $A_{f_{kl}} = \boldsymbol{e}_{kl}$. Hence $\mathcal{L}(M_1 \times M_2, R)$ is free of rank $m_1 m_2$.*

2) *Let $M_1 = M = M_2$, $m_1 = m = m_2$, and $\mathcal{A} = \mathcal{B}$. Then TFH:*

   a) *$\Psi_{\mathcal{A}}(f) \in R_{\mathrm{sym}}^{m \times m}$ iff $f \in \mathcal{L}^2_{\mathrm{sym}}$, hence $\Psi_{\mathcal{A}} : \mathcal{L}^2_{\mathrm{sym}}(M, R) \to R_{\mathrm{sym}}^{m \times m}$ is an isomorphism of R-modules. Thus $\mathcal{L}^2(M, R)$ is free of rank $m(m+1)/2$, having as a basis $(f_{kl})_{k \leqslant l}$ with $A_{f_{kk}} = \boldsymbol{e}_{kk}$, and $A_{f_{kl}} = \boldsymbol{e}_{kl} + \boldsymbol{e}_{kl}$ if $k \neq l$.*

   b) *$\mathcal{L}^2_{\mathrm{alt}}(M, N)$ is a free R-module having $(f_{kl})_{k<l}$ as a basis, where $A_{f_{kk}} = \mathbf{0}_{m \times m}$ and $A_{f_{kl}} = \boldsymbol{e}_{kl} - \boldsymbol{e}_{lk}$ for $k < l$. Hence $\mathcal{L}^2_{\mathrm{sym}}(M, R)$ is a free R-module of rank $m(m-1)/2$.*

Proof. Exercise... □

Example 5.12. Let $\mathcal{E} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_m)$ be the standard basis of $M := R^m$, and for a bilinear form $f := R^m \times R^m \to R$, let $\Psi(f) = A_f = (a_{ij})_{i,j} := \big(f(\boldsymbol{e}_i, \boldsymbol{e}_j)\big)_{i,j} \in R^{m\times m}$ be its matrix. Then for $\boldsymbol{x} =: (x_1, \dots, x_m)$ and $\boldsymbol{y} := (y_1, \dots, y_m)$ from R^m one has:

$$f(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}\, A_f\, \boldsymbol{y}^\tau = \textstyle\sum_{i,j} x_i y_j a_{ij}$$

Hence $f(\boldsymbol{x}, \boldsymbol{y})$ is a polynomial function in the coordinates $\boldsymbol{x}, \boldsymbol{y}$ such that each variable has degree at most one. Further, f is symmetric iff $a_{ij} = a_{ji}\ \forall\, i, j$, and f is alternating iff ...

5.3. **Multilinear maps / Multilinear forms.**

Let $M_1, \dots, M_n, N$ be R-modules, and $f \in \mathrm{Maps}(M_1 \times \cdots \times M_n, N)$ be a fixed map. For fixed indices $\imath \in \{1, \dots, n\}$ and fixed $\boldsymbol{x}_\imath := (x_i)_{i\neq\imath} \in \prod_{i\neq\imath} M_i$, consider:

$$\mathcal{M}_{\boldsymbol{x}_\imath} = \{(z_i)_i \in \textstyle\prod_{i=1}^n M_i \mid z_i = x_i \ \text{ for } i \neq \imath \text{ and } z_i \in M_i \text{ arbitrary}\}.$$

Then the $\imath^{\text{th}}$ projection $pr_\imath : \mathcal{M}_{\boldsymbol{x}_\imath} \to M_\imath$, $(z_i)_i \mapsto z_\imath$ is a bijection (WHY). Let $\boldsymbol{\imath}_{\boldsymbol{x}_\imath} : M_\imath \to \mathcal{M}_{\boldsymbol{x}_\imath}$ be the inverse of map of $pr_\imath$, i.e., $pr_\imath\big(\boldsymbol{\imath}_{\boldsymbol{x}_\imath}(z_\imath)\big) = z_\imath$ for all $z_\imath \in M_\imath$.

Definition 5.13.

1) In the above notations, $f : M_1 \times \cdots \times M_n \to N$ is called multilinear, if for all $\imath = 1, \dots, n$ and all $\boldsymbol{x}_\imath$ the map $f \circ \imath_{\boldsymbol{x}_\imath} : M_\imath \to N$ is a morphism of R-modules.

Notation: $\mathcal{L}(M_1 \times \cdots \times M_n, N) := \{f \in \mathrm{Maps}(M_1 \times \cdots \times M_n, N) \,|\, f \ \text{ is } R\text{-multilinear}\}$. Further, if $M_1 = \cdots = M_n = M$ we denote: $\mathcal{L}^n(M, N) := \mathcal{L}(M^n, N)$.

2) $f \in \mathcal{L}^n(M, N)$ is called symmetric, if $f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = f(x_1, \dots, x_n)$ for all $\sigma \in S_n$.

Notation: $\mathcal{L}^n_{\mathrm{sym}}(M, N) = \{\, f \in \mathcal{L}^n(M, R) \mid f \text{ symmetric}\,\}$

3) $f \in \mathcal{L}^n(M, N)$ is called anti-symmetric, or alternating, if $f(x_1, \dots, x_n) = 0_N$, provided $x_i = x_j$ for some $i \neq j$, and $f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \epsilon(\sigma)\, f(x_1, \dots, x_n)$ for all $\sigma \in S_n$.

Notation: $\mathcal{L}^n_{\mathrm{alt}}(M, N) = \{\, f \in \mathcal{L}^n(M, R) \mid f \text{ alternating}\,\}$.

Proposition 5.14. (Properties) *In the above notations one has the following:*

1) $\mathcal{L}(M_1 \times \cdots \times M_n, N) \subset \mathrm{Maps}(M_1 \times \cdots \times M_n, N)$ *is an R-submodule.*
Furthermore, $\mathcal{L}^n_{\mathrm{sym}}(M, N), \mathcal{L}^n_{\mathrm{alt}}(M, N) \subseteq \mathcal{L}^n(M, N)$ are R-submodules for all M, N.
2) *Let $\phi : N \to P$ and $\psi_1 \times \ldots \times \psi_n : L_1 \times \ldots \times L_n \to M_1 \times \ldots \times M_n$ be R-morphisms. Then $\phi \circ f \circ \psi \in \mathcal{L}(L_1 \times \cdots \times L_n, P)$ for all $f \in \mathcal{L}(M_1 \times \cdots \times M_n, N)$.*
3) *Furthermore, if $L_i = L$, $M_i = M$, $\psi_i = \psi$ for $1 \leqslant i \leqslant n$, thus $\psi^n : L^n \to M^n$, then:*
   a) *$\phi \circ f \circ \psi^n \in \mathcal{L}^n_{\mathrm{sym}}(L,P)$ for $f \in \mathcal{L}^n_{\mathrm{sym}}(M,N)$.*
   b) *$\phi \circ f \circ \psi^n \in \mathcal{L}^n_{\mathrm{alt}}(L, P)$ for $f \in \mathcal{L}^n_{\mathrm{alt}}(M, N)$.*

Proof. Exercise (direct verification) ... □

Remark 5.15.
It is an interesting & important question, whether given $M_1, \dots, M_n$ as above, there exists an R-module $\mathcal{M}$ ***depending on*** $M_1, \dots, M_n$ ***only*** together with a ***R-multilinear map*** $\boldsymbol{\imath} : M_1 \times \cdots \times M_n \to \mathcal{M}$, such that the following holds:

For all R-modules N there exists an ***isomorphism*** *of R-modules*

$$\boldsymbol{\imath}_N : \mathrm{Hom}_R(\mathcal{M}, N) \to \mathcal{L}(M_1 \times \cdots \times M_n, N) \ \textit{defined by} \ \imath_N(\varphi) = \varphi \circ \boldsymbol{\imath}$$

It is not difficult to prove that such a module $\mathcal{M}$ and the multilinear map $\boldsymbol{\imath}$ exist indeed. The R-module $\mathcal{M}$ is denoted $M_1 \otimes_R \cdots \otimes_R M_n$ and called the tensor product of the R-modules $M_1, \ldots, M_n$. Further, $\boldsymbol{\imath} : M_1 \times \cdots \times M_n \to M_1 \otimes_R \cdots \otimes_R M_n$ is the universal multilinear map. If $M_1 = \cdots = M_n = M$, hence $M_1 \times \cdots \times M_n = M^n$, and one denotes $M \otimes_R \cdots \otimes_R M =: M^{\otimes n}$ and calls it the n^{th} tensor power of M. Hence in this case the universal multilinear map

$$\boldsymbol{\imath} : M^n \to M^{\otimes n}$$

gives rise for every R-module N to a bijection

$$\boldsymbol{\imath}_N : \mathrm{Hom}_R(M^{\otimes n}, N) \to \mathcal{L}^n(M, N), \ \ \textit{defined by} \ \ \imath_N(\varphi) = \varphi \circ \boldsymbol{\imath}\,.$$

Moreover, there are submodules $\mathcal{N}_{\mathrm{sym}} \subset M^{\otimes n}$ and $\mathcal{N}_{\mathrm{alt}} \subset M^{\otimes n}$ such that the canonical maps

$$\boldsymbol{\imath}_{\mathrm{sym}} : M^n \to M^{\otimes n}/\mathcal{N}_{\mathrm{sym}} =: \mathcal{M}_{\mathrm{sym}}, \quad \boldsymbol{\imath}_{\mathrm{alt}} : M^n \to M^{\otimes n}/\mathcal{N}_{\mathrm{alt}} =: \mathcal{M}_{\mathrm{alt}}$$

induced by $\boldsymbol{\imath} : M^n \to M^{\otimes n}$ satisfy: Let $f = \varphi \circ \boldsymbol{\imath} : M^n \xrightarrow{\boldsymbol{\imath}} M^{\otimes n} \xrightarrow{\varphi} N$ be the multilinear form defined by some $\varphi \in \mathrm{Hom}_R(M^{\otimes n}, N)$ as above. Then the fact that f is multilinear symmetric or alternating is encoded via φ and $\boldsymbol{\imath}_{\mathrm{sym}}$, respectively $\boldsymbol{\imath}_{\mathrm{alt}}$ as follows:

- $f \in \mathcal{L}^n_{\mathrm{sym}}(M, N)$ iff $\mathcal{N}_{\mathrm{sym}} \subset \mathrm{Ker}(\varphi)$ iff there is $\varphi_{\mathrm{sym}} : \mathcal{M}_{\mathrm{sym}} \to N$ s.t. $f = \varphi_{\mathrm{sym}} \circ \boldsymbol{\imath}_{\mathrm{sym}}$ (WHY).
- $f \in \mathcal{L}^n_{\mathrm{alt}}(M, N)$ iff $\mathcal{N}_{\mathrm{alt}} \subset \mathrm{Ker}(\varphi)$ iff there is $\varphi_{\mathrm{alt}} : \mathcal{M}_{\mathrm{alt}} \to N$ s.t. $f = \varphi_{\mathrm{alt}} \circ \boldsymbol{\imath}_{\mathrm{alt}}$ (WHY).

• **From now on**, we fix notations as follows:
- M as a finitely generated free R-module
- $\mathcal{A} = (\alpha_1, \ldots, \alpha_m)$ is a fixed basis of M
- Recall that for $n \in \mathbb{N}_{>0}$, we set $[n] := \{1, \ldots, n\}$.

• For every $n \geqslant 1$, consider the sets:
- $\mathcal{I} := \{\underline{\boldsymbol{\imath}} \mid \underline{\boldsymbol{\imath}} : [n] \to [m] \text{ map}\} = \{\,\underline{\boldsymbol{\imath}} = (i_1, \ldots, i_n) \,|\, 1 \leqslant i_1, \ldots, i_n \leqslant m\} = [m]^{[n]}$
- $\mathcal{X} := \{\alpha_{\underline{\boldsymbol{\imath}}} \,|\, \underline{\boldsymbol{\imath}} \in \mathcal{I}\,\} = \{\alpha_1, \ldots, \alpha_m\}^n \subset M^n$ with $\alpha_{\underline{\boldsymbol{\imath}}} := (\alpha_{i_1}, \ldots, \alpha_{i_n})$ for $\underline{\boldsymbol{\imath}} = (i_1, \ldots, i_n)$.
- $\mathcal{I}_{\leqslant} := \{\,\underline{\boldsymbol{\imath}} = (i_1, \ldots, i_n) \in \mathcal{I} \,|\, 1 \leqslant i_1 \leqslant \cdots \leqslant i_n \leqslant m\} \subset \mathcal{I}$
- $\mathcal{X}_{\leqslant} := \{\alpha_{\underline{\boldsymbol{\imath}}} \in \mathcal{X} \,|\, \underline{\boldsymbol{\imath}} \in \mathcal{I}_{\leqslant}\,\} \subset \mathcal{X}$.
- $\mathcal{I}_{<} := \{\underline{\boldsymbol{\imath}} = (i_1, \ldots, i_n) \in \mathcal{I} \,|\, 1 \leqslant i_1 < \cdots < i_n \leqslant m\} \subset \mathcal{I}$
- $\mathcal{X}_{<} := \{\alpha_{\underline{\boldsymbol{\imath}}} \in \mathcal{X} \,|\, \underline{\boldsymbol{\imath}} \in \mathcal{I}_{<}\,\} \subset \mathcal{X}$.

Remark 5.16. In the above notations, one has the following:

1) $|\mathcal{I}| = m^n$, $|\mathcal{I}_{\leqslant}| = \binom{n+m-1}{n}$, and $|\mathcal{I}_{<}| = \binom{m}{n}$. Hence $\mathcal{I}_{<} = \emptyset$ iff $\mathcal{X}_{<} = \emptyset$ iff $m < n$ (WHY).
2) For every $\sigma \in S_n$, let $\overline{\sigma} := \sigma^{-1}$ denote its inverse.
Every $\sigma \in S_n$ defines a bijection $\sigma : \mathcal{I} \to \mathcal{I}$, $\underline{\boldsymbol{\imath}} = (i_1, \ldots, i_n) \mapsto (i_{\overline{\sigma}(1)}, \ldots, i_{\overline{\sigma}(n)}) =: \sigma(\underline{\boldsymbol{\imath}})$.
3) For $x = (x_j)_j \in M^n$ and $\sigma \in S_n$, set $\sigma(x) := (x_{\overline{\sigma}(1)}, \ldots, x_{\overline{\sigma}(n)})$. Therefore one has:

$$(*)_\sigma \qquad \sigma(\alpha_{\underline{\boldsymbol{\imath}}}) = \sigma(\alpha_{i_1}, \ldots, \alpha_{i_n}) \overset{\text{def}}{:=} (\alpha_{i_{\overline{\sigma}(1)}}, \ldots, \alpha_{i_{\overline{\sigma}(n)}}) = \alpha_{\sigma(\underline{\boldsymbol{\imath}})} \quad \forall\ \alpha_{\underline{\boldsymbol{\imath}}} \in \mathcal{X} \text{ (WHY)}.$$

4) For $x = (x_j)_j$ with $x_j = \sum_{i=1}^m a_{ij}\alpha_i$, one has $\sigma(x) = (y_j)_j$ with $y_j = \sum_{i=1}^m a_{i\overline{\sigma}(j)}\alpha_i$ (WHY).
5) Given $x = (x_j)_j$ with $x_j = \sum_{i=1}^m a_{ij}\alpha_i$, for every $\underline{\boldsymbol{\imath}}$, set $\boldsymbol{a}_{\underline{\boldsymbol{\imath}}} := a_{i_1 1} \ldots a_{i_n n} = \prod_j a_{i_j j} \in R$.

Notice that $\sigma(r)=s$ iff $r=\overline{\sigma}(s)$, hence $\boldsymbol{a}_{\overline{\sigma}(\underline{\boldsymbol{\imath}})} := a_{i_{\sigma(1)}1}\dots a_{i_{\sigma(n)}n} = a_{i_1\overline{\sigma}(1)}\dots a_{i_n\overline{\sigma}(n)}$ (WHY).

Proposition 5.17. (Explicit Form) *In the above notations, for $f\in\mathcal{L}^n(M,N)$ one has:*

1) *$f: M^n\to N$ is given explicitly as follows:*

$$f(x_1,\dots,x_n)=\sum\nolimits_{\underline{\boldsymbol{\imath}}\in\mathcal{I}}\boldsymbol{a}_{\underline{\boldsymbol{\imath}}}f(\alpha_{\underline{\boldsymbol{\imath}}})=\sum\nolimits_{(i_1,\dots,i_n)\in\mathcal{I}}a_{i_11}\dots a_{i_nn}\,f(\alpha_{i_1},\dots,\alpha_{i_n}).$$

Therefore, $f: M^n\to N$ is completely determined by $f|_{\mathcal{X}}$, i.e., by $\{f(\alpha_{\underline{\boldsymbol{\imath}}})\mid\alpha_{\underline{\boldsymbol{\imath}}}\in\mathcal{X}\}$.

2) *For $\sigma\in S_n$ one has: $f\big(\sigma(x)\big)=\sum_{\underline{\boldsymbol{\imath}}\in\mathcal{I}}\boldsymbol{a}_{\underline{\boldsymbol{\imath}}}f(\alpha_{\sigma(\underline{\boldsymbol{\imath}})})$. In particular, the following hold:*

a) *$f\in\mathcal{L}^n_{\mathrm{sym}}(M,N)$ if an only if for all $\underline{\boldsymbol{\imath}}\in\mathcal{I}$, $\sigma\in S_n$ one has: $f(\alpha_{\underline{\boldsymbol{\imath}}})=f(\alpha_{\sigma(\underline{\boldsymbol{\imath}})})$.*

(∗) *In particular, f is completely determined by $f|_{\mathcal{X}_{\leqslant}}$, i.e., by $\{f(\alpha_{\underline{\boldsymbol{\imath}}})\mid\alpha_{\underline{\boldsymbol{\imath}}}\in\mathcal{X}_{\leqslant}\}$, because one has $f(\alpha_{\underline{\boldsymbol{\imath}}'})=f(\alpha_{\underline{\boldsymbol{\imath}}})$ for $\underline{\boldsymbol{\imath}}'=\sigma(\underline{\boldsymbol{\imath}})$ and all $\sigma\in S_n$, $\underline{\boldsymbol{\imath}}\in\mathcal{I}_{\leqslant}$.*

b) *$f\in\mathcal{L}^n_{\mathrm{alt}}(M,N)$ if and only if for all $\underline{\boldsymbol{\imath}}\in\mathcal{I}$ and $\sigma\in S_n$ one has: $f(\alpha_{\underline{\boldsymbol{\imath}}})=0$ for $\underline{\boldsymbol{\imath}}\notin\mathcal{I}_<$ and $f(\alpha_{\sigma(\underline{\boldsymbol{\imath}})})=\epsilon(\sigma)f(\alpha_{\underline{\boldsymbol{\imath}}})$ for $\underline{\boldsymbol{\imath}}\in\mathcal{I}_<$.*

(†) *In particular, f is completely determined by $f|_{\mathcal{X}_<}$, i.e., by $\{f(\alpha_{\underline{\boldsymbol{\imath}}})\mid\alpha_{\underline{\boldsymbol{\imath}}}\in\mathcal{X}_<\}$, because one has $f(\alpha_{\underline{\boldsymbol{\imath}}'})=f(\alpha_{\underline{\boldsymbol{\imath}}})$ for $\underline{\boldsymbol{\imath}}'=\sigma(\underline{\boldsymbol{\imath}})$ and all $\sigma\in S_n$, $\underline{\boldsymbol{\imath}}\in\mathcal{I}_<$.*

(‡) *Therefore, $\mathcal{L}^n_{\mathrm{alt}}(M,N)$ consists of the zero-map $0_{\mathcal{L}^n(M,N)}$, provided $m<n$.*

Proof. To 1): Exercise (induction on n).

To 2): Since $x_j=\sum_{i=1}^m a_{ij}\alpha_i$, by item 4) above one has that $x_{\sigma(j)}=\sum_{i=1}^m a_{i\overline{\sigma}(j)}\alpha_i$. Further, recall that for all $\sigma\in S_n$ one has that $a_{i_1\overline{\sigma}(1)}\dots a_{i_n\overline{\sigma}(n)}=\boldsymbol{a}_{\overline{\sigma}(\underline{\boldsymbol{\imath}})}$ and $\overline{\sigma}:\mathcal{I}\to\mathcal{I}$, $\boldsymbol{\imath}\mapsto\overline{\sigma}(\boldsymbol{\imath})$, is a bijection. Therefore one has:

$$f\big(\sigma(x)\big)=f(x_{\overline{\sigma}(1)},\dots,x_{\overline{\sigma}(n)})=\sum\nolimits_{\underline{\boldsymbol{\imath}}\in\mathcal{I}}a_{i_1\overline{\sigma}(1)}\dots a_{i_n\overline{\sigma}(n)}\,f(\alpha_{\underline{\boldsymbol{\imath}}})=\sum\nolimits_{\underline{\boldsymbol{\imath}}\in\mathcal{I}}\boldsymbol{a}_{\overline{\sigma}(\underline{\boldsymbol{\imath}})}f(\alpha_{\underline{\boldsymbol{\imath}}})=\sum\nolimits_{\underline{\boldsymbol{\imath}}\in\mathcal{I}}\boldsymbol{a}_{\underline{\boldsymbol{\imath}}}f(\alpha_{\sigma(\underline{\boldsymbol{\imath}})}).$$

To 2a) and 2b): Both assertions follow directly from the formula above and the definition of R-multilinear symmetric, respectively alternating, map (HOW). **(Fill in all the details!)** □

Multilinear forms

We finally discuss R-multilinear forms, i.e., that is the case when N is the R-module $R,+$.

Proposition 5.18. (Multilin) *In the above hypotheses, the following hold:*

1) *For every $\underline{\boldsymbol{\imath}}\in\mathcal{I}$ there exists a unique map $f_{\underline{\boldsymbol{\imath}}}\in\mathcal{L}^n(M,R)$ such that $f_{\underline{\boldsymbol{\imath}}}(\alpha_{\underline{\boldsymbol{\imath}}'})=\delta_{\underline{\boldsymbol{\imath}}\underline{\boldsymbol{\imath}}'}$, i.e.,*
   i) *$f_{\underline{\boldsymbol{\imath}}}(\alpha_{\underline{\boldsymbol{\imath}}})=1_R$.*
   ii) *$f_{\underline{\boldsymbol{\imath}}}(\alpha_{\underline{\boldsymbol{\imath}}'})=0_R$ for $\underline{\boldsymbol{\imath}}'\neq\underline{\boldsymbol{\imath}}$.*
2) *The system of maps $(f_{\underline{\boldsymbol{\imath}}})_{\underline{\boldsymbol{\imath}}\in\mathcal{I}}$ is an R-basis of $\mathcal{L}^n(M,R)$ with m^n entries.*
3) *$\phi:\mathcal{L}^n(M,R)\to\mathrm{Maps}(\mathcal{X},R)$ by $f\mapsto f|_{\mathcal{X}}$ is an isomorphism of R-modules.*

Proof. To 1): For and $\underline{\boldsymbol{\imath}}=(i_1,\dots,i_n)\in\mathcal{I}$, and any $x=(x_j)_j$, $x_j=\sum_{i=1}^m a_{ij}\alpha_i\in M$, define $\tilde{f}:M^n\to R$ by $\tilde{f}(x)=\boldsymbol{a}_{\underline{\boldsymbol{\imath}}}:=a_{i_11}\dots a_{i_nn}$. Let $x'_j=\sum_i a'_{ij}\alpha_i\in M$, $x''_j=\sum_i a''_{ij}\alpha_i\in M$ and $r\in R$ be given. Then one obviously has $x'_j+x''_j=\sum_i(a'_i+a''_i)\,\alpha_i$, $rx_j=\sum_i(ra_{ij})\,\alpha_i$, and therefore:

$$\tilde{f}(\dots,x'_j+x''_j,\dots)=a_{i_11}\dots(a'_{ij}+a''_{ij})\dots a_{i_nn}=a_{i_11}\dots a'_{ij}\dots a_{i_nn}+a_{i_11}\dots a''_{ij}\dots a_{i_nn}=\tilde{f}(\dots,x'_j,\dots)+\tilde{f}(\dots,x''_j,\dots)$$

and $\tilde{f}(\dots,rx_j,\dots)=a_{i_11}\dots(ra_{i_j})\dots a_{i_nn}=r\tilde{f}(\dots x_j\dots)$, showing that $\tilde{f}$ is R-multilinear form.

We next prove that $\tilde{f}(\alpha_{\underline{\boldsymbol{\imath}}'})=\delta_{\underline{\boldsymbol{\imath}}\underline{\boldsymbol{\imath}}'}$. Indeed, let $\alpha_{\underline{\boldsymbol{\imath}}'}=(\alpha_{i'_1},\dots,\alpha_{i'_n})\in\mathcal{X}$ be given. Then $x'_j:=\alpha_{i'_j}$ can be written uniquely in the form $x'_j=\sum_i a'_{ij}\alpha_i$ with $a'_{ij}=\sum_i\delta_{i,i'_j}\alpha_i$ (WHY). Hence one has:

$$\tilde{f}(\alpha_{\underline{\boldsymbol{\imath}}'})=\tilde{f}(x'_1,\dots,x'_n)\overset{\text{def}}{:=}a_{i_11}\dots a_{i_nn}\overset{\text{why}}{=}\delta_{i_1i'_1}\dots\delta_{i_ni'_n}\overset{\text{why}}{=}\delta_{\underline{\boldsymbol{\imath}}\underline{\boldsymbol{\imath}}'}.$$

Hence $f_{\underline{\imath}} := \tilde{f}$ is an R-multilinear form with the desired properties. Finally, $f_{\underline{\imath}}$ with the above properties is unique (WHY).

To 2): We first prove that the system $(f_{\underline{\imath}})_{\underline{\imath}\in\mathcal{I}}$ is R-free: Let $f = \sum_{\underline{\imath}\in\mathcal{I}} c_{\underline{\imath}} f_{\underline{\imath}}$ be the zero map in $\mathrm{Maps}(M^n, R)$. Sine $f_{\underline{\imath}}(\alpha_{\underline{\imath}'}) = \delta_{\underline{\imath}\underline{\imath}'} 1_R$ for all $\underline{\imath}, \underline{\imath}' \in \mathcal{I}$, for a fixed $\underline{\imath}' \in \mathcal{I}$ one has:

$$0 = f(\alpha_{\underline{\imath}'}) = \textstyle\sum_{\underline{\imath}\in\mathcal{I}} c_{\underline{\imath}} f_{\underline{\imath}}(\alpha_{\underline{\imath}'}) = \sum_{\underline{\imath}\in\mathcal{I}} c_{\underline{\imath}} \delta_{\underline{\imath}\underline{\imath}'} 1_R = c_{\underline{\imath}'}$$

and therefore, $\sum_{\underline{\imath}\in\mathcal{I}} c_{\underline{\imath}} f_{\underline{\imath}}$ is the trivial linear combination, as claimed. We next show that $(f_{\underline{\imath}})_{\underline{\imath}\in\mathcal{I}}$ is a system of generators: Given $f \in \mathcal{L}^n(M, R)$, let $c_{\underline{\imath}} := f(\alpha_{\underline{\imath}})$. Then setting $g := \sum_{\underline{\imath}\in\mathcal{I}} c_{\underline{\imath}} f_{\underline{\imath}}$, for every $\underline{\imath}' \in \mathcal{I}$ one has:

$$g(\alpha_{\underline{\imath}'}) = \textstyle\sum_{\underline{\imath}\in\mathcal{I}} c_{\underline{\imath}} f_{\underline{\imath}}(\alpha_{\underline{\imath}'}) = \sum_{\underline{\imath}\in\mathcal{I}} c_{\underline{\imath}} \delta_{\underline{\imath}\underline{\imath}'} 1_R = c_{\underline{\imath}'} = f(\alpha'_{\underline{\imath}})$$

which means that g and f are R-multilinear maps with $f(\alpha_{\underline{\imath}}) = g(\alpha_{\underline{\imath}})$ for all $\underline{\imath} \in \mathcal{I}$. But then $f = g$ (WHY). Conclude that $(f_{\underline{\imath}})_{\underline{\imath}\in\mathcal{I}}$ is a system of generators of $\mathcal{L}^n(M, R)$.

To 3): This is just a reformulation of assertions 1), 2) (WHY). □

Proposition 5.19. (Sym) *In the above hypotheses, the following hold:*

1) *For every $\underline{\imath} \in \mathcal{I}_{\leqslant}$ there exists a unique map $g_{\underline{\imath}} \in \mathcal{L}^n_{\mathrm{sym}}(M, R)$ satisfying the following:*
   i) *$g_{\underline{\imath}}(\alpha_{\sigma(\underline{\imath})}) = 1_R$ for all $\sigma \in S_n$.*
   ii) *$g_{\underline{\imath}}(\alpha_{\underline{\imath}'}) = 0_R$ provided $\underline{\imath}' \neq \sigma(\underline{\imath})$ for all $\sigma \in S_n$, or equivalently, $\underline{\imath}' \notin S_n(\underline{\imath})$.*
2) *The system of maps $(g_{\underline{\imath}})_{\underline{\imath}\in\mathcal{I}_{\leqslant}}$ is an R-basis of $\mathcal{L}^n_{\mathrm{sym}}(M, R)$ with $\binom{n+m-1}{n}$ elements.*
3) *$\phi_{\mathrm{sym}} : \mathcal{L}^n_{\mathrm{sym}}(M, R) \to \mathrm{Maps}(\mathcal{X}_{\leqslant}, R)$ by $f \mapsto f|_{\mathcal{X}_{\leqslant}}$ is an isomorphism of R-modules.*

Proof. To 1): For every $\underline{\imath} \in \mathcal{I}_{\leqslant}$, define $g'_{\underline{\imath}} : \mathcal{X} \to R$ by $g'_{\underline{\imath}}(\underline{\imath}') = 1_R$ if $\underline{\imath}' = \sigma(\underline{\imath})$ for some $\sigma \in S_n$ and $u_{\underline{\imath}'} = 0_R$ else. By Proposition (Multilin) above, there exists a unique R-multilinear form $g_{\underline{\imath}} : M^n \to R$ such that $(g_{\underline{\imath}})|_{\mathcal{X}} = g'_{\underline{\imath}}$. And notice that $g_{\underline{\imath}}(\alpha_{\underline{\imath}'}) = g'_{\underline{\imath}}(\alpha_{\underline{\imath}'}) = 1_R$ iff $\underline{\imath}' = \sigma(\underline{\imath})$ for some $\sigma \in S_n$ and $f_{\underline{\imath}}(\alpha_{\underline{\imath}'}) = g'_{\underline{\imath}}(\alpha_{\underline{\imath}'}) = 0_R$ else. Hence for an arbitrary $\underline{\imath}' \in \mathcal{I}$ and $\tau' \in S_n$, we have: First, $\tau'(\underline{\imath}') = \sigma(\underline{\imath})$ for some $\sigma \in S_n$ iff $\underline{\imath}' = (\tau'^{-1} \circ \sigma)(\underline{\imath})$ iff $\underline{\imath}' \in S_n(\underline{\imath})$ (WHY) iff $g_{\underline{\imath}}(\alpha_{\underline{\imath}'}) = 1_R$. And $\tau(\underline{\imath}') \neq \sigma(\underline{\imath})$ for all $\sigma \in S_n$ iff $\underline{\imath}' \notin S_n(\underline{\imath})$ iff $g_{\underline{\imath}}(\alpha_{\underline{\imath}'}) = 0_R$. Thus conclude that $g_{\underline{\imath}}\big(\sigma(\underline{\imath}')\big) = g_{\underline{\imath}}(\underline{\imath}')$ for all $\sigma \in S_n$ and all $\underline{\imath}' \in \mathcal{I}$ (WHY). But then by Proposition (Explicit Form), 2a), it follows that $g_{\underline{\imath}}$ is symmetric. For the uniqueness of $g_{\underline{\imath}}$ with the given properties i), ii), let $g \in \mathcal{L}^n(M, R)$ satisfy $g(\alpha_{\underline{\imath}'}) = 1_R$ if $\underline{\imath}' \in S_n(\underline{\imath})$ and $g(\alpha_{\underline{\imath}'}) = 0_R$ else. Then $(g_{\underline{\imath}} - g)(\alpha_{\underline{\imath}'}) = 0_R$ for all $\underline{\imath}' \in \mathcal{I}$ (WHY). Hence $g_{\underline{\imath}} = g$ (WHY).

To 2): We first prove that the system $(g_{\underline{\imath}})_{\underline{\imath}\in\mathcal{I}_{\leqslant}}$ is free. Indeed, let

$$f := \textstyle\sum_{\underline{\imath}\in\mathcal{I}_{\leqslant}} c_{\underline{\imath}} g_{\underline{\imath}} \quad c_{\underline{\imath}} \in R$$

be the zero map in $\mathrm{Maps}(M^n, R)$. Since $g_{\underline{\imath}}(\underline{\imath}') = \delta_{\underline{\imath}\underline{\imath}'} 1_R$ for all $\underline{\imath}, \underline{\imath}' \in \mathcal{I}_{\leqslant}$, for fixed $\underline{\imath}' \in \mathcal{I}_{\leqslant}$ we get:

$$0_R = f(\underline{\imath}') = \textstyle\sum_{\underline{\imath}\in\mathcal{I}_{\leqslant}} c_{\underline{\imath}} g_{\underline{\imath}}(\alpha_{\underline{\imath}'}) = \sum_{\underline{\imath}\in\mathcal{I}_{\leqslant}} c_{\underline{\imath}'} \delta_{\underline{\imath}\underline{\imath}'} = c_{\underline{\imath}'}.$$

This means simply that $c_{\underline{\imath}'} = 0_R$ for all $\underline{\imath}' \in \mathcal{I}_{\leqslant}$, thus $f := \sum_{\underline{\imath}\in\mathcal{I}_{\leqslant}} c_{\underline{\imath}} g_{\underline{\imath}}$ is the trivial linear combination.

Finally, $(g_{\underline{\imath}})_{\underline{\imath}\in\mathcal{I}_{\leqslant}}$ is a system of generators of $\mathcal{L}^n_{\mathrm{sym}}(M, R)$: Indeed, given any $f \in \mathcal{L}^n_{\mathrm{sym}}(M, R)$, by Proposition (Explicit Form), it follows that setting $c_{\underline{\imath}'} := f(\alpha_{\underline{\imath}'})$, one has $f = \sum_{\underline{\imath}'\in\mathcal{I}} c_{\underline{\imath}'} f_{\underline{\imath}'}$, where $(f_{\underline{\imath}'})_{\underline{\imath}'\in\mathcal{I}}$ is the basis of $\mathcal{L}^n(M, R)$ defined in Proposition (Multilin), i.e., $f_{\underline{\imath}'}(\alpha_{\underline{\imath}''}) = \delta_{\underline{\imath}'\underline{\imath}''} 1_R$ for all $\underline{\imath}', \underline{\imath}'' \in \mathcal{I}$. Since f is symmetric, by Proposition (Explicit Form), 2b), it follows that $c_{\sigma(\underline{\imath}')} = f(\alpha_{\sigma(\underline{\imath}')}) = f(\sigma_{\underline{\imath}'}) = c_{\underline{\imath}'}$ for all $\underline{\imath}' \in \mathcal{I}_{\leqslant}$ and $\sigma \in S_n$. Conclude that $f = \sum_{\underline{\imath}\in\mathcal{I}_{\leqslant}} c_{\underline{\imath}} g_{\underline{\imath}}$ (WHY), thus $(g_{\underline{\imath}})_{\underline{\imath}\in\mathcal{I}_{\leqslant}}$ is a system of generators of $\mathcal{L}^n_{\mathrm{sym}}(M, R)$. In order to conclude the proof of assertion 2), recall that $|\mathcal{I}_{\leqslant}| = \binom{n+m-1}{n}$ (WHY).

To 3): This is a reformulation of assertions 1), 2) above (WHY). □

Proposition 5.20. (Alt)
First, if $n > m$, then $\mathcal{L}^{\mathrm{alt}}(M^n, R) = \{0\}$. Second, if $n \leqslant m$, the following hold:

1) *For every $\underline{\imath} \in \mathcal{I}_<$ there exists a unique map $h_{\underline{\imath}} \in \mathcal{L}^n_{\mathrm{alt}}(M, R)$ which satisfies the following:*
   i) *$h_{\underline{\imath}}(\alpha_{\sigma(\underline{\imath})}) = \epsilon(\sigma) \cdot 1_R$ for all $\sigma \in S_n$.*
   ii) *$h_{\underline{\imath}}(\alpha_{\underline{\imath}'}) = 0_R$ provided $\underline{\imath}' \neq \sigma(\underline{\imath})$ for all $\sigma \in S_n$, or equivalently, $\underline{\imath}' \notin S_n(\underline{\imath})$.*
2) *The system of maps $(h_{\underline{\imath}})_{\underline{\imath} \in \mathcal{I}_<}$ is an R-basis of $\mathcal{L}^n_{\mathrm{alt}}(M, R)$ with $\binom{m}{n}$ entries.*
3) *$\phi_{\mathrm{alt}} : \mathcal{L}^n_{\mathrm{alt}}(M, R) \to \mathrm{Maps}(\mathcal{X}_<, R)$, $f \mapsto f|_{\mathcal{X}_<}$ is an isomorphism of R-modules.*

Proof. Let $f \in \mathcal{L}^n_{\mathrm{alt}}(M, R)$ be arbitrary, and let $f = \sum_{\underline{\imath}' \in \mathcal{I}} c_{\underline{\imath}'} f_{\underline{\imath}'}$ be the representation of f as a linear combination of the basis $(f_{\underline{\imath}'})_{\underline{\imath}' \in \mathcal{I}}$ given in Proposition (Multilin), 2). Recall that by loc.cit. we must have $c_{\underline{\imath}'} = f(\alpha_{\underline{\imath}'})$ for all $\underline{\imath}' \in \mathcal{I}$. First suppose that $n > m$. Since the R-basis $\mathcal{A} = (\alpha_1, \ldots, \alpha_m)$ has $m < n$ entries, in every $\alpha_{\underline{\imath}'} = (\alpha_{i'_1}, \ldots, \alpha_{i'_n})$ we must have $\alpha_{i'_k} = \alpha_{i'_l}$ for some $k < l$ (WHY). Since f is alternating, we must have $c_{\underline{\imath}'} = f(\alpha_{\underline{\imath}'}) = 0_R$ (WHY). Thus f is the trivial linear combination of the maps $(f_{\underline{\imath}'})_{\underline{\imath}'}$, thus it is the zero map (WHY). Conclude that $\mathcal{L}^n_{\mathrm{alt}}(M, R) = \{0\}$.

Now suppose that $n \leqslant m$. Proceed similarly to the proof of the previous Proposition (Sym) above, but starting as follows: For every $\underline{\imath} \in \mathcal{I}_<$ define $h_{\underline{\imath}} := \sum_{\sigma \in S_n} \epsilon(\sigma) f_{\sigma(\underline{\imath})}$, where $(f_{\underline{\imath}'})_{\underline{\imath}' \in \mathcal{I}}$ is the basis of $\mathcal{L}^n(M, R)$ defined in Proposition (Multilin). Then show that $(h_{\underline{\imath}})_{\underline{\imath} \in \mathcal{I}_<}$ is an R-basis of $\mathcal{L}^n_{\mathrm{alt}}(M, R)$, etc. Finally, to conclude, recall that $|\mathcal{I}_<| = \binom{m}{n}$ (WHY), etc. □

As a first corollary of Proposition (Alt), we have the following:

Theorem 5.21. *Let M be a finitely generated free R-module. Then all the R-bases of M have the same number of entries.*

Proof. Let $\mathcal{A} = (\alpha_1, \ldots, \alpha_m)$ and $\mathcal{B} = (\beta_1, \ldots, \beta_n)$ be R-bases. By contradiction, suppose that $m < n$. Then by Proposition (Alt) applied to M endowed with the basis $\mathcal{A}$ it follows that $\mathcal{L}^n_{\mathrm{alt}}(M, R) = \{0\}$, whereas working with the R-basis $\mathcal{B}$, it follows that $\mathcal{L}^n_{\mathrm{alt}}(M, R) \neq \{0\}$, because the latter has an R-basis with $1 = \binom{n}{n}$ elements. Contradiction! □

Definition 5.22. Let M be a finitely generated free R-module, and V be an F-vector space.
1) The cardinality of any basis of M is called the *rank* of M. **Notation:** $\mathrm{rk}_R(M)$ or $\mathrm{rk}(M)$.
2) The cardinality of a basis of V is the *dimension* of V. **Notation:** $\dim_F(V)$ or $\dim(V)$.

6. Determinants

6.1. Definitions, basic facts.

• We work in the following context/notations:

- M is a finitely generated free module of rank m
- $\mathcal{A} = (\alpha_1, \ldots, \alpha_m)$ is an R-basis of M.
- ${}^mR := R^{m \times 1}$ and $R^m := R^{1 \times m}$ have standard bases $\mathcal{E} := {}_m\mathcal{E}$, respectively $\mathcal{E}_m$.
- $R^m = ({}^mR)^\vee$ and ${}^mR = (R^m)^\vee$ canonically via the multiplication $R^m \times {}^mR \to R$ of matrices.
- In particular, $\mathcal{E}$ and $\mathcal{E}_m$ are dual basses to each other, i.e., $\mathcal{E}_m = \mathcal{E}^\vee$ and $\mathcal{E}_m^\vee = \mathcal{E}$.
- Finally, every $A \in R^{m \times m}$ gives rise to ${}_A\varphi \in \mathrm{End}_R({}^mR)$ and ${}_A\varphi^\vee \in \mathrm{End}_R(R^m)$ defined by

$$ {}_A\varphi(\boldsymbol{x}) = A\boldsymbol{x} \quad \text{and} \quad {}_A\varphi^\vee(\boldsymbol{x}^\tau) = \boldsymbol{x}^\tau A $$

• Recall that for $m = n$, the set $\mathcal{I}_<$ has just the element $\underline{\imath}_0 = (1,...,m)$. Thus every R-multilinear alternating map $f \in \mathcal{L}^m(M,R)$ is determined by $f(\alpha_1, \dots \alpha_m)$, because by Proposition (Alt) the map f satisfies $f(\alpha_{\sigma(1)}, \dots \alpha_{\sigma(m)}) = \epsilon(\sigma) f(\alpha_1, \dots \alpha_m)$ for all $\sigma \in S_n$.

Hence there exists a unique R-multilinear alternating map $f_{\mathcal{A}} : M^m \to R$ defined by

$$ (*) \qquad f_{\mathcal{A}}(\alpha_{\underline{\imath}}) = 0 \text{ if } \underline{\imath} \notin S_n(\underline{\imath}_0), \quad f_{\mathcal{A}}(\alpha_{\underline{\imath}}) = \epsilon(\sigma) 1_R \text{ if } \underline{\imath} = \sigma(\underline{\imath}_0) \text{ for all } \sigma \in S_n\,. $$

Proposition 6.1. *In the above notation, the following hold:*

1) *For every $f \in \mathcal{L}^m_{\mathrm{alt}}(M,R)$ there exists a unique $a_f \in R$ such that $f = a_f f_{\mathcal{A}}$. Moreover, one has $a_f = f(\alpha_1, \dots, \alpha_m)$.*
2) *For every $\varphi \in \mathrm{End}_R(M)$ there exits a unique $a_\varphi \in R$ such that $f_{\mathcal{A}} \circ \varphi^m = a_\varphi f_{\mathcal{A}}$. Moreover, one has that $a_\varphi = f_{\mathcal{A}}\big(\varphi(\alpha_1), \dots, \varphi_m(\alpha_m)\big)$.*

Proof. To 1): Since $\mathcal{I}_<$ has just the element $\underline{\imath}_0 = (1,...,n)$, by Proposition (Alt) one has that $\mathcal{L}^m_{\mathrm{alt}}(M,R)$ is a free R-module of rank 1 having $(f_{\mathcal{A}})$ with $f_{\mathcal{A}} = h_{\underline{\imath}_0}$ as basis (WHY). In particular, if $f_1 \in \mathcal{L}^m(M,R)$ is another basis, then $f_{\mathcal{A}} = a f_1$ with $a \in R$ as maps $M^m \to R$, thus and $a \in R^\times$ (WHY). Therefore, for every $f \in \mathcal{L}^m_{\mathrm{alt}}(M,R)$ there exists a unique $a_f \in R$ such that $f = a_f f_{\mathcal{A}}$ as maps $M^m \to R$. In order to compute a_f, one plugs $(\alpha_1, \dots, \alpha_m)$ in, and gets:

$$ a_f = a_f\, 1_R = a_f f_{\mathcal{A}}(\alpha_1, \dots, \alpha_m) = f(\alpha_1, \dots, \alpha_m). $$

To 2): Since $f_{\mathcal{A}} \in \mathcal{L}^m_{\mathrm{alt}}(M,R)$, it follows by Proposition (Properties), 3), b), that $f_\varphi := f_{\mathcal{A}} \circ \varphi^m$ lies in $\mathcal{L}^m_{\mathrm{alt}}(M,R) = R f_{\mathcal{A}}$ and therefore, there exists a unique $a_\varphi \in R$ such that $f_\varphi = f_{\mathcal{A}} \circ \varphi^m = a_\varphi f_{\mathcal{A}}$ as maps $M^m \to R$. OTOH, by definitions one has that $\varphi^m(\alpha_1, \dots, \alpha_m) = \big(\varphi(\alpha_1), \dots, \varphi(\alpha_m)\big)$ (WHY). Thus for a_φ, by assertion 1), one has: $a_\varphi = (f_{\mathcal{A}} \circ \varphi^m)(\alpha_1, \dots, \alpha_m) = f_{\mathcal{A}}\big(\varphi(\alpha_1), \dots, \varphi(\alpha_m)\big)$. □

Definition 6.2.

1) In the above notations, the element $a_\varphi \in R$ such that $f_{\mathcal{A}} \circ \varphi^m = a_\varphi f_{\mathcal{A}}$ is called the determinant of φ in the basis $\mathcal{A}$, and it is denoted $\det_{\mathcal{A}}(\varphi)$. The resulting map

$$ \det_{\mathcal{A}} : \mathrm{End}_R(M) \to R \text{ defined by } \varphi \mapsto \det_{\mathcal{A}}(\varphi) $$

is called the determinant map on $\mathrm{End}_R(M)$ in basis $\mathcal{A}$. Notice that one has:

$$ \det_{\mathcal{A}}(\varphi) = f_{\mathcal{A}}\big(\varphi(\alpha_1), \dots, \varphi(\alpha_m)\big). $$

2) Correspondingly, for $A \in R^{m\times m}$, we set $\det(A) := \det_{\mathcal{E}}({}_A\varphi)$, and call it the determinant of A, thus defining the determinant map

$$ \det : R^{m\times m} \to R \text{ defined by } A \mapsto \det(A). $$

Notice that one has: $\det(A) = f_{\mathcal{E}}(x_1, \dots, x_m)$, where $x_j := {}_A\varphi(e_j) = Ae_j = \boldsymbol{c}_j$ is the j^{th} column of A.

Remarks 6.3. The determinant $\det : R^{m\times m} \to R$ is the unique map which satisfies (WHY):

i) $\det(\cdot)$ is R-multilinear in the columns (WHAT DOES THAT MEAN?).
ii) $\det(A) = 0_R$ if A has two proportional/identical columns.
iii) $\det(\boldsymbol{I}_m) = 1_R$.

Exercise. Make sure that you understand/know why the above remarks hold.

6.2. **Facts/Properties of Determinants.**

1) Expansion formulas:

Let $\varphi \in \mathrm{End}_R(M)$, $A := [\varphi]_{\mathcal{A}} = (a_{ij})_{i,j} \in R^{m\times m}$ be its matrix in basis $\mathcal{A}$. Then one has:

a) $\det_{\mathcal{A}}(\varphi) = \sum_{\sigma\in S_m} \epsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(m)m}$

b) $\det(A) = \sum_{\sigma\in S_m} \epsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(m)m} = \sum_{\sigma\in Sm} \epsilon(\sigma) a_{1\sigma(1)} \dots a_{m\sigma(m)} = \det(A^\tau)$

Proof. Since $\varphi(\alpha_j) = \sum_i a_{ij}\alpha_j$, one has: $\det_{\mathcal{A}}(\varphi) = f_{\mathcal{A}}\big((\varphi(\alpha_1),\dots,\varphi(\alpha_m)\big) = f_{\mathcal{A}}\big((\sum_i a_{ij}\alpha_i)_j\big) = \sum_{\underline{\imath}} a_{\underline{\imath}} f_{\mathcal{A}}(\alpha_{\underline{\imath}})$. Since by $(*)$ above, $f_{\mathcal{A}}(\alpha_{\underline{\imath}}) = 0$ for $\underline{\imath} \notin S_m(\underline{\imath}_0)$ and $f_{\mathcal{A}}(\alpha_{\underline{\imath}}) = \epsilon(\sigma)1_R$ for $\underline{\imath} = \sigma(\underline{\imath}_0)$ for $\sigma \in S_m$, one gets:

$$\det_{\mathcal{A}}(\varphi) = \sum_{\sigma\in S_m} a_{\sigma(\underline{\imath}_0)}\epsilon(\sigma)1_R = \sum_{\sigma\in S_m} \epsilon(\sigma) a_{\sigma(1)1}\dots a_{\sigma(m)m}.$$

Similarly, $\det(A) :\overset{\text{def}}{=} \det_{\mathcal{E}}({}_A\varphi) = f_{\mathcal{E}}\big((\boldsymbol{c}_j)_j\big) = f_{\mathcal{E}}\big((\sum_i a_{ij}e_i)_j\big) = \sum_{\underline{\imath}} a_{\underline{\imath}} f_{\mathcal{E}}(e_{\underline{\imath}})$ (WHY). Hence reasoning as above, $\det(A) = \det_{\mathcal{E}}({}_A\varphi) = \sum_{\sigma\in S_m} a_{\sigma(\underline{\imath}_0)}\epsilon(\sigma)1_R = \sum_{\sigma\in S_m} \epsilon(\sigma) a_{\sigma(1)1}\dots a_{\sigma(m)m}$. Further, for all $\sigma \in S_m$ one has $a_{\sigma(1)1}\dots a_{\sigma(m)m} = a_{1\overline{\sigma}(1)}\dots a_{m\overline{\sigma}(m)}$ (WHY) and second, $A^\tau = (a^\tau_{ij})_{i,j}$ with $a^\tau_{ij} = a_{ji}$. Hence since $\epsilon(\sigma) = \epsilon(\overline{\sigma})$, one gets $\epsilon(\sigma)a_{\sigma(1)1}\dots a_{\sigma(m)m} = \epsilon(\overline{\sigma})a_{1\overline{\sigma}(1)}\dots a_{m\overline{\sigma}(m)} = \epsilon(\overline{\sigma})a^\tau_{\overline{\sigma}(1)1}\dots a^\tau_{\overline{\sigma}(m)m}$ (WHY). OTOH, for a sum $\sum_{g\in G} x_g$ indexed by $g \in G$, G finite, one has $\sum_{g\in G} x_g = \sum_{g\in G} x_{g^{-1}}$ (WHY). Since $\overline{\sigma} = \sigma^{-1}$, conclude:

$$\det(A) = \sum_{\sigma\in S_m} \epsilon(\sigma) a_{\sigma(1)1}\dots a_{\sigma(m)m} = \sum_{\sigma\in S_m} \epsilon(\sigma) a^\tau_{\sigma(1)1}\dots a^\tau_{\sigma(m)m} = \det(A^\tau).$$

□

Remarks 6.4. Taking into account the Remark 6.3 above and that $\det(A) = \det(A^\tau)$, it follows that $\det : R^{m\times m} \to R$ is the unique map which satisfies (WHY):

i) $\det(\cdot)$ is R-multilinear in the rows (WHAT DOES THAT MEAN?).

ii) $\det(A) = 0_R$ if A has two proportional/identical rows.

iii) $\det(\boldsymbol{I}_m) = 1_R$.

2) Multiplicativity:

$\det_{\mathcal{A}} : \mathrm{End}_R(M) \to R$ and $\det : R^{m\times m} \to R$ are multiplicative, i.e., one has:

$$\det_{\mathcal{A}}(\varphi\circ\psi) = \det_{\mathcal{A}}(\varphi)\det_{\mathcal{A}}(\psi), \qquad \det(AB) = \det(A)\det(B)$$

Therefore, one has the following:

a) $\det_{\mathcal{A}}$ is a *morphism of monoids* from the monoid $\mathrm{End}_R(M), \circ$ to the monoid $R, \cdot$

b) $\det : R^{m\times m} \to R$ is a *morphism of monoids* from the monoid $R^{m\times m}, \cdot$ to $R, \cdot$.

Proof. Let $\varphi, \psi \in \mathrm{End}_R(M)$ be given. Then $\varphi\circ\psi \in \mathrm{End}_R(M)$, and $(\varphi\circ\psi)^m = \varphi^n \circ \psi^n$ on M^m (WHY). OTOH, $\det_{\mathcal{A}}(\varphi)f_{\mathcal{A}} = f_{\mathcal{A}}\circ\varphi^n$, $\det_{\mathcal{A}}(\psi)f_{\mathcal{A}} = f_{\mathcal{A}}\circ\psi^n$ and $\det_{\mathcal{A}}(\varphi\circ\psi)f_{\mathcal{A}} = f_{\mathcal{A}}\circ(\varphi\circ\psi)^n$ on M^m, and therefore:

$$\begin{aligned}\det_{\mathcal{A}}(\varphi\circ\psi)f_{\mathcal{A}} &= f_{\mathcal{A}}\circ(\varphi\circ\psi)^m = (f_{\mathcal{A}}\circ\varphi^m)\circ\psi^n = \\ &= \big(\det(\varphi)f_{\mathcal{A}}\big)\circ\psi^n = \det(\varphi)(f_{\mathcal{A}}\circ\psi^m) = \det_{\mathcal{A}}(\varphi)\det_{\mathcal{A}}(\psi)f_{\mathcal{A}},\end{aligned}$$

from which we conclude that $\det_{\mathcal{A}}(\varphi\circ\psi) = \det_{\mathcal{A}}(\varphi)\det_{\mathcal{A}}(\psi)$ (WHY), as claimed. □

As an essential consequence of multiplicativity of $\det_{\mathcal{A}}$ and expansion formulas for $\det_{\mathcal{A}}$ one has the following.

Proposition 6.5. *The maps* $\det_{\mathcal{A}} : \mathrm{End}_R(M) \to R$ *and* $\det : R^{m\times m} \to R$ *satisfy:*

1) *Let* $\varphi \in \mathrm{End}_R(M)$ *be an isomorphism. Then* $\det_{\mathcal{A}}(\varphi) \in R^\times$ *and* $\det_{\mathcal{A}}(\varphi^{-1}) = \big(\det_{\mathcal{A}}(\varphi)\big)^{-1}$. *In particular, if* $A \in \mathrm{GL}_m(R)$, *then* $\det(A) \in R^\times$ *and* $\det(A^{-1}) = \big(\det(A)\big)^{-1}$.

2) $\det_{\mathcal{A}}(\varphi)$ *is independent of the concrete basis* $\mathcal{A}$*, i.e., if* $\mathcal{A}'$ *is another basis of* M*, then one has that* $\det_{\mathcal{A}}(\varphi) = \det_{\mathcal{A}'}(\varphi)$.

Notation. *One denotes* $\det_{\mathcal{A}}(\varphi)$ *simply by* $\det(\varphi)$.

Proof. To 1): Since the identity map id_M has matrix $[\mathrm{id}_M]_{\mathcal{A}} = \boldsymbol{I}_m$, one has $\det_{\mathcal{A}}(\mathrm{id}_M) = \det(\boldsymbol{I}_m) = 1$ (WHY). Hence by the multiplicativity of $\det_{\mathcal{A}}$, one gets:

$$1 = \det_{\mathcal{A}}(\mathrm{id}_M) = \det_{\mathcal{A}}(\varphi \circ \varphi^{-1}) = \det_{\mathcal{A}}(\varphi)\det_{\mathcal{A}}(\varphi^{-1}).$$

In particular, $\det_{\mathcal{A}}(\varphi) \in R$ is invertible and its inverse is $\det_{\mathcal{A}}(\varphi^{-1})$, and which proves the first assertion.

To 2): Recalling that $[\varphi]_{\mathcal{A}'} = S_{\mathcal{A}'\mathcal{A}}[\varphi]_{\mathcal{A}}S_{\mathcal{A}'\mathcal{A}}^{-1}$, using the assertion 1) of the expansion formulas, one has:

$$\det_{\mathcal{A}'}(\varphi) \overset{\text{why}}{=} \det([\varphi]_{\mathcal{A}'}) = \det\left(S_{\mathcal{A}'\mathcal{A}}[\varphi]_{\mathcal{A}}S_{\mathcal{A}'\mathcal{A}}^{-1}\right) \overset{\text{why}}{=} \det(S_{\mathcal{A}'\mathcal{A}})\det([\varphi]_{\mathcal{A}})\det(S_{\mathcal{A}'\mathcal{A}}^{-1}) \overset{\text{why}}{=} \det([\varphi]_{\mathcal{A}}) \overset{\text{why}}{=} \det_{\mathcal{A}}(\varphi).$$

□

Minors and cofactors

Let $A = (a_{ij})_{i,j} \in R^{m\times m}$ be given. For every $1 \leqslant k, l \leqslant m$ we consider the $(m\text{-}1) \times (m\text{-}1)$ matrix A_{kl} obtained from A by erasing the k^{th} row and the l^{th} column of A. The determinant $\Delta_{kl} := \det(A_{kl})$ is called the kl-**minor** of A, and $(-1)^{k+l}\Delta_{kl}$ is the **cofactor** of a_{kl}.

3) Column cofactor expansion / Row cofactor expansion:

a) For $A = (a_{ij})_{i,j} \in R^{m\times m}$, let Δ_{ij} be its minors. Then for all $k = 1, \ldots, m$ one has:

$$\sum_{i=1}^{m}(-1)^{i+k}a_{ik}\Delta_{ik} = \det(A) = \sum_{i=1}^{m}(-1)^{k+i}a_{ki}\Delta_{ki}$$

b) Moreover, for all $k, l = 1, \ldots, m$ with $k \neq l$ one has:

$$\sum_{i=1}^{m}(-1)^{i+k}a_{il}\Delta_{ik} = 0_R = \sum_{i=1}^{m}(-1)^{i+k}a_{li}\Delta_{ki}$$

Proof. To a): Begin by noticing that for all $A = (a_{ij})_{i,j} \in R^{m\times m}$ and all $1 \leqslant k, l \leqslant m$ one has $(A_{kl})^{\tau} = A_{lk}^{\tau}$ (WHY). Hence the cofactor $(-1)^{k+l}\Delta_{kl}$ of the coefficient a_{kl} of A equals the cofactor $(-1)^{l+k}\Delta_{lk}^{\tau}$ of the coefficient a_{lk} of A^{τ} (WHY). Therefore, it is sufficient to prove either the r.c.e. or the c.c.e. holds. We prove that the latter holds, i.e., that for all $A \in R^{m\times m}$ and $k = 1, \ldots, m$ the following holds:

$$(*)_{A,k} \qquad\qquad \det(A) = \textstyle\sum_{i=1}^{m}(-1)^{i+k}a_{ik}\Delta_{ik}$$

Claim 1. $(*)_{A,k}$ *holds for all* $k = 1,\ldots,m$*,* $A \in R^{m\times m}$ *iff* $(*)_{A,m}$ *holds for all* $A \in R^{m\times m}$.

Indeed, we make induction on $n := m - k$. If $n = 0$, there is nothing to prove (WHY). If $n > 0$, that is, $k < m$, let $A' := (a'_{ij})_{i,j}$ be the matrix obtained from A by interchanging the columns $\boldsymbol{c}_k$ and $\boldsymbol{c}_{k+1}$ (e.g. if $k = m-1$, thus $n = 1$, then interchange $\boldsymbol{c}_{m-1}$ and $\boldsymbol{c}_m$). Then $a'_{ij} = a_{ij}$ for $j \neq k, k+1$ (WHY) and $a'_{ik} = a_{i\,k+1}$, $a'_{i\,k+1} = a_{ik}$ (WHY). Hence $A_{ik} = A'_{i\,k+1}$ (WHY), thus $\Delta_{ik} = \Delta'_{i\,k+1}$ for all $i = 1, \ldots, m$. Therefore one has:

$$\textstyle\sum_{i=1}^{m}(-1)^{i+k}a_{ik}\Delta_{ik} = \sum_{i=1}^{m}(-1)^{i+k}a'_{i\,k+1}\Delta'_{i\,k+1} = -\sum_{i=1}^{m}(-1)^{i+k+1}a'_{i\,k+1}\Delta'_{i\,k+1}.$$

Further, since the determinant is alternating, one has $\det(A') = -\det(A)$ (WHY). Hence we conclude that $(*)_{A,k}$ holds iff $(*)_{A',k+1}$ holds (WHY). Hence setting $A_n := A$ and $A_{n-1} := A'$, conclude by induction on $n > 0$ that $(*)_{A_n,k}$ holds for all $A_n \in R^{m\times m}$ iff $(*)_{A_1,m}$ holds for all $A_1 \in R^{m\times m}$ (WHY). The Claim 1 is proved.

Let $A = (a_{ij})_{i,j} = \left(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_m\right) \in R^{m\times m}$ have as m^{th} column $\boldsymbol{c}_m = \sum_i a_{im}\boldsymbol{e}_i$, where $(\boldsymbol{e}_i)_{1\leqslant i\leqslant m}$ is the standard basis of ${}^mR$. Since $\det(A)$ is multilinear alternating in each column, thus in the column $\boldsymbol{c}_m$ as well, one has that $\det(A) = \sum_{i=1}^{m} a_{im}\det\left(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_{m-1}, \boldsymbol{e}_i\right)$. Hence the m^{th}-row cofactor expansion formula for all matrices $A \in R^{m\times m}$ follows (WHY) from the following:

Claim 2. If $B = (b_{ij})_{i,j} \in R^{m\times m}$ has $\boldsymbol{e}_k$ as its m^{th} column, then $\det(B) = (-1)^{k+m}\Delta_{km}$.

First, setting $n = m - k$, an argument by induction on n as in the proof of Claim 1 (HOW), reduces the problem to the verification step, i.e., to considering the case $n = 0$, i.e., $k = m$ (WHY). Finally, for $k = m$,

setting $B = (b_{ij})_{i,j}$, by the expansion formula one has $\det(B) = \sum_{\sigma\in S_m} \epsilon(\sigma) b_{\sigma(1)1} \dots b_{\sigma(m)m}$. OTOH, since the m^{th} column $\boldsymbol{c}_m = (b_{im})_i$ of B equals $\boldsymbol{e}_m \in {}^mR$, one has $b_{im} = \delta_{im}$, hence $b_{\sigma(m)m} = \delta_{m\sigma(m)}$ (WHY). In particular, setting $\tilde{S}_m := \{\sigma \in S_m \,|\, \sigma(m) = m\}$, one has: First, $b_{\sigma(m)m} \neq 0_R$ iff $\sigma \in \tilde{S}_m$ iff $b_{\sigma(m)m} = 1_R$, and second, the map $\tilde{S}_m \to S_{m-1}$, $\sigma \mapsto \sigma|_{[m-1]}$ is an isomorphism of groups (WHY). Therefore, one has:

$$\det(B) \overset{\text{why}}{=} \sum_{\sigma\in\tilde{S}_m} \epsilon(\sigma) b_{\sigma(1)1} \dots b_{\sigma(m-1)\,m-1} \overset{\text{why}}{=} \sum_{\tau\in S_{m-1}} \epsilon(\tau) a_{\tau(1)1} \dots b_{\tau(m-1)\,m-1} \overset{\text{why}}{=} (-1)^{m+m} \Delta_{mm}.$$

To b): For the row case, consider the matrix $A' = (a'_{ij})_{i,j}$ obtained form $A = (a_{ij})_{i,j}$ by replacing the k^{th} row $\boldsymbol{\mathcal{R}}_k$ by the l^{th} row $\boldsymbol{\mathcal{R}}_l$ row and keeping $\boldsymbol{\mathcal{R}}_i$ in place for all $i \neq k, l$, i.e., $a'_{ij} = a_{ij}$ for $i \neq l, k$ and $a'_{kj} = a_{lj} = a'_{lj}$ for $j = 1, \dots, m$. Then $\boldsymbol{\mathcal{R}}'_k = \boldsymbol{\mathcal{R}}_l$ implies $\det(A') = 0_R$ (WHY), and further one has: $A'_{kj} = A_{kj}$ for all $j = 1, \dots, m$ (WHY). Therefore, the minors Δ'_{kj} of A', and Δ_{kj} of A satisfy: $\Delta'_{kj} = \Delta_{kj}$ (WHY). The cofactor expansion of $\det(A')$ gives:

$$0_R = \det(A') = \sum_{j=1}^m (-1)^{k+j} a'_{kj} \Delta'_{kj} = \sum_{j=1}^m (-1)^{j+k} a_{lj} \Delta_{kj}.$$

The proof of the column expansion is absolutely similar, but working with columns instead of rows. □

4) The classical adjoint/adjugate:

Definition 6.6. Let $A = (a_{ij})_{i,j} \in R^{m\times m}$ have minors Δ_{ij}. The matrix

$$A^* := \big((-1)^{i+j}\Delta_{ji}\big)_{i,j} \in R^{m\times m}$$

is called the (***classical***) ***adjoint*** or the (***classical***) ***adjugate*** of A.

Theorem 6.7. *In the above notations, the following hold:*

1) $A^*A = \det(A)\boldsymbol{I}_m = A\,A^*$
2) *A is invertible iff* $\det(A) \in R^\times$. *If so, then* $A^{-1} = \det(A)^{-1}A^*$.

Proof. To a): Since $A^* := \big((-1)^{i+j}\Delta_{ji}\big)_{i,j}$, it assertion a) is equivalent to:

$$\sum_{j=1}^m (-1)^{k+i} a_{kj}\Delta_{lj} = \det(A)\delta_{kl}, \quad \sum_{i=1}^m (-1)^{k+i} a_{ik}\Delta_{li} = \det(A)\delta_{kl} \text{ for } 1 \leqslant k, l \leqslant m.$$

This follows instantly from the row/column expansion above. □

6.3. The Cayley–Hamilton Theorem.

Recall the usual notations:

- R is a commutative ring with 1_R.
- M is an R-free module, $\mathcal{A} = (\alpha_1, \dots, \alpha_m)$ and R-basis of M.
- $R^{m\times m}$ and $\mathrm{End}_R(M)$ the corresponding R-algebras.
- We set $\tilde{R} := R[t]$ the ring of polynomials in t over R.

Note: The inclusion $R \hookrightarrow \tilde{R}$ gives rise to an inclusion of rings $R^{m\times m} \hookrightarrow \tilde{R}^{m\times m}$.

Definition 6.8.

1) Let $A \in R^{m\times m}$ be a fixed matrix, and consider $\tilde{A} := t\boldsymbol{I}_m - A \in \tilde{R}^{m\times m}$. The polynomial $P_A(t) := \det(\tilde{A}) = \det(t\boldsymbol{I}_m - A) \in R[t] = \tilde{R}$ is called the characteristic polynomial of A.
2) For $\varphi \in \mathrm{End}(M)$, let A_φ be its matrix in the basis $\mathcal{A}$. We set $P_\varphi(t) := P_{A_\varphi}(t)$, and call it the characteristic polynomial of φ.

Note: $P_\varphi(t)$ does not depend on the basis $\mathcal{A}$. Indeed, if $\mathcal{A}'$ is another basis, and A'_φ is the matrix of φ in the basis $\mathcal{A}'$, then $A_\varphi = S^{-1}A'_\varphi S$, where $S := S_{\mathcal{A}'\mathcal{A}}$ the base change matrix. Hence one has the following:

$$\tilde{A}_\varphi = t\boldsymbol{I}_m - A_\varphi = t\boldsymbol{I}_m - S^{-1}A'_\varphi S = S^{-1}(t\boldsymbol{I}_m - A'_\varphi)S = S^{-1}\tilde{A}'_\varphi S,$$

and therefore: $P_\varphi(t) := \det(\tilde{A}_\varphi) = \det(S^{-1})\det(\tilde{A}'_\varphi)\det(S) = \det(\tilde{A}'_\varphi)$.

Construction: Consider the following maps:

- $\phi_A : R[t] \to R^{m\times m}$, $p(t) \mapsto p(A)$, is the evaluation morphism at $t \mapsto A$.
- $*$ Make sure to **check/know** that ϕ_A is in fact a ring homomorphism, and even more:

$$\phi_A\big(r\,p(t)\big) = r\phi_A\big(p(t)\big) \text{ for all } r \in R,\ p(t) \in R[t] \text{ (WHY)}.$$

- **Example:** $\phi_A(1_{R[t]}) = \boldsymbol{I}_m$; $\phi_A(a+bt) = a\boldsymbol{I}_m + bA$, etc.

- Let $\phi_\varphi : R[t] \to \mathrm{End}_R(M)$, $p(t) \mapsto p(\varphi)$, i.e., ϕ_φ is the evaluation morphism at $t \mapsto \varphi$.
- $*$ Make sure to **check/know** that ϕ_φ is in fact a ring homomorphism, and even more:

$$\phi_\varphi\big(r\,p(t)\big) = r\phi_\varphi\big(p(t)\big) \text{ for all } r \in R,\ p(t) \in R[t] \text{ (WHY)}.$$

- **Example:** $\phi_\varphi(1_{R[t]}) = \mathrm{id}_M$; $\phi_\varphi(a+bt) = a\,\mathrm{id}_M + b\varphi$, etc.

- Define an *outer multiplication* of $R[t]$ on M by $p(t)\cdot x := p(\varphi)(x)$, for all $x \in M$.
- **Example:** $1_{R[t]}\cdot x = \mathrm{id}_M(x) = x$; $(a+bt)\cdot x = ax + b\varphi(x)$ (WHY), etc.

Proposition 6.9. *The above outer multiplication makes M into an $R[t]$-module.*

Proof. **Ex** . . . □

In the above notation and situation, one of the fundamental and most important facts concerning the characteristic polynomial $P_\varphi(t) \in R[t]$ of an endomrphism $\varphi \in \mathrm{End}_R(M)$, respectively the characteristic polynomial of a matrix $A \in R^{m\times m}$ is the following.

Theorem 6.10. (Cayley–Hamilton Thm)
In the above notation, $P_A(A) = \mathbf{0}_{m\times m}$ in $R^{m\times m}$ and $P_\varphi(\varphi) = 0_{\mathrm{End}_R(M)}$ in $\mathrm{End}_R(M)$.

Proof. Let $\mathcal{A} = (\alpha_1, \ldots, \alpha_m)$ be a fixed basis of M, and $A_\varphi = (a_{ij})_{i,j} \in R^{m\times m}$ be the matrix of the endomorphism $\varphi : M \to M$ in the basis $\mathcal{A}$. Then one has the following description of the outer multiplication: $t\cdot\alpha_j = \varphi(\alpha_j) = \sum_{i=1}^m a_{ij}\alpha_i$ for all $j = 1, \ldots, m$ (WHY). Hence one has:

$$t\cdot\mathcal{A} = \varphi(\mathcal{A}) = \mathcal{A}A_\varphi \text{ (WHY)}.$$

Hence recalling that $\tilde{A}_\varphi := t\boldsymbol{I}_m - A_\varphi \in R[t]^{m\times m}$, we have:

$$\mathcal{A}\,\tilde{A}_\varphi = \mathcal{A}\,(t\boldsymbol{I}_m - A_\varphi) = t\cdot\mathcal{A} - \mathcal{A}A_\varphi = (0_M, \ldots, 0_M) \text{ (WHY)}.$$

Multiplying on the right by the (canonical) adjoint $\tilde{A}^*_\varphi \in R[t]^{m\times m}$ of $\tilde{A}_\varphi$, we get:

$$(0_M, \ldots, 0_M) = (0_M, \ldots, 0_M)\tilde{A}^*_\varphi = \mathcal{A}\,\tilde{A}_\varphi\,\tilde{A}^*_\varphi = \mathcal{A}\,\det(\tilde{A}_\varphi)\boldsymbol{I}_m = \det(\tilde{A}_\varphi)\cdot\mathcal{A} \text{ (WHY)}.$$

Thus since by definition we have $\det(\tilde{A}_\varphi) = P_\varphi(t)$, it follows that

$$P_\varphi(t)(\alpha_1, \ldots, \alpha_m) = (0_M, \ldots, 0_M),$$

i.e., $P_\varphi(t)\alpha_j = 0_M$ for all $j = 1, \ldots, m$ (WHY). Hence by the definition of the outer multiplication we have: $P_\varphi(\varphi)(\alpha_j) = 0_M$ for all $j = 1, \ldots, m$, hence $P_\varphi(\varphi) = 0_{\mathrm{End}_R(M)}$ (WHY).

Finally, for $A \in R^{m\times m}$ and ${}_A\varphi : R^{m\times m} \to R^{m\times m}$, ${}_A\varphi(\boldsymbol{x}) := A\boldsymbol{x}$, one has: $A = A_{{}_A\varphi}$ is the matrix of ${}_A\varphi$ in the standard basis $\mathcal{E}$ of $R^{m\times 1}$. Hence $P_A(A) = P_{{}_A\varphi}({}_A\varphi) = 0_{R^{m\times m}}$. □

7. Diagonalization & Canonical Forms of Matrices

7.1. Basics: Eigenvectors & Eigenvalues, Invariant subspaces.

In the section we consider the following context context:

- V is an F-vector space, $\varphi : V \to V$ is an F-endomorphism of V, and $\lambda \in F$.
  Every $A \in F^{m\times m}$ defines an F-endomorphism ${}_A\varphi : {}^mF \to {}^mF$, ${}_A\varphi(x) = Ax$.

Definition/Remark 7.1. In the above context, define/notice the following:

1) $V_\lambda := \{v \in V \mid \varphi(v) = \lambda v\} = \mathrm{Ker}(\lambda\cdot \mathrm{id}_V - \varphi) \subset V$ is called the λ subspace of V.
2) $\lambda \in F$ is called eigenvalue for φ if there exists $v \neq 0_V$ such that $\varphi(v) = \lambda v$.
   Note that $V_\lambda \neq \{0_V\}$ iff λ is an eigenvalue (WHY).
3) $v \in V$ is called eigenvector for φ if $v \neq 0_V$ and $\exists\ \lambda \in F$ such that $\varphi(v) = \lambda v$.
   Note that $\lambda = 0_F$ is allowed, but **eigenvectors must be** $\neq 0_V$.
4) A subspace $W \subset V$ is called invariant subspace of φ, if $\varphi(W) \subset W$.
   Note that $V_\lambda \subset V$ is and invariant subspace (WHY).

Remark 7.2. In the above notations, the following hold:

1) Let $W \subset V$ be an invariant subspace, and $\overline{V} := V/W$ be the quotient of V by its subspace W. Then $\varphi : V \to V$ gives rise to an endomorphism
$$\overline{\varphi} : \overline{V} \to \overline{V}, \quad \overline{\varphi}(\overline{v}) = \varphi(v) + W, \quad \text{where} \quad \overline{v} := v + W \ \text{(WHY)}$$
2) Suppose that $\dim(V) = m$. Then the ollowing hold:
   a) Let $\psi := \varphi_{|_W}$ be the restriction of φ to W. Then $\psi(W) = \varphi(W) \subset W$, hence $\psi \in \mathrm{End}_F(W)$. Hence if $\mathcal{B} = (w_1, \dots, w_p)$ be a basis of W, one has
   $$\psi(\mathcal{B}) = \mathcal{B}\cdot[\psi]_\mathcal{B} \ \text{ for a unique } \ [\psi]_\mathcal{B} \in F^{p\times p} \ \text{(WHY)}$$
   b) If $\overline{\mathcal{C}} = (\overline{\gamma}_1, \dots, \overline{\gamma}_q)$ is a basis of $\overline{V}$. Since $\overline{\varphi} \in \mathrm{End}_F(\overline{V})$, one has
   $$\overline{\varphi}(\overline{\mathcal{C}}) = \mathcal{C}\cdot[\overline{\varphi}]_{\overline{\mathcal{C}}} \ \text{ for a unique } \ [\overline{\varphi}]_{\overline{\mathcal{C}}} \in F^{q\times q} \ \text{(WHY)}$$

Proposition 7.3 (Characteristic polynomial and Invariant subspaces).
In the above context and notation the following hold:

1) *For any preimage $\mathcal{C} = (\gamma_1, \dots, \gamma_q)$ of $\overline{\mathcal{C}}$ in V, the system $\mathcal{A} = (\beta_1, \dots, \beta_p, \gamma_1, \dots, \gamma_q)$ is a basis of V, ence $m = p + q$. Further, the matrix of φ in the basis $\mathcal{A}$ is of the form*
$$[\varphi]_\mathcal{A} = \begin{pmatrix} [\psi]_\mathcal{B} & * \\ \mathbf{0}_{q\times p} & [\overline{\varphi}]_{\overline{\mathcal{C}}} \end{pmatrix}, \quad \textit{for some matrix } \ * \in F^{p\times q}$$
2) *Hence the characteristic polynomials $P_\psi(t)$, $P_{\overline{\varphi}}(t)$ and $P_\varphi(t)$ satisfy:*
$$P_\varphi(t) = \det\begin{pmatrix} t\boldsymbol{I}_p - [\psi]_\mathcal{B} & -* \\ \mathbf{0}_{q\times p} & t\boldsymbol{I}_q - [\overline{\varphi}]_{\overline{\mathcal{C}}} \end{pmatrix} = \det\Big(t\boldsymbol{I}_p - [\psi]_\mathcal{B}\Big)\ \det\Big(t\boldsymbol{I}_q - [\overline{\varphi}]_{\overline{\mathcal{C}}}\Big) = P_\psi(t)P_{\overline{\varphi}}(t)$$

Proof. To 1): First, $\mathcal{A}$ is free: Let $x := w + v = 0_V$, where $w := a_1\alpha_1 + \dots + a_p\alpha_p \in W$, $w' := b_1\gamma_1 + \dots + b_q\gamma_q$. Then $\overline{w} = 0_{\overline{V}}$, hence $0_{\overline{V}} = \overline{x} = \overline{w}' = b_1\overline{\gamma}_1 + \dots + b_q\overline{\gamma}_q$, thus $b_1 = \dots = b_q = 0_F$ (WHY), and $x = w$. Hence $0_V = x = a_1\alpha_1 + \dots + a_p\alpha_p$, thus $a_1, \dots, a_p = 0_F$ (WHY). Hence $\mathcal{A}$ is free. Second, $\langle\mathcal{A}\rangle_F = V$: Indeed, if $x \in V$, one has $\overline{x} = b_1\overline{\gamma}_1 + \dots + b_q\overline{\gamma}_q$ for some $b_1, \dots, b_q \in F$ (WHY). We setting $w' := b_1\gamma_1 + \dots + b_q\gamma_q$,

one has $\overline{x} = \overline{w}'$. Therefore, $w := x - w'$ satisfies $\overline{w} = \overline{x} - \overline{w}' = 0_{\overline{V}}$ (WHY), thus $w = x - w' \in W$. Hence $w = a_1\beta_1 + \cdots + a_p\beta_p$ for some $a_i \in F$ (WHY), and finally conclude that $x = w + w'$ is the linear combination $x = a_1\beta_1 + \cdots + a_p\beta_p + b_1\gamma_1 + \cdots + b_q\gamma_q$. Notice as well that if $W' := \langle \mathcal{C} \rangle_F$, Then $V = W + W'$ is direct, i.e., $W \cap W' = \{0_V\}$ (WHY). Hence very $x \in V$ has a unique representation of the form:

$$x = w + w', \quad w = \mathcal{B} \cdot [w]_{\mathcal{B}} \in W, \quad w' = \mathcal{C} \cdot [w']_{\mathcal{C}} \in W' \text{ and moreover, } [w']_{\mathcal{C}} = [\overline{x}]_{\overline{\mathcal{C}}} \text{ (WHY).}$$

Hence the coordinate vectors $[w]_{\mathcal{B}}$, $[w']_{\mathcal{C}}$ and $[x]_{\mathcal{A}}$ satisfy:

$$[w]_{\mathcal{A}} = \begin{pmatrix} [w]_{\mathcal{B}} \\ \mathbf{0}_p \end{pmatrix}, \quad [w']_{\mathcal{A}} = \begin{pmatrix} \mathbf{0}_q \\ [w']_{\mathcal{C}} \end{pmatrix}, \quad [x]_{\mathcal{A}} = \begin{pmatrix} [w]_{\mathcal{B}} \\ [\overline{x}]_{\overline{\mathcal{C}}} \end{pmatrix}$$

Finally, the columns of $[\varphi]_{\mathcal{A}}$ are $\boldsymbol{c}_j := [\varphi(\beta_j)]_{\mathcal{A}}$, $1 \leqslant j \leqslant p$, $\boldsymbol{c}_{p+j};= [\varphi(\gamma_j)]_{\mathcal{A}}$, $1 \leqslant j \leqslant q$ (WHY). Hence:

- $\varphi(\beta_j) = \psi(\beta_j)$, thus $\boldsymbol{c}_j = \begin{pmatrix} [\psi(\beta_j)]_{\mathcal{B}} \\ \mathbf{0}_p \end{pmatrix}, 1 \leqslant j \leqslant p$. Therefore $(\boldsymbol{c}_j)_{1 \leqslant j \leqslant p} = \begin{pmatrix} [\psi]_{\mathcal{B}} \\ \mathbf{0}_{q \times p} \end{pmatrix}$ (WHY).

- $\varphi(\gamma_j) = \overline{\varphi}(\overline{\gamma}_j)$, thus $\boldsymbol{c}_{p+j} = \begin{pmatrix} * \\ [\overline{\varphi}(\overline{\gamma}_j)]_{\overline{\mathcal{C}}} \end{pmatrix}$. Therefore, $(\boldsymbol{c}_{p+j})_{1 \leqslant j \leqslant q} = \begin{pmatrix} * \\ [\overline{\varphi}]_{\overline{\mathcal{C}}} \end{pmatrix}$ (WHY).

Therefore, one finally gets $[\varphi]_{\mathcal{A}} = (\boldsymbol{c}_j)_{1 \leqslant j \leqslant p+q} = \begin{pmatrix} [\psi]_{\mathcal{B}} & * \\ \mathbf{0}_{q \times p} & [\overline{\varphi}]_{\overline{\mathcal{C}}} \end{pmatrix}$.

To 2): Recall that for $A \in F^{p \times p}$, $C \in F^{p \times q}$, $B \in F^{q \times q}$ and $O := \mathbf{0}_{q \times p}$, one has: $\det\left(\begin{smallmatrix} A & C \\ O & B \end{smallmatrix}\right) = \det(A)\det(B)$ (WHY). Hence conclude by using assertion 2) concerning $[\varphi]_{\mathcal{A}}$, and the definitions. □

Proposition 7.4. (Basic Facts on Invariant Subspaces)

In the above context and notation the following hold:

1) *Let $\lambda_1, \ldots, \lambda_n$ be distinct eigenvalues of φ. Then the sum $V_{\lambda_1} + \cdots + V_{\lambda_n}$ is a direct sum, i.e., if $v_i \in V_{\lambda_i}$, then $\sum v_i = 0_V$ iff $v_i = 0_V$ for all $i = 1,\ldots,n$.*
2) *Hence if V is finite dimensional, φ has at most finitely many eigenvalues, maybe none. And if $\lambda_1, \ldots, \lambda_n$ are all the eigenvalues of φ, and $\mathcal{A}_k = (v_{kj})_{1 \leqslant j \leqslant d_k}$ is an F-basis of V_{λ_k} with $d_k = \dim(V_{\lambda_k})$, then the system $\mathcal{A} := (\mathcal{A}_1, \ldots \mathcal{A}_n) := (v_{kj})_{1 \leqslant k \leqslant n, 1 \leqslant j \leqslant d_k}$ is F-free.*

$(*)$ *In particular, $n \leqslant \sum_{k=1}^{n} \dim(V_{\lambda_k}) \leqslant \dim(V)$.*

Proof. To 1): Induction on n: For $n = 1$ there is nothing to prove. We prove that "$n \Rightarrow (n+1)$": Let $\sum_{i=1}^{n+1} v_i = 0_V$ with $v_i \in V_{\lambda_i}$. We show that $v_i = 0$ for all $1 \leqslant i \leqslant n+1$. Since $\sum_i v_i = 0_V$, we get

$$0_V \overset{\text{why}}{=} \varphi(\textstyle\sum_{i=1}^{n+1} v_i) \overset{\text{why}}{=} \sum_{i=1}^{n+1} \varphi(v_i) \overset{\text{why}}{=} \sum_{i=1}^{n+1} \lambda_i v_i.$$

Second, $\sum_{i=1}^{n+1} v_i = 0_V$ implies $0_V = \lambda_{n+1} 0_V = \sum_i \lambda_{n+1} v_i$. Hence $\sum_i \lambda_{n+1} v_i = 0_V = \sum_i \lambda_i v_i$ (WHY), thus concluding that $0_V = \sum_{i=1}^{n} (\lambda_i - \lambda_{n+1}) v_i$ (WHY). Then setting $v_i' := (\lambda_i - \lambda_{n+1}) v_i$, one has $v_i' \in V_{\lambda_i}$ for $i = 1, \ldots, n$ and $\sum_{i=1}^{n} v_i' = 0_V$ (WHY). By the induction hypothesis we get: $v_k' = 0_V$ for all $i \leqslant n$ (WHY). Equivalently, $(\lambda_{n+1} - \lambda_i) v_i = 0_V$, and since $\lambda_{n+1} \neq \lambda_i$ for all $i = 1, \ldots, n$, we must have $v_i = 0_V$ (WHY). Conclude that $v_{n+1} = 0_V$ as well (WHY).

To 2): We notice that if $\lambda_1, \ldots, \lambda_n$ are distinct eigenvalues of φ, then $V_{\lambda_k} \neq \{0_V\}$ (WHY), and we claim that $\mathcal{A} = (v_{kj})_{1 \leqslant k \leqslant n, 1 \leqslant j \leqslant d_k}$ is an F-basis of $V_{\lambda_1} + \cdots + V_{\lambda_n}$. Indeed, we first show that $\mathcal{A}$ is a free system. Let namely $\sum_k \sum_{j=1}^{d_k} a_{kj} v_{kj} = 0_V$ be a linear combination of the elements of $\mathcal{A}$. Then for each fixed k one has: $v_k := \sum_{j=1}^{d_k} a_{kj} v_{kj} \in V_{\lambda_k}$ (WHY). Further, $\sum_k v_k = \sum_k \sum_{j=1}^{d_k} a_{kj} v_{kj} = 0_V$ (WHY). Thus by assertion 1) it follows that $v_k = 0_V$ for each $k = 1, \ldots, n$. Since each $\mathcal{A}_k = (v_{kj})_{1 \leqslant j \leqslant d_k}$ is an F-basis of V_k, conclude that $a_{kj} = 0_F$ for all $j = 1, \ldots, d_k$ (WHY), thus $a_{kj} = 0_F$ for all k, j. Finally, since $d_k = \dim(V_{\lambda_k}) \geqslant 1$ and $\mathcal{A}$ has $\sum_{k=1}^{n} d_k$ entries and $\mathcal{A}$ is free, it follows that $n \leqslant \sum_{k=1}^{n} \dim(V_{\lambda_k}) \leqslant \dim(V)$ (WHY). □

7.2. Diagonalization of endomorphisms/matrices.

Let V be a finite dimensional F-vector space, $m = \dim_F(V)$. We next discuss the diagonalization of an endomorphism $\varphi \in \mathrm{End}_F(V)$, respectively of a matrix $A \in F^{m\times m}$.

Definition 7.5. Consider the context/notations from the previous section.

1) We say that $\varphi \in \mathrm{End}_F(V)$ is diagonalizable, if there is a basis $\mathcal{V} = (v_1, \dots, v_m)$ of V such that $[\varphi]_{\mathcal{V}} = (a_{ij})_{i,j} \in F^{m\times m}$ is a diagonal matrix, i.e., $a_{ij} = 0_F$ for all $i \neq j$.
2) We say that $A \in F^{m\times m}$ is diagonalizable, if A is similar to a diagonal matrix, i.e., there exists $S \in \mathrm{GL}_m(F)$ such that SAS^{-1} is diagonal.

Proposition 7.6. (Diagonalization & Eigenspaces)
For an endomorphism $\varphi \in \mathrm{End}_F(V)$ of a finite dimensional F-vector space V, TFAE:

i) *φ is diagonalizable, i.e., $\exists$ an F-basis $\mathcal{V}$ of V such that $[\varphi]_{\mathcal{A}}$ is a diagonal matrix.*
ii) *$\sum_{\lambda\in F} V_\lambda = V$.*
iii) *V has an F-basis $\mathcal{V}$ consisting of eigenvectors.*

Proof. i) $\Rightarrow$ ii): Let $\mathcal{V} = (v_1, \dots, v_m)$ be an F-basis such that $[\varphi]_{\mathcal{V}}$ is diagonal, say $[\varphi]_{\mathcal{A}} = D(a_1, \dots, a_m)$ is the diagonal matrix with entries $a_1, \dots, a_m$ on the diagonal. Then by definitions, $\varphi(\mathcal{V}) = \mathcal{V}[\varphi]_{\mathcal{V}}$ is equivalent to $\varphi(v_i) = a_i v_i$ for $i = 1, \dots, m$. Let $\lambda_1, \dots, \lambda_n$ be the distinct elements of the diagonal, and suppose that $d_1, \dots, d_n$ are the number to times for which $a_i = \lambda_k$ for each $k = 1, \dots, n$. **Terminology**: d_k is called the geometric multiplicity of λ_i. Then setting $\mathcal{A}_k = (v_{kj})_{a_j=\lambda_k}$, it follows that $\varphi(v_{kj}) = \lambda_k v_{kj}$ for all elements v_{kj} from $\mathcal{A}_k$, thus the F-subspace $W_k := \langle \mathcal{A}_k \rangle_F$ is contained in V_{λ_k} for all $k = 1, \dots, n$. Further, since $\mathcal{V} = (\mathcal{A}_1, \dots, \mathcal{A}_n)$ (WHY), it follows that $W_1 + \cdots + W_n = \langle \mathcal{A} \rangle$. Hence we have the following:

$$V = W_1 + \cdots + W_n \subset V_{\lambda_1} + \cdots + V_{\lambda_n} \subset V.$$

Hence we conclude that all inclusions must be equalities (WHY), hence $W_k = V_{\lambda_k}$ for each $k = 1, \dots, n$, and finally $\dim(V) = \sum_k \dim(V_{\lambda_k}$.
ii) $\Rightarrow$ iii): By the assertion 3) of the previous Proposition, it follows there are only finitely many eigenvalues $\lambda_1, \dots, \lambda_n$ of φ and if $\mathcal{A}_k$ be an F-basis of V_{λ_k} for each k, then $\mathcal{V} := (\mathcal{A}_1, \dots, \mathcal{A}_n)$ is an F-basis of V. Further, for every vector v in $\mathcal{V}$ one has: If v is an element of $\mathcal{A}_k$, then $\varphi(v) = \lambda_k v$, thus v is an eigenvector to the eigenvalue λ_k. Thus $\mathcal{V}$ is an F-basis consisting of eigenvectors of φ.
iii) $\Rightarrow$ i): Let $\mathcal{V} = (v_1, \dots, v_m)$ be an F-basis of V consisting of eigenvectors. Then by definition, for every v_i, there exists some $\lambda_i \in F$ such that $\varphi(v_i) = \lambda_i v_i$. Thus by the definition of $[\varphi]_{\mathcal{V}}$ it follows that $[\varphi]_{\mathcal{V}} = D(\lambda_1, \dots, \lambda_m)$ is the diagonal matrix with entries $\lambda_1, \dots, \lambda_m$ on the diagonal. □

Recall that for every polynomial $P(t) \in F[t]$, we say that $\lambda \in F$ is a *root* of $P(t)$, if $P(\lambda) = 0$. Further, λ is a root of $P(t)$ iff $(t - \lambda)$ divides $P(t)$ in $F[t]$. Finally, we say that $\lambda \in F$ is a root of (algebraic) multiplicity m_λ of $P(t)$, if m_λ is the maximal natural number such that $(t - \lambda)^{m_\lambda}$ divides $P(t)$. Note that $m_\lambda = 0 \overset{\text{def}}{\longleftrightarrow}$ λ is not a root of $P(t)$.

In particular, if $\lambda_1, \dots, \lambda_n$ are all the distinct roots of $P(t)$ in F, and $m_1, \dots, m_n$ are their (algebraic) multiplicities, then $P_0(t) := \prod_k (t - \lambda_k)^{m_k}$ is the largest product of degree one polynomials dividing $P(t)$. Hence $P(t) = P_0(t)\, Q(t)$ in $F[t]$, and $Q(t)$ has not roots in F.

Proposition 7.7. (Diagonalization & Characteristic polynomial)
Let $\varphi \in \mathrm{End}_F(V)$ be an F-endomorphism of a finite dimensional F-vector space V. Let $\lambda_1, \dots, \lambda_n \in F$ be the distinct roots of the characteristic polynomial $P_\varphi(t)$ in F, and for each $k = 1, \dots, n$ let m_k be the multiplicity of λ_k. Then the following hold:

1) *$\lambda \in F$ is an eigenvalue of φ iff λ is one of the roots λ_k of $P_\varphi(t)$.*
2) *For every root λ_k of $P_\varphi(t)$, one has that $\dim(V_{\lambda_k}) \leqslant m_k$. In particular, the following assertions are equivalent:*
   i) *φ is diagonalizable.*
   ii) *$\dim(V_{\lambda_k}) = m_k$ for each $k = 1, \ldots, n$ and $\sum_k m_k = \dim(V)$.*
3) *Finally, if φ is diagonalizable, and $\mathcal{V} = (\mathcal{A}_1, \ldots, \mathcal{A}_n)$ with $\mathcal{A}_k$ an F-basis of V_{λ_k} for each k, then*

$$[\varphi]_{\mathcal{V}} = \begin{pmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix}$$

is diagonal such that each λ_k appears precisely m_k times on the diagonal for $k = 1, \ldots, n$.

Proof. To 1): Let $\mathcal{A}$ be any F-basis of V and A_φ be the matrix of φ in the F-basis $\mathcal{A}$. The one has: $\lambda \in F$ is eigenvalue iff $\exists v \neq 0_V$ such that $\varphi(v) = \lambda v$ (WHY) iff $(\lambda \operatorname{id}_V - \varphi)(v) = 0_V$ (WHY) iff $\lambda \operatorname{id}_V - \varphi$ is not invertible in $\operatorname{End}_F(V)$ (WHY) iff $\lambda I_m - A_\varphi$ is not invertible in $F^{m\times m}$ (WHY) iff $\det(\lambda I_m - A_\varphi) = 0_F$ (WHY) iff $P_\varphi(\lambda) = 0_F$ (WHY).

To 2): Let $\mathcal{A}_k = (v_{kj})_{1\leqslant j\leqslant d_k}$ be an F-basis of V_{λ_k}, and $\mathcal{B} = \mathcal{A}_k \cup (v_{d_k+1}, \ldots, v_m)$ be any F-basis of V containing $\mathcal{A}_k$. Then the matrix of φ in the basis $\mathcal{B}$ is of the form $[\varphi]_{\mathcal{B}} = \left(\begin{smallmatrix} D & D' \\ O & D'' \end{smallmatrix}\right)$, where setting $p := d_k$, $q := m - d_k$, one has: D is the diagonal matrix $D = \lambda_k \boldsymbol{I}_p$, and $D' \in F^{p\times q}$, $O = \mathbf{0}_{p\times q}$ and $D'' \in F^{q\times q}$ (WHY). Therefore, $P_\varphi(t) = P_{[\varphi]_{\mathcal{B}}}(t) = P_D(t)P_{D''}(t) = (t-\lambda_k)^{d_k} P_{D''}(t)$ (WHY). In particular, since $(t-\lambda_k)^{d_k}$ divides $P_\varphi(t)$, it follows that $\dim(V_{\lambda_k}) = d_k \leqslant m_k$ (WHY).

Next let $P_0(t) = \prod_k (t-\lambda_k)^{m_k}$ be the maximal product of linear factors dividing $P_\varphi(t)$. Then $P_0(t)$ divides $P_\varphi(t)$ in $F[t]$, and $Q(t) := P_\varphi(t)/P_0(t)$ has not linear factors in $F[t]$, or equivalently, no roots in F (WHY). In particular, one has:

$$\dim(V) = \deg P_\varphi(t) = \deg P_0(t) + \deg Q(t) \text{ (WHY)},$$
$$\deg P_0(t) = \textstyle\sum_k m_k \geq \sum_k \dim(V_{\lambda_k}) \text{ (WHY)}.$$

Thus using Diagonalization & Eigenspaces, the following are equivalent:

- φ is diagonalizable
- $\dim(V) = \sum_k \dim(V_{\lambda_k})$
- $\dim(V) = \deg P_\varphi(t) = \deg P_0(t) + \deg Q(t)$
  $\geq \deg P_0(t) = \sum_k m_k \geq \sum_k \dim(V_{\lambda_k}) = \dim(V)$
- all the inequalities above are equalities
- $\dim(V) = \sum_k \dim(V_{\lambda_k})$ and $m_k = \dim(V_{\lambda_k})$ for all $k = 1, \ldots, n$.

To 3): Clear by the fact that $\varphi(v_{kj}) = \lambda_k v_{kj}$ for every k and $1 \leqslant j \leqslant d_k$, etc.
This concludes the proof of the Proposition. □

Diagonalization Procedure

- First recall that diagonalizing an emdomorphism $\varphi \in \operatorname{End}_F(V)$ is equivalent to diagonalizing its matrix $A_\varphi := [\varphi]_{\mathcal{A}} \in F^{m\times m}$, where $\mathcal{A}$ is some F-basis of V.
- Diagonalizing matrices $A \in F^{m\times m}$
  - Compute $P_A(t)$ and its roots $\lambda_1, \ldots, \lambda_n$ with their multiplicities $m_1, \ldots, m_n$.
  - Check whether $\sum_k m_k = m$.

($*$) If not, **STOP**: *A cannot be diagonalized.*

- If $\sum_k m_k = m$, compute a basis $\mathcal{A}_k$ for $V_{\lambda_k} = \{v \in {}^mF \mid \lambda_k v - Av = 0\}$
- Check whether $\dim(V_{\lambda_k}) = m_k$ for each k.

$(*)$ If not, **STOP**: *A cannot be diagonalized.*

- If $\dim(V_{\lambda_k}) = m_k$ for each k, let S be the matrix $S = (v_1, \ldots, v_m)$ be the matrix whose columns are the eigenvectors computed above for $1 \leqslant k \leqslant n$.

- If $\mathcal{V} := (v_1, \ldots, v_m)$ is the corresponding F-basis of ${}^mF$, and $\mathcal{E}$ is the standard basis of ${}^mF$, then $\mathcal{V} = \mathcal{E}S$ (WHY), hence S is the base change matrix from $\mathcal{V}$ to $\mathcal{E}$ in ${}^mF$ (WHY).
- Conclude that $D = S^{-1}AS = \begin{pmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix}$ is the diagonalization of A, where each λ_k appears m_k times on the diagonal (WHY).

7.3. Upper-triangular form of an endomorphism / matrix.

Definition 7.8. In the notations and context from the Subsection 6.1 one defines:

1) $\varphi \in \mathrm{End}_F(V)$ is called triangulable, if there exists a basis $\mathcal{V}$ of V such that $[\varphi]_{\mathcal{V}} = (c_{ij})_{i,j}$ is an upper triangular matrix, i.e., $c_{ij} = 0_F$ for $i > j$.
2) A matrix $A = (a_{ij})_{i,j} \in F^{m\times m}$ is called triangulable, if there exists $S \in \mathrm{GL}_m(F)$ such that $S^{-1}AS$ is an upper triangular matrix.

Proposition 7.9. *Let $\varphi \in \mathrm{End}_F(V)$ be a fixed endomorphism. The following hold:*

1) *For $\varphi \in \mathrm{End}_F(V)$ the following are equivalent:*
   i) *$P_\varphi(t)$ splits in linear factors in $F[t]$, i.e., $P_\varphi(t)$ has all its roots in F.*
   ii) *φ is triangulable.*
   - *Moreover, if* i) *is satisfied and the roots of $P_\varphi(t)$ are explicitly given, there is an explicit procedure to compute a basis $\mathcal{V}$ such that $[\varphi]_{\mathcal{V}}$ is upper triangular.*
2) *The same holds similarly for $m \times m$ matrices A over F.*

Proof. Clearly, assertion 2) follows from assertion 1) (WHY), therefore it is sufficient to prove assertion 1).

i) $\Rightarrow$ ii): Suppose that φ is triangulable, i.e., there exists a basis $\mathcal{V}$ of V such that $[\varphi]_{\mathcal{V}} = (c_{ij})_{i,j}$ is upper triangular. Then one has:

$$P_\varphi(t) = \det\left(t\boldsymbol{I}_m - (c_{ij})_{i,j}\right) = \prod_{i=1}^m (t - c_{ii}) \quad \text{(WHY)}$$

Hence $P_\varphi(t)$ splits in linear factors in $F[t]$, and its roots (counted with multiplicities) are c_{ii}, $1 \leqslant i \leqslant m$.

i) $\Rightarrow$ ii): Suppose that $P_\varphi(t) = \prod_i (t - \lambda_i)^{m_i}$ where $\lambda_1, \ldots, \lambda_n \in F$ are the distinct roots of $P_\varphi(t)$, and $m_1, \ldots, m_n$ are their multiplicities. We make induction on $m = m_1 + \cdots + m_n$ as follows:

$m = 1$: Then $V = V_{\lambda_1}$, and $[\varphi]_{(v_1)} = (\lambda_1) \in F^{1\times 1}$ (WHY) is upper triangular.

$m \Rightarrow (m+1)$: Let $v_1 \in V_{\lambda_1}$ be nonzero eigenvector to the eigenvalue λ_1.

• **Note** *that if φ is concretely given, e.g., via its matrix A_φ is a basis $\mathcal{A}$ of V, the eigenvectors to every eigenvalue λ_i of φ, that is of A_φ, can be* **effectively computed**, *as discussed in the previous subsection.*

Letting $W := Fv_1 \subset V = \langle v_1 \rangle_F$, one has $\dim(W) = 1$, and W is a φ-invariant subspace (WHY). Further, $\psi := \varphi_{|W} : W \to W$ satisfies $\psi(v) = \lambda_1 v$ for all $v \in W$ (WHY), hence if $\mathcal{B} := (v_1)$, then $[\psi]_{\mathcal{B}} = (\lambda_1) \in F^{1\times 1}$ (WHY). Setting $\overline{V} := V/W$, one has: $\dim(\overline{V}) = \dim(V) - \dim(W) = (m+1) - 1 = m$ (WHY), and $\overline{\varphi} : \overline{V} \to \overline{V}$ satisfies:

$$P_\varphi(t) = P_\psi(t)\, P_{\overline{\varphi}}(t)$$

in $F[t]$. Since $P_\varphi(t) = \prod_i (t-\lambda_i)^{m_i}$, and $P_\psi(t) = t - \lambda_1$, it follows that $P_{\overline{\varphi}}(t) = (t-\lambda_1)^{m_1 - 1} \prod_{i>1}^m (t-\lambda_i)^{m_i}$. Conclude that $\overline{\varphi} : \overline{V} \to \overline{V}$ satisfies condition i), and $\dim(\overline{V}) = \dim(V) - 1 = m$, thus i) $\Rightarrow$ ii) holds for $\overline{\varphi}$.

Hence by the induction hypothesis, there exists a basis $\overline{\mathcal{C}}$ of $\overline{V}$ such that $[\overline{\varphi}]_{\overline{\mathcal{C}}} \in F^{m\times m}$ is in upper triangular. Hence by Proposition 7.3, it follows that letting $\mathcal{C}$ be any preimage of $\overline{\mathcal{C}}$ in V, one has: $\mathcal{V} := (v_1, \mathcal{C})$ is a basis of V such that

$$[\varphi]_{\mathcal{V}} = \begin{pmatrix} \lambda_1 & \boldsymbol{a} \\ \mathbf{0}_m & [\overline{\varphi}]_{\overline{\mathcal{C}}} \end{pmatrix}, \quad \text{where } \mathbf{0}_m = 0_{\,mF} \text{ and } \boldsymbol{a} \in F^m \text{ (WHY).}$$

And since $[\overline{\varphi}]_{\overline{\mathcal{C}}} \in F^{m\times m}$ is upper triangular, so is $\begin{pmatrix} \lambda_1 & \boldsymbol{a} \\ \mathbf{0}_m & [\overline{\varphi}]_{\overline{\mathcal{C}}} \end{pmatrix} = [\varphi]_{\mathcal{V}}$ (WHY).

- **Note** *that if φ is concretely given, e.g., via its matrix A_φ is a basis $\mathcal{A}$ of V, so is $\overline{\varphi}$, and $\overline{\mathcal{C}}$ can be* ***inductively effectively computed****. In particular, one can* **effectively compute** *the basis $\mathcal{V}$ and $[\varphi]_{\mathcal{V}}$.*

□

7.4. Upper-triangular block form of an endomorphism/matrix.

There is a refinement of the above facts concerning the upper-triangular for of an endomorphism/matrix as follows.

Definition 7.10.

1) A λ-upper-tringular block is any matrix square matrix over F which is upper-triangular and has all the diagonal entries equal to λ.
2) A matrix $A \in F^{m\times m}$ is said to be in upper-triangular block form, if there exist distinct $\lambda_1, \ldots, \lambda_n \in F$, $m_1, \ldots, m_n \in \mathbb{N}_{>0}$, and upper triangular blocks $D_{\lambda_k} \in F^{m_k \times m_k}$ such that:

$$A = \big(A_{kl}\big)_{1\leqslant k,l\leqslant n}, \qquad A_{kk} = D_{\lambda_k}, \qquad A_{kl} = \mathbf{0}_{m_k \times m_l} \text{ for } k \neq l.$$

Proposition 7.11. *Let $\varphi \in \mathrm{End}_F(V)$ be a fixed endomorphism. The following hold:*

1) *For every $\lambda \in F$ there exists a unique maximal φ-invariant subspace $W_\lambda \subset V$ such that setting $\psi := \varphi|_{W_\lambda}$ one has: $P_\psi(t) = (t-\lambda)^n$ is the highest power of $t-\lambda$ dividing $P_\varphi(t)$.*
2) *If $\lambda_1, \ldots, \lambda_n \in F$ are the distinct roots of $P_\varphi(t)$ in F, and $W_i := W_{\lambda_i} \subset V$ are the corresponding subspaces, then $W_1 + \cdots + W_n$ is a direct sum.*

Proof. To 1): By the characterization of eigenvalues of φ one has: $t - \lambda$ divides $P_\varphi(t)$ iff λ is an eigenvalue of φ. Second, by Proposition 7.3, it follows that $P_\psi(t)$ divides $P_\varphi(t)$. Hence if $P_\psi(t) = (t-\lambda)^n$ one has: $n = \deg\big(P_\psi(t)\big) = \dim_F(W_\lambda) > 0$ iff $t-\lambda$ divides $P_\varphi(t)$ iff λ is an eigenvalue of φ. Thus from now let suppose that $n = \deg\big(P_\psi(t)\big) = \dim_F(W) > 0$, thus equivalently, λ is an eigenvalue of φ. Then reasoning as in the proof of Proposition 7.3, one has: Let $v \in V_\lambda$ be $\neq 0_V$, hence an eigenvector to λ, and set $W_1 := Fv \subset V_\lambda$, and $\overline{V} := V/W_1$. Then $\dim(\overline{V}) = \dim(V) - 1$ and $P_\varphi(t) = (t-\lambda)\, P_{\overline{\varphi}}(t)$. Hence reasoning by induction on $m = \dim(V)$, it follows that there exists a unique maximal $\overline{\varphi}$-invariant subspace $\overline{W}_\lambda \subset \overline{V}$ of $\overline{V}$ such that setting $\overline{\psi} := \overline{\varphi}|_{\overline{W}_\lambda}$, one has: $P_{\overline{\psi}}(t) = (t-\lambda)^{\overline{n}}$ is the maximal power of $t-\lambda$ which divides $P_{\overline{\varphi}}(t)$, thus $\overline{n} = \dim(\overline{W}_\lambda)$ as well (WHY). Hence setting $\overline{\overline{V}} := \overline{V}/\overline{W}_\lambda$ and letting $\overline{\overline{\varphi}} : \overline{\overline{V}} \to \overline{\overline{V}}$ be the corresponding morphism, one has

$$P_{\overline{\varphi}}(t) = P_{\overline{\psi}}(t)\, P_{\overline{\overline{\varphi}}} = (t-\lambda)^{\overline{n}} P_{\overline{\overline{\varphi}}}$$

and $t - \lambda$ does not divide $P_{\overline{\overline{\varphi}}}$ (WHY).

Next let $W \subset V$ be the preimage of $\overline{W}_\lambda \subset \overline{V}$ under the canonical surjective projection $V \to \overline{V}$. Then $W \subset V$ is φ-invariant (WHY), and $W_1 = Fv \subset W$ is φ-invariant as well (WHY). Hence setting $\psi := \varphi|_W$, one has that $W_1 \subset W$ is a ψ-invariant subspace, and further:

$$P_\psi(t) = (t-\lambda)\, P_{\overline{\psi}} = (t-\lambda)^{1+\overline{n}}, \quad P_\varphi(t) = (t-\lambda)\, P_{\overline{\varphi}}(t) = (t-\lambda) P_{\overline{\psi}}(t)\, P_{\overline{\overline{\varphi}}} = (t-\lambda)^{1+\overline{n}}\, P_{\overline{\overline{\varphi}}}$$

Moreover, since $t-\lambda$ does not divide $P_{\overline{\varphi}}$, it follows that $n := 1+\overline{n}$ is the highest power of $t-\lambda$ dividing $P_\varphi(t)$, and $W_\lambda := W \subset V$ is a φ-invariant subspace with $\dim(W_\lambda) = n$ and $P_\psi(t) = (t-\lambda)^n$, where $\psi := \varphi|_{W_\lambda}$.

Finally, to prove the uniqueness of $W_\lambda = W$, let $W' \subset V$ be a further subspace which is φ-invariant, and $\psi' := \varphi_{W'}$ has $P_{\psi'}(t) = (t-\lambda)^n$. *By contradiction,* suppose that $W \neq W'$. Then one has: First, $W'' := W + W'$ is φ-invariant (WHY), and $\overline{W}'' := W''/W \subset V/W =: \overline{V}$ is a non-zero subspace of $\overline{V}$ (WHY), etc...

To 2): Let m_i be the multiplicity of λ_i for $i = 1, \dots, m$, hence by assertion 1) one has $\dim_F(W_i) = m_i$ for $i = 1, \dots, n$. Denote $P_0(t) := (t-\lambda_1)^{m_1} \dots (t-\lambda_n)^{m_n}$, and notice that $P_0(t)$ divides $P_\varphi(t)$ (WHY), and setting $P_\varphi(t) = P_0(t)P_1(t)$, it follows that $P_1(t)$ has no roots on F (WHY). We set

$$V_0 := W_1 + \dots + W_n \subset V$$

and notice the following:

- Since each W_i is an F-subspace of V, so is V_0 (WHY).
- Since each W_i is φ-invariant, so is V_0 (WHY). Define $\varphi_0 := \varphi|_{V_0}$.
- Since $W_i \subset V_0 \subset V$, one has that $\varphi_0|_{W_i} = \varphi|_{W_i} = \psi_i$, and therefore:

$$P_{\psi_i}(t) = (t-\lambda_i)^{m_i} \text{ divides } P_{\varphi_0}(t) \text{ (WHY)}.$$

Hence, since $\lambda_1, \dots, \lambda_n$ are distinct, we conclude that the product $P_0 := (t-\lambda_1)^{m_1} \dots (t-\lambda_n)^{m_n}$ divides $P_{\varphi_0}(t)$ (WHY). Thus one finally has:

$$\dim_F(V_0) = \deg\big(P_{\varphi_0}(t)\big) \geqslant \deg\big(P_0(t)\big) = m_1 + \dots + m_r = \dim_F(W_1) + \dots + \dim_F(W_n) \geq \dim_F(V_0) \text{ (WHY)}.$$

Therefore, all the above inequalities are equalities (WHY), thus concluding $\dim_F(V_0) = \dim_F(W_1) + \dim_F(W_n)$. From this it follows that $V_0 = W_1 + \dots + W_n$ is a direct sum (WHY) [proof exercise!]. □

Theorem 7.12. *Let V be a finite dimension F-vector space, and $\mathcal{A}$ be a basis. One has:*

1) *The following are equivalent:*
   i) *$P_\varphi(t)$ is a product of linear factors in $F[t]$.*
   ii) *There exists a basis $\mathcal{V}$ of V such that $[\varphi]_{\mathcal{V}}$ is in upper-triangular block form.*

- *Moreover, if condition* i) *is satisfied, there is an* ***effective inductive procedure*** *to compute $\mathcal{V}$ from the data $\mathcal{A}$, $[\varphi]_{\mathcal{A}}$.*

2) *Correspondingly, for $A \in F^{m\times m}$ the following are equivalent:*
   i) *$P_A(t)$ is a product of linear factors in $F[t]$.*
   ii) *There exists an invertible matrix $S \in \mathrm{GL}_m(F)$ such that $S^{-1}AS$ is in upper-triangular block form.*

Proof. Clearly, assertions 1) and 2) are equivalent (WHY).

ii) ⇒ i): We notice that a matrix in upper-triangular block form is obviously upper triangular (WHY). Hence by Proposition 7.9, it follows that P_φ is a product of linear factors (WHY).

i) ⇒ ii): Let $\lambda_1, \dots, \lambda_n$ be the distinct roots of $P_\varphi(t)$ or $P_A(t)$, and $m_1, \dots, m_n$ be their multiplicities, hence $P_\varphi(t) = \prod_i (t-\lambda_i)^{m_i}$ (WHY). Further let $W_i := W_{\lambda_i} \subset V$ be the corresponding λ_i-subspaces. Then by Proposition 7.11, 2) above, one has $V = W_1 + \dots + W_n$ as a direct sum of φ-invariant subspaces. Further, setting $\psi_i := \varphi|_{W_i}$, it follows that $P_{\psi_i}(t) = (t-\lambda_i)^{m_i}$. Hence since $P_{\psi_i}(t)$ splits in linear factors over F, by Proposition 7.9 there exists an **effective procedure** to find a basis $\mathcal{V}_i = (v_{i1}, \dots, v_{im_i})$ of W_i such that

$$D_{\lambda_i} := [\psi_i]_{\mathcal{V}_i}$$

is an (upper) triangular matrix. Moreover, since on the elements the diagonal of D_{λ_i} are the eigenvalues of $P_{\psi_i}(t)$ with their multiplicities, it follows that D_{λ_i} is actually an upper λ_i-block (WHY). Therefore one has:

• Let $\mathcal{V}_i = (v_{i1}, \dots, v_{im_i})$ be the above basis of W_i, thus such that each $[\psi_i]_{\mathcal{V}_i}$ is an upper λ_i-block. Setting

$$\mathcal{V} = (\mathcal{V}_1, \dots, \mathcal{V}_n) = (v_{ij})_{1\leqslant i\leqslant n, 1\leqslant j\leqslant n_i}$$

the following hold:

a) First, $\mathcal{V}$ is a basis of V.

Indeed, every $v \in V$ has a unique writing in the form $v = w_1 + \cdots + w_n$ with $w_i \in W_i$ (WHY). And each w_i has a unique writing in the form $w_i = a_{i1}v_{i1} + \cdots + a_{im_i}v_{im_i}$, $a_{ij} \in F$ (WHY).

b) From this we deduce that $\mathcal{V}$ is a system of generators of V, and that $\mathcal{V}$ is free (WHY); hence $\mathcal{V}$ is a basis.

c) For every $i = 1, \ldots, n$ one has $\varphi(\mathcal{V}_i) = \psi_i(\mathcal{V}_i) = \mathcal{V}_i \cdot D_{\lambda_i}$, and therefore:

$$[\varphi]_{\mathcal{V}} = (A_{kl})_{k,l}, \quad \text{where } A_{kk} := D_{\lambda_k},\ A_{kl} = \boldsymbol{O}_{m_k m_l} \text{ for } k \neq l \text{ (WHY)}$$

Hence finally, $[\varphi]_{\mathcal{V}}$ is in upper triangular block form.

- *Moreover, if φ is concretely given, e.g., via its matrix A_φ is some basis $\mathcal{A}$ of V, then the basis $\mathcal{V}$ and the matrix $[\varphi]_{\mathcal{V}}$ are* **effectively computable.**

□

7.5. **The Jordan canonical form of endomorphisms/matrices** (Special Case).

We next want to address the question concerning the "simplest" form of the matrix $[\varphi]_{\mathcal{A}}$ of an endomorphism $\varphi \in \mathrm{End}_F(V)$, respectively that of the conjugates $S^{-1}A\,S$, $S \in \mathrm{GL}_m(F)$, of a matrix $A \in F^{m\times m}$ in the case $P_\varphi(t)$, respectively $P_A(t)$ split in linear factors in $F[t]$, or equivalently, have all their roots in F. That simplest form is called the Jordan canonical form, and it is a refinement of the upper triangular form given above.

Definition 7.13.

1) A primitive Jordan λ-block, $\lambda \in F$, is any upper triangular matrix $J_\lambda^{(p)} = (a_{ij})_{i,j} \in F^{p\times p}$ satisfying $a_{ii} = \lambda$ for all i, $a_{i\,i+1} = 1_F$ for all $i < p$, and $a_{ij} = 0_F$ else. Hence one has:

$$J_\lambda^{(1)} = (\lambda); \quad J_\lambda^{(2)} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}; \quad J_\lambda^{(3)} = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}; \quad \text{etc.},$$

2) A Jordan λ-block, $\lambda \in F$, is any upper triangular matrix of the form $J_\lambda = \big(J_{rs}\big)_{r,s}$ where $J_{rr} = J_\lambda^{(m_r)} \in F^{p_r\times p_r}$ are primitive Jordan λ-blocks, and $J_{rs} = \mathbf{0}_{p_r\times p_s}$ for $r \neq s$.

3) Finally, $A \in F^{m\times m}$ is said to be in Jordan normal form, if there exist distinct $\lambda_1, \ldots, \lambda_n$ and positive $m_1, \ldots, m_n \in \mathbb{N}$, and Jordan λ_k-blocks $J_{\lambda_k} \in F^{m_k\times m_k}$ such that:

$$A = \big(J_{kl}\big)_{1\leqslant k,l\leqslant n}, \quad J_{kk} = J_{\lambda_k}, \quad J_{kl} = \mathbf{0}_{m_k\times m_l} \text{ for } k \neq l.$$

Theorem 7.14. *Let V be a finite dimension F-vector space, and $\mathcal{A}$ be a basis. TFH:*

1) *The following are equivalent:*
   i) *$P_\varphi(t)$ is a product of linear factors in $F[t]$.*
   ii) *There exists a basis $\mathcal{V}$ of V such that $[\varphi]_{\mathcal{V}}$ is in Jordan canonical form.*

- *Moreover, if condition* i) *is satisfied, there is an* ***effective inductive procedure*** *to compute $\mathcal{V}$ from $\mathcal{A}$ and $[\varphi]_{\mathcal{A}}$.*

2) *Correspondingly, for $A \in F^{m\times m}$ the following are equivalent:*
   i) *$P_A(t)$ is a product of linear factors in $F[t]$.*
   ii) *There exists an invertible matrix $S \in \mathrm{GL}_m(F)$ such that $S^{-1}A\,S$ is in Jordan canonical form.*

Proof. Clearly, assertions 1) and 2) are equivalent (WHY). Further, for the implication ii) $\Rightarrow$ i), notice that a matrix in Jordan canonical form is obviously upper triangular (WHY). Hence ii) $\Rightarrow$ i) follows from the previous Proposition (WHY). Thus it is left to prove that i) $\Rightarrow$ ii). Let $\lambda_1, \dots, \lambda_n$ be the distinct roots of $P_\varphi(t)$ or $P_A(t)$, and $m_1, \dots, m_n$ be their multiplicities, hence $P_\varphi(t) = \prod_i (t-\lambda_i)^{m_i}$ (WHY).

- We make induction n as follows:

$n = 1$: In particular, $m = m_1$, and setting $\lambda := \lambda_1$, one has $P_\varphi(t) = (t-\lambda)^m$ an/or $P_A(t) = (t-\lambda)^m$.

- We next make induction on m as follows:

- Verification step $m = 1$: Then $P_\varphi(t) = t - \lambda$ and $V = V_\lambda = F\,v$ for all $v \neq 0_V$ (WHY). Hence setting $\mathcal{V} := (v)$ with $v \in V$, $v \neq 0_V$, one has $[\varphi]_\mathcal{V} = (\lambda) \in F^{1\times 1}$, which in Jordan normal form (WHY).
- Induction step $m \Rightarrow (m+1)$: We begin by proceeding as in the proof of Proposition 6.10, namely: Let v_1 be an eigenvector v_1 to the eigenvalue λ. Then $W := Fv_1$ is an invariant subspace of φ, and set $\overline{V} := V/W$ and consider $\overline{\varphi} : \overline{V} \to \overline{V}$. Then as in loc.cit., one has:

$$P_\varphi(t) = P_\psi(t) P_{\overline{\varphi}}(t) = (t-\lambda) P_{\overline{\varphi}}(t), \quad \text{hence } P_{\overline{\varphi}}(t) = (t-\lambda)^m \quad \text{(WHY)}.$$

Hence $\overline{\varphi} : \overline{V} \to \overline{V}$ has a unique root λ with multiplicity $m = \dim(\overline{V})$, thus satisfies condition i) from the Proposition, and $\dim(\overline{V}) = m$. Thus by the induction hypothesis (in m), there exists a basis $\overline{\mathcal{V}}$ of $\overline{V}$ such that $C := (c_{ij})_{i,j} := [\overline{\varphi}]_{\overline{\mathcal{C}}}$ is in Jordan canonical form, having a unique Jordan block, which is a λ-block $\overline{J}_\lambda \in F^{m\times m}$. Hence if $\mathcal{C} = (v_2, \dots, v_{m+1})$ is a preimage of $\overline{\mathcal{C}}$ in V, then reasoning as in the proof of Proposition 6.10, setting $\mathcal{A} := (v_1, \mathcal{C})$, it follows that

$$(*) \qquad A := [\varphi]_\mathcal{A} = \begin{pmatrix} \lambda_1 & \boldsymbol{a} \\ \mathbf{0}_m & [\overline{\varphi}]_{\overline{\mathcal{C}}} \end{pmatrix} = \begin{pmatrix} \lambda_1 & \boldsymbol{a} \\ \mathbf{0}_m & C \end{pmatrix}, \quad \mathbf{0}_m = 0_{m_F}, \quad \boldsymbol{a} = (a_{12}, \dots, a_{1\,m+1}) \in F^m \quad \text{(WHY)}.$$

Hence in order to conclude, it is enough to effectively construct an invertible matrix $E \in F^{m\times m}$ such that $E^{-1}AE$ is in Jordan canonical form (WHY). In order to do so, we will effectively construct a sequence of elementary matrices $E_1, \dots, E_s$ such $E := E_1 \dots E_s$ does the job. The number of matrices s depends on m only, and the matrices involve the entries of $\boldsymbol{a} = (a_{12}, \dots, a_{1\,m+1}) \in F^m$ in a concrete way as follows. First notice that the entries c_{ij} of C satisfy: $c_{rr} = \lambda$ for $r = 1, \dots, m+1$, and $c_{r-1\,r} = 0_F$ or $c_{r-1\,r} = 1_F$, and $c_{ij} = 0_F$ else.

Step 1. For $1 \leqslant r < m$ such that $c_{r\,r+1} = 1_F$ consider/do the following:

Recall that for all $x \in F$ one has: $E_{1r}^{-1}(x) = E_{1r}(-x)$, and multiplication by $E_{1r}(x)$ to the right replaces $\boldsymbol{\mathcal{C}}_r$ by $\boldsymbol{\mathcal{C}}_r + x\,\boldsymbol{\mathcal{C}}_1$, and leaves the other columns $\boldsymbol{\mathcal{C}}_j$ of A unchanged. Hence since $\boldsymbol{\mathcal{C}}_1 = (a_{i1})_i$ and $a_{11} = \lambda_1$, $a_{i1} = 0_F$ for $i > 1$, the matrix $AE_{1r}(x)$ is obtained from A by replacing a_{1r} by $a_{1r} + x\lambda$ (WHY). Second, multiplication by $E_{1r}(-x)$ to the left replaces $\boldsymbol{\mathcal{R}}_1$ by $\boldsymbol{\mathcal{R}}_1 - x\,\boldsymbol{\mathcal{R}}_r$ (WHY), hence replaces a_{1j} by $a_{1j} - xc_{rj}$ for $1 \leqslant j \leqslant m+1$ (WHY), and leaves the other rows $\boldsymbol{\mathcal{R}}_i$ unchanged. And since $C = (c_{kl})_{k,l}$ is a Jordan λ-block, and $c_{rr} = \lambda$, $c_{r\,r+1} = 1$, the coefficients a'_{ij} of $A' := E_{1r}(-x)AE_{1r}$ are as follows:

$$a'_{1j} = a_{1j} \quad \text{for} \quad j \neq r+1, \quad a'_{1\,r+1} = a_{1\,r+1} - x, \quad a_{ij} = a'_{ij}, \quad i > 1 \quad \text{(WHY)}.$$

Hence choosing $x := x_r := a_{1\,r+1}$ one gets that $a'_{1\,r+1} = 0_F$. Thus repeating this process for all $1 \leqslant r < m$ such that $a_{r\,r+1} = 1$, if follows that one gets an effectively constructible sequence of (at most m) elementary matrices $E_{1r}(x_r)$ such that setting $E = \prod_r E_{1r}(x_r)$, one gets that $A' := E^{-1}AE$ has:

$$(*) \qquad \text{If} \quad a_{j\,j+1} = 1_F, \quad \text{then} \quad a'_{1\,j+1} = 0_F, \quad \text{and} \quad a'_{ij} = a_{ij} \quad \text{otherwise} \quad \text{(WHY)}.$$

Hence replacing A by $A' = E^{-1}AE$, w.l.o.g. we can suppose that $\boldsymbol{a} = (a_{12}, \dots, a_{1\,m+1}) \in F^m$ satisfies $(*)$.

Case 1. $a_{12} = 0$, and let $r > 2$ be minimal such that $a_{1r+1} \neq 0_F$, thus $a_{r\,r+1} = 0_F$.

Then the rows of the matrix $A' := E(1,r)^{-1}\,A\,E(1,r)$ satisfy: $\boldsymbol{\mathcal{R}}_1 = (a'_{1j})_j$ with $a'_{11} = \lambda$ and $a'_{1j} = 0_F$ for $j > 1$; $\boldsymbol{\mathcal{R}}_r = (a'_{rj})_j$ has $a'_{rj} = a_{rj}$ for $j \leqslant r$ and $a'_{rj} = a_{rj}$ for $j > r$; and $\boldsymbol{\mathcal{R}}'_s = \boldsymbol{\mathcal{R}}_s$ for $s \neq 1, r$ (WHY). In particular, A' is of the form

$$(*) \qquad A' = \begin{pmatrix} \lambda & \mathbf{0}_m \\ \mathbf{0}_m & C' \end{pmatrix}$$

with $C' \in F^{m\times m}$ an upper triangular matrix having the diagonal entries equal to λ (WHY). Hence induction by on m, there exist a product $E' = \prod_i E'_i$ of effectively computable elementary matrices $E'_i \in F^{m\times m}$ such that $E'^{-1} C\, E'$ is in Jordan canonical form. Thus setting

$$E_i := \begin{pmatrix} 1_F & \mathbf{0}_m \\ \mathbf{0}_m & E'_i \end{pmatrix}, \quad E := \begin{pmatrix} 1_F & \mathbf{0}_m \\ \mathbf{0}_m & E' \end{pmatrix} = \prod_i E_i$$

one has that $E^{-1} A\, E$ is in Jordan canonical form (WHY).

Case 2. $a_{12} \neq 0_F$.

Let $x := a_{12} \neq 0_F$ and recall that the inverse of the (scaling) elementary matrix $E_i(x)$ is $E_i(x^{-1})$. Therefore, A and $A' := E_1(x) A E_1^{-1}(x)$ have identical rows $\boldsymbol{\mathcal{R}}_i$ for $i > 1$, and $a'_{11} = a_{11} = \lambda$, and $a'_{1j} = a_{1j}/x$ for $j > 1$. Hence $a'_{1\,r+1} = 1_F$, and $a'_{1j} \neq 0_F$ iff $a_{1j} \neq 0_F$, $1 \leqslant j \leqslant m+1$.

- Hence replacing A by A', w.l.o.g. we can suppose that $a_{12} = 1_F$.

For $x := a_{1k} \neq 0_F$ with $k > 2$, one has: The matrix $A' := A\, E_{2k}(-x)$ has columns $\boldsymbol{\mathcal{C}}'_j$ identical with the columns $\boldsymbol{\mathcal{C}}_j$ of A for $j \neq k$, and $\boldsymbol{\mathcal{C}}'_k = \boldsymbol{\mathcal{C}}_k - x\, \boldsymbol{\mathcal{C}}_2$ (WHY), i.e., setting $\boldsymbol{\mathcal{C}}'_k = (a'_{ik})_i$ has $a'_{1k} = 0_F$, $a'_{2k} = -x\lambda$, and $a'_{ik} = a_{ik}$ for $i > 2$ (WHY). Further, $A'' := E_{2k}(x)\, A'$ and A' have rows $\boldsymbol{\mathcal{R}}''_i = \boldsymbol{\mathcal{R}}'_i$ for $i \neq 2$, and $\boldsymbol{\mathcal{R}}''_2 = \boldsymbol{\mathcal{R}}'_2 + x\, \boldsymbol{\mathcal{R}}'_k$ (WHY). Hence $A'' = (a''_{ij})_{i,j}$ satisfies: $a''_{1k} = 0_F$ and $a''_{2\,k+1} = x\, a_{k\,k+1}$, and $a''_{ij} = a_{ij}$ else (WHY). Repeating this argument inductively for each $k > 2$ with $a_{1k} \neq 0_F$, we end up with a sequence of elementary matrices $E_{2k}(-a_{1k})$ such that letting E be their product, $A' := E^{-1} A\, E = (a_{ij})_{i,j}$ has rows $\boldsymbol{\mathcal{R}}'_i$ satisfying:

- $\boldsymbol{\mathcal{R}}'_1 = (a'_{1j})_j$ has $a'_{11} = \lambda$, $a'_{12} = 1_F$, $a'_{1j} = 0_F$ for $j > 2$.
- $\boldsymbol{\mathcal{R}}'_2 = (a'_{2j})_j$ has $a'_{21} = 0_F$, $a'_{22} = \lambda$, $a'_{23} = a_{23}$, $a'_{2\,j+1} \neq 0_F$ iff $a'_{1j} \neq 0_F$.
- $\boldsymbol{\mathcal{R}}'_i = \boldsymbol{\mathcal{R}}_i$ for $i > 2$.

Subcase 2a: $a_{23} = 0_F$

Subcase 2a: $a_{23} \neq 0_F$, hence $a_{23} = 1_F$ (WHY). Then repeating the process above, but using $E_{3k}(x)$ for $x = a'_{2k} \neq 0_F$, $k > 3$, we get a matrix A'' having rows $\boldsymbol{\mathcal{R}}''_j$ satisfying $\boldsymbol{\mathcal{R}}''_1 = \boldsymbol{\mathcal{R}}'_1$, $\boldsymbol{\mathcal{R}}''_2 = (a''_{2j})_j$ with $a''_{2j} = a'_{2j}$, $1 \leqslant j \leqslant 3$, $a''_{2j} = 0_F$ for $j > 3$, and $\boldsymbol{\mathcal{R}}''_i = \boldsymbol{\mathcal{R}}'_i$ for $i > 3$.
consider $A' := E^{-1}(1,r)\, A'\, E(1,r)$, and recall that $E(1,r)^{-1} = E(1,r)$. Then multiplying A to the left by $E(1,r)$ interchanges the rows $\boldsymbol{\mathcal{R}}_1$ and $\boldsymbol{\mathcal{R}}_r$, and multiplying $A'\, E(1,r)$ to the left by $E(1,r)$ interchanges the columns $\boldsymbol{\mathcal{C}}_1$ and $\boldsymbol{\mathcal{C}}_r$. Hence conclude that Recall that $A = [\varphi]_{\mathcal{A}} = \begin{pmatrix} \lambda_1 & \boldsymbol{a} \\ \mathbf{0}_m & C \end{pmatrix}$, with C in Jordan normal form with Jordan blocks $J_{\lambda_1} \in F^{m_1-1\times m_1-1}$, $J_{\lambda_i} \in F^{m_i\times m_i}$ for $1 < i \leqslant n$. Setting $m' := m - m_1$, we rewrite A as follows:

$$A := \begin{pmatrix} \lambda_1 & \boldsymbol{a} \\ \mathbf{0}_m & C \end{pmatrix} = \begin{pmatrix} \widetilde{J}_{\lambda_1} & \mathbf{0}_{m_1\times m'} \\ \mathbf{0}_{m'\times m_1} & C' \end{pmatrix}$$

□

8. Real/Complex linear algebra

8.1. Normed vector spaces.

8.2. Spaces with inner product.

8.3. Othogonal/unitary transformations.

8.4. Normal operators.

8.5. The least squares method.

9. **Quadratic Forms & Bilinear Forms** (revisited)

9.1. **Bilinear symmetric forms vs Quadratic forms.**

9.2. **Diagonalization.**

9.3. **Sylvester's law of inertia.**

9.4. **Positive/negative (semi)definite forms, etc.** ...

E-mail address: pop@math.upenn.edu
URL: http://math.penn.edu/~pop