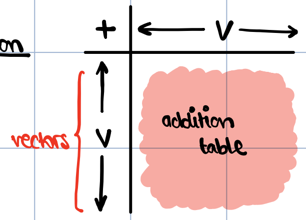
**Vector Spaces** — a type of number system obeying a set of axioms

- You need a field to define your vector space over:

$V$ over $F$ is a vector space $V = \{\ldots\}$ (a set of symbols) such that we have 2 composition laws:
elements are called vectors

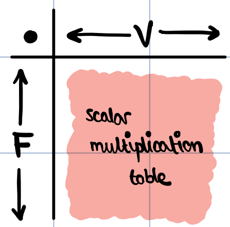
(1) Addition




obeying the axioms of an Abelian Group

$\Rightarrow$ the 0-vector exists, i.e. $0_V \in V$.
(we also have $0_F \in F$)

(2) Scalar multiplication




satisfying:

(a) Closure: $\forall c \in F, \forall v \in V$, $cv \in V$ (result is in the vector space over $F$)

(b) Associativity: $\forall c_1, c_2 \in F, \forall v \in V$, $\underbrace{c_1(c_2 v)}_{\text{scalar multiplication}} = \underbrace{(c_1 c_2)}_{\substack{\text{multiplication in the field} \\ \in F}} v$

(c) Multiplicative Identity: For $1_F \in F$ multiplicative identity, $\forall v \in V$, $1_F v = v$.

(d) Distributivity: $\forall c \in F, \forall v, \bar{v} \in V$, $c(v+\bar{v}) = cv + c\bar{v}$

$\forall c_1, c_2 \in F, \forall v \in V$, $(c_1 + c_2)v = c_1 v + c_2 v$

example: a field $F$ can form a vector space over itself. (Check definitions).
$(V = F)$

**Constructing $F^n$ where $n \in \mathbb{N}$** (Analogous to / Generalization of $\mathbb{R}^1$, $\mathbb{R}^2$, etc.)
vector space
we consider any abstract field instead of just analyzing $\mathbb{R}$

We need to:
- Construct a set of vectors } take arrangements of $n$ elements of $F$: $n$-tuples of elements of $F$, i.e.
$F^n = \{\underbrace{(x_1, x_2, \ldots, x_n)}_{n\text{-tuple}} : x_i \in F\}$ ANY combination
(column form: $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$) for now, rows and columns have no distinct meanings
- Define $+_V$ and $\cdot_V$ based on elements of $F$

Addition: let $v = (x_1, \ldots, x_n)$ and $\bar{v} = (\bar{x}_1, \ldots, \bar{x}_n)$

Then define $\underbrace{v + \bar{v}}_{\text{addition in } V=F^n} = (x_1 \underbrace{+}_{\text{addition in } F} \bar{x}_1, x_2 + \bar{x}_2, \ldots, x_n + \bar{x}_n)$. We can check that $+$ satisfies axioms of an abelian group.

Closure (1) $\forall v, \bar{v} \in F^n$, $v + \bar{v} \in F^n$: $v + \bar{v} = (x_1 + \bar{x}_1, \ldots, x_n + \bar{x}_n)$ and $x_i + \bar{x}_i \in F$. $\Delta$

Associative (2) $\forall v, \bar{v}, \bar{\bar{v}} \in F^n$, $(v + \bar{v}) + \bar{\bar{v}} = v + (\bar{v} + \bar{\bar{v}})$: $((x_1 + \bar{x}_1) + \bar{\bar{x}}_1, \ldots, (x_n + \bar{x}_n) + \bar{\bar{x}}_n)$
$= (x_1 + (\bar{x}_1 + \bar{\bar{x}}_1), \ldots, x_n + (\bar{x}_n + \bar{\bar{x}}_n))$ by associativity in $F$. $\Delta$

0-vector (3) $\forall v \in F^n$, $0 + v = v$: $(x_1 + 0, \ldots, x_n + 0) = (x_1, \ldots, x_n) = v$. $\Delta$

Inverses (4) Take $(x_1, \ldots, x_n)$ and $(-x_1, \ldots, -x_n)$ by choosing additive inverses for $x_i \in F$. Define $-v = (-x_1, \ldots, -x_n)$. $\Delta$

Commutative (5) $\forall v, \bar{v} \in F^n$, $v + \bar{v} = \bar{v} + v$. Trivial by commutativity of $+$ in $F$. $\Delta$

**Scalar multiplication:** Let $c \in F$ and $v \in F^n$, so $v = (x_1, \dots, x_n)$.

Then define $cv = c(x_1, \dots, x_n) = (cx_1, \dots, cx_n)$. We can check that $\cdot$ satisfies the axioms we require:

(a) $cv \in F^n$ trivial since $c \in F$ and $x_i \in F \Rightarrow cx_i \in F$, so $cv \in F^n$. $\Delta$

(b) $\forall c_1, c_2 \in F, \forall v \in F^n, c_1(c_2 v) = (c_1 c_2)v$. Also trivial: $c_1(c_2(x_1, \dots, x_n)) = c_1(c_2 x_1, \dots, c_2 x_n)$
$= (c_1(c_2 x_1), \dots, c_1(c_2 x_n))$
$= ((c_1 c_2)x_1, \dots, (c_1 c_2)x_n) = (c_1 c_2)v$. $\Delta$

(c) $1_F v = v$. Trivial by $1_F x_i = x_i \in F$.

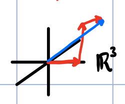
(d) Check definitions. Not very complicated. $\Delta$

$(\mathbb{R}^3 = \{(x_1, x_2, x_3) : x_i \in \mathbb{R}\}$

So we see that $F^n = \{(x_1, \dots, x_n) : x_i \in F\}$ forms a vector space over $F$. For instance, $\mathbb{R}^3$ is a vector space over $\mathbb{R}$: $F = \mathbb{R}$, $n = 3$.

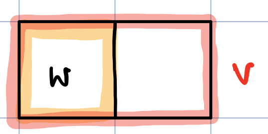
$\downarrow$

visualizable! gives intuition




## Subspaces of Vector Spaces

Let $V$ be a vector space over a field $F$.




A subspace $W$ is a subset $W \subseteq V$ satisfying the axioms of vector spaces with the inherited composition laws $+_V$ and $\cdot_V$.

We are restricting properties of $V$ to elements of $W$. We write $W \leq V$ for $W$ being subspace of $V$.

Remark: $(W, +)$ is also a subgroup of $(V, +)$ by the inherited laws.

**Theorem (Subspace Criteria):** Let $W \subseteq V$ be a subset of a vector space $V$. Then if the following are satisfied:

(1) $\forall w, \bar{w} \in W$, $w + \bar{w} \in W$;

(2) $\forall c \in F$, $\forall w \in W$, $cw \in W$;

(3) $0_V \in W$ exists;

then $W$ is a subspace of $V$.

**Proof:** We must check that:
- $+$ satisfies (abelian) group axioms,
- $\times$ satisfies vector space axioms over $F$.

We will not go over this proof. $\square$

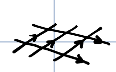
Remark: $V \leq W$ and $0 \leq V$ are always subspaces of a vector space $V$.

**Example:** Recall $\mathbb{R}^3 = \{(x_1, x_2, x_3) : x_i \in \mathbb{R}\}$ the set of real-valued 3-tuples (triplets) over $F = \mathbb{R}$.

Consider a vector $v \in \mathbb{R}^3$ and let $c \in \mathbb{R}$ vary over the real numbers. Then $cv$ forms a line.

We can then consider the subset $\{0, v\} \subset \mathbb{R}^3$. How can we turn this into a subspace of $\mathbb{R}^3$? (All lines passing through $0 \in \mathbb{R}^3$ are subspaces of $\mathbb{R}^3$)

$\{0, v\}$ • lines passing the origin $0 \in \mathbb{R}^3$; what if we add some other vector $\bar{v}$ in our set such that $\bar{v}$ is not parallel to $v$?



$\{0, v, \bar{v}\}$ • planes passing the origin $0 \in \mathbb{R}^3$; adding yet another vector $\bar{\bar{v}}$ such that $\bar{\bar{v}}$ is not in the plane forms $\mathbb{R}^3$ itself:

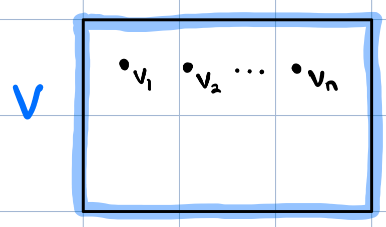
$\{0, v, \bar{v}, \bar{\bar{v}}\}$ • $\mathbb{R}^3$ itself, since $0$, $\bar{v}$ and $\bar{\bar{v}}$ end up spanning the space.

## Linear Combinations and Span

Let $V$ be a vector space over a field $F$. Let $S \subseteq V$ be a subset of our vector space such that $S = \{v_1, v_2, \dots, v_n\}$ (note that $S$ need NOT be necessarily finite)

Then we say $\underbrace{c_1 v_1}_{\in V} + \underbrace{c_2 v_2}_{\in V} \underbrace{+ \dots +}_{\text{closed under } +} \underbrace{c_n v_n}_{\in V} \in V$ where $c_i \in F$ is a linear combination of $v_1, \dots, v_n \in V$. It is a vector constructible in this form.

We define $\mathrm{Span}(S) = \{$all vectors in $V$ that are linear combinations of vectors in $S\} = \{c_1v_1 + \ldots + c_nv_n : c_i \in F,\ v_i \in S\} \subseteq V$.




Let us consider $S = \{v_1\}$ so $\mathrm{Span}(S) = \mathrm{Span}(\{v_1\}) = \{c_1v_1 : c_1 \in F\}$.

We are effectively constructing the line $\lambda v_1$ (see IB Geometry).

We know $0 \in \mathrm{Span}(S)$ always, since $0 \in F$ exists and $0v = 0 \in V$ our zero element.

Considering $S = \{v_1, v_2\}$, we have $\mathrm{Span}(\{v_1, v_2\}) = \{c_1v_1 + c_2v_2 : c_1, c_2 \in F\}$ and by choosing $c_1, c_2 = 0$, we have

$\mathrm{Span}(\{v_1\}) \subseteq \mathrm{Span}(\{v_1, v_2\})$ and $\mathrm{Span}(\{v_2\}) \subseteq \mathrm{Span}(\{v_1, v_2\})$ being subsets. (When is $\mathrm{Span}(\{v_1, v_2\}) = \mathrm{Span}(\{v_2\}) = \mathrm{Span}(\{v_1\})$?)

**Theorem:** Let $V$ be a vector space over a field $F$. Let $S \subseteq V$ be any subset of $V$.

Then $\mathrm{Span}(S) \leq V$ is a subspace of $V$.

**Proof:** $\mathrm{Span}(S) \subseteq V$ by definition, since $\mathrm{Span}(S) = \{c_1v_1 + \ldots + c_nv_n : c_i \in F\}$ and $c_1v_1 + \ldots + c_nv_n \in V$, $\forall v_i \in V$, $c_i \in F$.

Suffices to prove that $\mathrm{Span}(S)$ obeys the axioms of a vector space.

(1) Group axioms on $+$: (a) Closed under addition. Take $v, \bar{v} \in \mathrm{Span}(S)$ such that $v = c_1v_1 + \ldots + c_nv_n$, $\bar{v} = \bar{c}_1v_1 + \ldots + \bar{c}_nv_n$.

Then $v + \bar{v} = c_1v_1 + \ldots + c_nv_n \underbrace{+}_{\text{addition in } V} \bar{c}_1v_1 + \ldots + \bar{c}_nv_n = (c_1 \underbrace{+}_{\text{addition in } F} \bar{c}_1)v_1 + \ldots + (c_n + \bar{c}_n)v_n \in \mathrm{Span}(S)$

since $(c_i + \bar{c}_i) \in F$. $\Delta$

(b) Additive inverse exists: choose $-(c_i)$ for coefficients $c_i \in F$. $\Delta$

... other axioms: similar argument. $\Delta$

(2) Scalar multiplication: easy exercise, same structure. Use properties of $F$.

## Linear Independence and Dependence

Let $V$ be a vector space over a field $F$. Let $S = \{v_1, \ldots, v_n\} \subseteq V$ be a (finite) subset of $V$.

We define a linear relation on $S$ to be a linear combination of $S$ such that the result is the zero-vector, i.e.

$c_1v_1 + \ldots + c_nv_n = 0$ where $c_i \in F$. *Remark:* the linear relation where $c_1 = c_2 = \ldots = c_n = 0$ always exists. (trivial linear relation)

Let $L = \{v_1, \ldots, v_n\} \subseteq V$. We say $L$ is linearly independent over $V$ if the only linear relation of $L$ is trivial.

i.e. $L \subseteq V$ is linearly independent over $V$ $\equiv$ $c_1v_1 + \ldots + c_nv_n = 0 \Leftrightarrow c_1 = c_2 = \ldots = c_n = 0$.

**Lemma (Uniqueness of Linear Combinations):** Let $L \subseteq V$ be linearly independent ($L = \{v_1, \ldots, v_n\}$). Let $c_i, \bar{c}_i \in F$.

Then $c_1v_1 + \ldots + c_nv_n = \bar{c}_1v_1 + \ldots + \bar{c}_nv_n \Leftrightarrow c_i = \bar{c}_i$, $\forall i : (1 \leq i \leq n)$. Coefficients must match.

**Proof:** Suppose for a contradiction that $c_1v_1 + \ldots + c_nv_n = \bar{c}_1v_1 + \ldots + \bar{c}_nv_n$ but not all $c_i = \bar{c}_i$.

Then $(c_1v_1 + \ldots + c_nv_n) - (\bar{c}_1v_1 + \ldots + \bar{c}_nv_n) = 0$ but by distributivity,

$\Rightarrow (c_1 - \bar{c}_1)v_1 + \ldots + (c_n - \bar{c}_n)v_n = 0$, but not all of $c_i - \bar{c}_i = 0$, contradicting linear independence. $\lightning$

Let $S = \{v_1, \dots, v_r\} \subseteq V$. We say $S$ is linearly dependent over $V$ if there exists a non-trivial linear relation.

Since $c_1, \dots, c_r$ are not all zero, suppose $c_r \in F$ is non-zero, and $c_1v_1 + \dots + c_{r-1}v_{r-1} + c_rv_r = 0$.

Then we can write $v_r$ as a linear combination of $v_1, \dots, v_{r-1}$.

**Lemma:** Let $S = \{v_1, \dots, v_r\} \subseteq V$ be linearly dependent over $V$. Then $v_r = \bar{c}_1 v_1 + \dots + \bar{c}_{r-1} v_{r-1}$ for some $\bar{c}_i \in F$.

*Proof:* We have $c_1v_1 + \dots + c_rv_r = 0$. Without loss of generality, let $c_r \neq 0$, so $c_r \in F^\times$.

$\Rightarrow c_rv_r = -c_1v_1 - c_2v_2 - \dots - c_{r-1}v_{r-1}$ where $(-c_i) \in F$. Since $c_r^{-1} \in F$ exists,

$\Rightarrow c_rc_r^{-1}v_r = (-c_1c_r^{-1})v_1 + \dots + (-c_{r-1}c_r^{-1})v_{r-1} = v_r$. Write $\bar{c}_i = (-c_ic_r^{-1}) \in F$. $\square$

In other words, $v_n \in \mathrm{Span}(\{v_1, \dots, v_{r-1}\})$. If $S$ is linearly independent, this is not possible.
This can be proven by contradiction. Left as easy exercise.

**Lemma:** Let $S = \{v_1, \dots, v_{n-1}, 0\} \subseteq V$ containing the $0$-vector. Then $S$ is linearly dependent.

*Proof:* Consider linear relations $c_1v_1 + \dots + c_{n-1}v_{n-1} + c_n0 = 0$. We have a non-trivial linear relation by choosing

$c_1 = \dots = c_{n-1} = 0 \in F$ and $c_n \in F \setminus \{0\}$ since $c_n 0 = 0 \in V$, $\forall c_n \in F$ (including $c_n$ non-zero). $\square$

**Lemma:** Let $S = \{v_1\} \subseteq V \setminus \{0\}$ contain exactly one nonzero vector. Then $S$ is linearly independent over $V$.

*Proof:* Suffices to show that, for $c \in F$, $cv_1 = 0 \Leftrightarrow c = 0 \in F$.

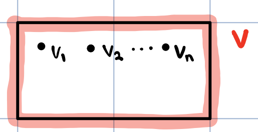
Suppose for a contradiction $c \in F^\times$ is nonzero and $cv_1 = 0$. Then $cc^{-1}v_1 = 0c^{-1} = v_1 = 0$, but we said $v_1$ nonzero. $\lightning$

In $\mathbb{R}^3$ (over $\mathbb{R}$): $\mathbb{R}^3 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : x_i \in \mathbb{R} \right\}$ where $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{pmatrix} = \begin{pmatrix} x_1 + \bar{x}_1 \\ x_2 + \bar{x}_2 \\ x_3 + \bar{x}_3 \end{pmatrix}$ and $c\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} cx_1 \\ cx_2 \\ cx_3 \end{pmatrix}$ for $c \in \mathbb{R}$.

How can we construct a linearly dependent set of vectors in $\mathbb{R}^3$?

- Put $0$ in the set;
- Put a linear combination of the other vectors in the set (e.g. $S = \left\{ \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -6 \\ -4 \\ 0 \end{pmatrix} \right\}$ is linearly dependent over $\mathbb{R}^3$.)

## Bases of Vector Spaces




Let $V$ be a vector space over a field $F$.

**Basis:** Let $B \subseteq V$ be a (finite) set $B = \{v_1, \dots, v_n\}$. If $\mathrm{Span}(B) = V$ and $B$ is linearly independent over $V$, we say $B$ is a basis for $V$.

i.e. $B \subseteq V$ satisfies:
- $\{c_1v_1 + \dots + c_nv_n : c_i \in F\} = V$; $(\mathrm{Span}(B) = V)$
- $c_1v_1 + \dots + c_nv_n = 0 \Leftrightarrow c_1 = \dots = c_n = 0$. (linear independence)

i.e. Any arbitrary vector $v \in V$ can be written uniquely as a linear combination of basis vectors $v = c_1v_1 + \dots + c_nv_n$.
(since $\mathrm{Span}(B) = V$) (since linearly independent)

Then we call $c_1, \dots, c_n$ the co-ordinates of $v$ with respect to $B$.

We can then construct the co-ordinate vector $\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$ of $v \in V$ with respect to $B$.

Let $v = c_1v_1 + \dots + c_nv_n$ (uniquely); then $v + \bar{v} = (c_1 + \bar{c}_1)v_1 + \dots + (c_n + \bar{c}_n)v_n$ with co-ordinate vector $\begin{pmatrix} c_1 + \bar{c}_1 \\ c_2 + \bar{c}_2 \\ \vdots \\ c_n + \bar{c}_n \end{pmatrix}$.
$\bar{v} = \bar{c}_1v_1 + \dots + \bar{c}_nv_n$

and $cv = c(c_1v_1) + \dots + c(c_nv_n)$ with co-ordinate vector $\begin{pmatrix} cc_1 \\ cc_2 \\ \vdots \\ cc_n \end{pmatrix}$.

What is the basis of the $0$ vector space $V_0 = \{0\}$ (over any arbitrary field $F$)? Since $0 \in V_0$ is our only vector, $B = \{0\}$. (!!) ×

$B = \{0\}$ fails linear independence, since there are non-trivial linear relations over $B$.

We define the basis $B$ of the $0$ vector space $V_0 = \{0\}$ to be the empty set $\emptyset$ $(B = \emptyset)$.

Let us look at the bases of $F^n$ over some field $F$. We have $F^n = \{(x_1, \dots, x_n) : x_i \in F\}$ the set of $n$-tuples in $F$.

"Standard Basis" is one of the bases of $F^n$, denoted by $E$, defined as $E = \{(1, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, 0, \dots, 1)\}$ $= \left\{ \underbrace{\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}}_{= e_1}, \underbrace{\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}}_{= e_2}, \dots, \underbrace{\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}}_{= e_n} \right\}$.

We can show that $\mathrm{Span}(E) = F^n$ easily. Left as exercise.

Remark: $E$ is not necessarily superior to other bases. It is just nomenclature to call it the Standard Basis.

The symbols we have given $e_1, \dots, e_n \in E$ are what makes it "special". Bases are utterly equivalent algebraically.

## Finite Dimensional Vector Spaces

Let $V$ be a vector space over a field $F$. We say $V$ is finite-dimensional if $\exists S \subseteq V$ where $|S| = n$ finite, $(S = \{v_1, \dots, v_n\})$ such that $\mathrm{Span}(S) = V$.

If we can find a finite subset of $V$ that spans $V$, then $V$ is finite-dimensional.

Lemma: Let $V$ be a finite dimensional vector space. Then a finite basis $B = \{v_1, \dots, v_n\} \subseteq F$ exists.

Proof: Since $S \subseteq V$ such that $\mathrm{Span}(S) = V$ exists, it suffices to prove that we can construct subset $B \subseteq S$ that is linearly independent.

Suppose $S$ is not linearly independent (else, we are done). Say $S = \{v_1, \dots, v_r\}$.

$\Rightarrow \exists\, c_1, \dots, c_r \in F$ not all zero, and $c_1 v_1 + \dots + c_r v_r = 0$. We want to reduce $S$ until it is linearly independent.

Suppose $c_r \neq 0 \in F$. $\Rightarrow c_r \in F^\times$ Then we can write $v_r$ as a linear combination $v_r = \underbrace{-(c_1 c_r^{-1})}_{a_1 \in F} v_1 - \dots \underbrace{-(c_{r-1} c_r^{-1})}_{a_{r-1} \in F} v_{r-1}$ and remove $v_r$ from $S$.

Let $S_1 = S \setminus \{v_r\} = \{v_1, \dots, v_{r-1}\}$. We want to prove that $S_1$ spans $V$, i.e. $\mathrm{Span}(S_1) = V$.

$\mathrm{Span}(S) = V \Rightarrow v = b_1 v_1 + \dots + b_{r-1} v_{r-1} + b_r v_r$ for any $v \in V$. Substituting our result:

$\Rightarrow v = b_1 v_1 + \dots + b_{r-1} v_{r-1} + b_r (a_1 v_1 + \dots + a_{r-1} v_{r-1}) = \underbrace{(b_r a_1 + b_1)}_{\in F} v_1 + \dots + \underbrace{(b_r a_{r-1} + b_{r-1})}_{\in F} v_{r-1}$ for arbitrary $v \in V$.
$(\Rightarrow \mathrm{Span}(S_1) = V)$.

We are done. If $S_1$ is linearly independent, $B = S_1$ exists.

Else, apply recursively. Since $|S| = r$ is finite, we will have some $n \in \mathbb{N} \cup \{0\}$ such that $S_n = B$ is a basis. $\square$
$(n = 0 \Rightarrow B = \emptyset)$

Lemma: Let $V$ be a finite-dimensional vector space with some bases $B_1, B_2 \subseteq V$. Then $|B_1| = |B_2|$. (All bases have the same size).

1) Lemma: Let $L = \{v_1, \dots, v_n\} \subseteq V$ be linearly independent over $V$. Let $S = \{\bar{v}_1, \dots, \bar{v}_m\} \subseteq V$ be a spanning set of $V$. Then $|L| \leq |S|$, i.e. $n \leq m$.

Proof: Suppose for a contradiction that $m < n$. Since $S$ spans $V$, we can write each $v_i = a_{i,1} \bar{v}_1 + a_{i,2} \bar{v}_2 + \dots + a_{i,m} \bar{v}_m \in L$

We want to find a nontrivial solution $c_1 v_1 + \dots + c_n v_n = 0$.

Substituting: $c_1 (a_{1,1} \bar{v}_1 + \dots + a_{1,m} \bar{v}_m) + \dots + c_n (a_{n,1} \bar{v}_1 + \dots + a_{n,m} \bar{v}_m)$

$= (c_1 a_{1,1}) \bar{v}_1 + \dots + (c_1 a_{1,m} \bar{v}_m) + \dots + (c_n a_{n,1}) \bar{v}_1 + \dots + (c_n a_{n,m}) \bar{v}_m$

$= \underbrace{(c_1 a_{1,1} + c_2 a_{2,1} + \dots + c_n a_{n,1})}_{k_1} \bar{v}_1 + \dots + \underbrace{(c_1 a_{1,m} + c_2 a_{2,m} + \dots + c_n a_{n,m})}_{k_m} \bar{v}_m$

If $k_1 = \dots = k_m = 0$, we are done, where not all of $c_i$ are zero.

We have $m$ equations and $n$ variables. Since $n > m$, we can always find nontrivial solutions. ↯

Proof: $B_1$ is a basis, so $B_1$ is linearly independent. $B_2$ is a basis, so $B_2$ spans $V$. Therefore $|B_1| \leq |B_2|$.

However, we also get $|B_2| \leq |B_1|$. So $|B_1| = |B_2|$. $\square$

We define the dimension of $V$ to be $\dim(V) = |B|$ where $B$ is a basis of $V$.

Lemma: Let $V$ be an $F$-vector space that is finite dimensional, and $W \subseteq V$ be a subspace. Then $W$ is finite-dimensional.

Proof: Let $B$ be a basis of $F$. Then $\mathrm{Span}(B) = F$ and $\dim(V) = |B|$ finite.

Since $W \subseteq V$, we have $B$ spanning $W$. Therefore, $B$ spans $W \Rightarrow \underset{\underset{|B_W|}{\|}}{\dim(W)} \leq |B| = \dim(V)$. $\square$

where $B_W$ is a basis of $W$.

Co-ordinate Vectors (with respect to a basis)

Let $V$ be a finite dimensional vector space over a field $F$. Then let $B = \{v_1, \dots, v_n\} \subseteq V$ be a basis (so $\dim(V) = n$ finite).

Then since we can write $v = c_1 v_1 + \dots + c_n v_n \in V$ uniquely, there is a bijection $f : V \longrightarrow F^n$, $c_1 v_1 + \dots + c_n v_n \longmapsto (c_1, \dots, c_n)$. In fact, $f$ defines an isomorphism.

Remember that $c_1, \dots, c_n$ are the co-ordinates of $v \in V$ with respect to $B$ basis.

Since we can ascribe a different isomorphism $f : V \to F^n$ for each basis, we see that there are automorphisms $\sigma : V \to V$ preserving structure.

(each basis $B_1$ of $V$ describes an automorphism $\sigma_1 \in \mathrm{Aut}(V)$).

Let $V$ over $F$ be a finite dimensional vector space. Then we have a basis $B = \{v_1, \dots, v_n\} \subseteq V$ $(\dim(V) = n)$.

Let $\bar{B} = \{\bar{v}_1, \dots, \bar{v}_n\} \subseteq V$ be another basis of $V$.

Then for a vector $v$, we have $v = c_1 v_1 + \dots + c_n v_n = \bar{c}_1 \bar{v}_1 + \dots + \bar{c}_n \bar{v}_n \in V$, where $c_1, \dots, c_n, \bar{c}_1, \dots, \bar{c}_n \in F$.

$B$ defines $\sigma : V \longrightarrow F^n$, $v \longmapsto \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ and $\bar{B}$ defines $\bar{\sigma} : V \longrightarrow F^n$, $v \longmapsto \begin{pmatrix} \bar{c}_1 \\ \vdots \\ \bar{c}_n \end{pmatrix}$ but how do $\sigma$ and $\bar{\sigma}$ relate?

We define a matrix to describe a change of basis or a transformation from $B$ to $\bar{B}$.

Rewrite $v_1, \dots, v_n \in B$ as linear combinations of $\bar{v}_1, \dots, \bar{v}_n \in \bar{B}$ first, so

$$v_1 = p_{11} \bar{v}_1 + p_{21} \bar{v}_2 + \dots + p_{n1} \bar{v}_n$$
$$v_2 = p_{12} \bar{v}_1 + p_{22} \bar{v}_2 + \dots + p_{n2} \bar{v}_2$$
$$\vdots$$
$$v_n = p_{1n} \bar{v}_1 + p_{2n} \bar{v}_2 + \dots + p_{nn} \bar{v}_n$$

where $p_{ij} \in F$

so $v = c_1(p_{11}\bar{v}_1 + \dots + p_{n1}\bar{v}_n) + c_2(p_{12}\bar{v}_1 + \dots + p_{n2}\bar{v}_n) + \dots + c_n(p_{1n}\bar{v}_1 + \dots + p_{nn}\bar{v}_n)$

$= (c_1 p_{11} + c_2 p_{12} + \dots + c_n p_{1n}) \bar{v}_1 + \dots + (c_1 p_{n1} + c_2 p_{n2} + \dots + c_n p_{nn}) \bar{v}_n$

$$\Rightarrow \begin{pmatrix} \bar{c}_1 \\ \vdots \\ \bar{c}_n \end{pmatrix} = \begin{pmatrix} c_1 p_{11} + c_2 p_{12} + \dots + c_1 p_{1n} \\ c_1 p_{21} + c_2 p_{22} + \dots + c_n p_{2n} \\ \vdots \\ c_1 p_{n1} + c_2 p_{n2} + \dots + c_n p_{nn} \end{pmatrix} =: \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & & & \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

defining a change of basis matrix.

QUESTION 2: PERHAPS A STRANGE APPROACH?

Let $V$ be the real vector space of continuous functions $f : \mathbb{R} \to \mathbb{R}$ and let $f(x) = \exp(x^2)$. Show that the set

$$\{f^{(j)}(x) \colon j \geq 0\}$$

is a (real) linearly independent subset of $V$.

Let $S = \{f^{(j)}(x) : j \geq 0\}$ denote this set. By observation, all $f^{(j)}(x) \in C^\infty(\mathbb{R})$, so $S \subseteq V$. Notice that $f(x) > 0,\ \forall x \in \mathbb{R}$. Take $a_0, ..., a_n \in \mathbb{R}$.

We want to show that $a_0 f^{(0)}(x) + ... + a_n f^{(n)}(x) = 0 \iff a_0 = ... = a_n = 0$.

**Claim:** For $k \geq 0$, we have $f(x) \mid f^{(k)}(x)$. In particular, for $k \in \mathbb{N},\ k \geq 2$, we can write

$$f^{(k)}(x) = 2(k-1)f^{(k-2)}(x) + 2xf^{(k-1)}(x)$$

*Proof.* We will proceed by induction on $k$.

Consider the base case for $k = 2$. We have

$$\begin{aligned} f^{(0)}(x) &= e^{x^2} \\ f^{(1)}(x) &= 2xe^{x^2} = 2xf^{(0)}(x) \\ f^{(2)}(x) &= 2f^{(0)}(x) + 2xf^{(1)}(x) \end{aligned}$$

so we see the equation holds for $k = 2$.

Now consider some $k \geq 2$ and suppose $f^{(k)}(x) = 2(k-1)f^{(k-2)}(x) + 2xf^{(k-1)}(x)$.

Then we get, for $n = k+1$,

$$\begin{aligned} f^{(k+1)}(x) &= \frac{\mathrm{d}}{\mathrm{d}x}(f^{(k)}(x)) \\ &= \frac{\mathrm{d}}{\mathrm{d}x}(2(k-1)f^{(k-2)}(x) + 2xf^{(k-1)}(x)) \\ &= 2(k-1)f^{(k-1)}(x) + 2f^{(k-1)}(x) + 2xf^{(k)}(x) \\ &= 2kf^{(k-1)}(x) + 2xf^{(k)}(x) \\ \implies f^{(n)}(x) &= 2(n-1)f^{(n-2)}(x) + 2xf^{(n-1)}(x). \end{aligned}$$

We are done by induction. Hence, $f(x) \mid f^{(k)}(x)$. □

**Claim:** For $k \geq 0$, the function $f^{(k)}(x)$ factors as $f^{(k)}(x) = f^{(0)}(k)P_k(x)$, where $P_k(x) \in \mathbb{R}[x]$ is an arbitrary polynomial with $\deg(P_k) = k$. Moreover, the coefficients of each $P_k(x) \in \mathbb{R}[x]$ are non-negative.

*Proof.* We perform induction on $k$.

We see the base case $k = 0$ is true with the constant polynomial $P_0(x) = 1$.

Suppose $f^{(k)}(x) = f^{(0)}(x)P_k(x)$ for some $k \geq 0$. Now we consider the case $n = k+1$.

$$\begin{aligned} f^{(k+1)}(x) &= \frac{\mathrm{d}}{\mathrm{d}x}(f^{(k)}(x)) \\ &= \frac{\mathrm{d}}{\mathrm{d}x}(f^{(0)}(x)P_k(x)) \\ &= P_k'(x)f^{(0)}(x) + 2xf^{(0)}(x)P_k(x) = f^{(0)}(x)(P_k'(x) + 2xP_k(x)). \end{aligned}$$

We are done, since $P_{k+1}(x) := (P_k'(x) + 2xP_k(x))$ satisfies $\deg(P_{k+1}) = k+1$ with non-negative coefficients. □

Now we are nearly finished. Suppose for a contradiction that $S$ is (real) linearly dependent. Then, for some $a_0, ..., a_n \in \mathbb{R}$ not all zero, we have

$$a_0 f^{(0)}(x) + ... + a_n f^{(n)}(x) = 0.$$

Without loss of generality*, say $a_n \neq 0$. Then we can rewrite as follows:

$$\begin{aligned}
& a_0 f^{(0)}(x) + ... + a_{n-1} f^{(n-1)}(x) = (-a_n) f^{(n)}(x) \\
\implies & (-a_n)^{-1}(a_0 f^{(0)}(x) + ... + a_{n-1} f^{(n-1)}(x)) = f^{(n)}(x) \\
\implies & (-a_n)^{-1}(a_0 f^{(0)} P_0(x) + ... + a_{n-1} f^{(0)} P_{n-1}(x)) = f^{(0)}(x) a_n P_n(x) \\
\implies & (-a_n)^{-1} f^{(0)}(x)(a_0 P_0(x) + ... + a_{n-1} P_{n-1}(x)) = f^{(0)}(x) a_n P_n(x)
\end{aligned}$$

but we see from above that we have a contradiction, since

$$\begin{aligned}
\implies \deg((-a_n)^{-1} f^{(0)}(x)(a_0 P_0(x) + ... + a_{n-1} P_{n-1}(x))) &= n - 1 \\
&= \deg(f^{(0)}(x) a_n P_n(x)) = n,
\end{aligned}$$

but $n - 1 \neq n$. $\perp$

Therefore, the set $\{f^{(j)}(x) \colon j \geq 0\}$ is a (real) linearly independent subset of $V$.

**Remark.** *Since a nonzero coefficient $a_i \in \mathbb{R}$ corresponding to the largest degree polynomial exists, i.e. $\{\deg(a_0 P_0(x), ..., \deg(a_n P_n(x)\}$ has a largest element.

example: Let $X \neq \emptyset$ be a nonempty set, $F$ be a field and $V = \{f : X \to F\}$.

Notice $+$ on $V$ can be defined as a pointwise addition: $\underbrace{(f+g)(x)}_{\in V} = \underbrace{f(x)}_{\in F} + \underbrace{g(x)}_{\in F}$, where $f(x), g(x) \in F$.

The additive identity $0_V \in V$ would satisfy $0_V : x \mapsto 0_F$, $\forall x \in X$.

Then $\cdot$ on $V$ could be such that $(\lambda \cdot f)(x) = \lambda \cdot f(x)$, $\forall \lambda \in F$, $f \in V$, $x \in X$.

Notice: when $X = \{1, 2, \dots, n\}$ and $V = \{f : X \to F\}$, we can fully characterize each $f \in V$ by a set $f \leftrightsquigarrow \{f(0), \dots, f(n)\}$.

then $V = \{f \mid f : X \to F\} \leftrightsquigarrow \{\{f(0), \dots, f(n)\} \mid f : X \to F\}$

Let $F$ be a field and $K/F$ be a field extension. Then $K$ is a vector space over $F$.

Let $F$ be a field. The ring $F[x]$ is an $F$-vector space.

The ring $F[x]_{\leq n} := \{f \in F[x] \mid \deg(f) \leq n,\ n \in \mathbb{N}\}$ is an $F$-vector space, and $F[x]_{\leq n} \subseteq F[x]$ (subspace).

# Linear Maps

Let $F$ be a field, and $V, W$ be $F$-vector spaces. Say $T: V \longrightarrow W$.

For $T$ to preserve the vector space structures, we need
① $T(v_1 + v_2) = T(v_1) + T(v_2)$, $\forall v_1, v_2 \in V$.
② $T(\underset{\in F}{\lambda} \underset{\in W}{v_1}) = \lambda T(v_1)$, $\forall \lambda \in F$, $\forall v_1 \in V$.
$\Rightarrow$ ③ $T(0_v) = 0_w$

We call these homomorphisms. In linear algebra,

a map $T$ with properties ① and ② are called linear maps (historically).

Given a linear map $T: V \to W$, we define $\ker(T) = \{v \in V : T(v) = 0_w\} \subseteq V$.

$\mathrm{im}(T) = \{T(v) : v \in V\} \subseteq W$.

**Lemma:** Let $T: V \to W$ be a linear map. Then $\ker(T) \subseteq V$ is a subspace of $V$ and $\mathrm{im}(T) \subseteq W$ is a subspace of $W$.

Proof: We check definitions. Trivially, $\ker(T) \subseteq V$ and $\mathrm{im}(T) \subseteq W$. We now show they are vector spaces.

We have $0_v \in \ker(T)$ since $T(0_v) = 0_w$. For $v_1, v_2 \in \ker(T)$, we have $T(v_1+v_2) = T(v_1) + T(v_2) = 0_w + 0_w$.

$\Rightarrow v_1 + v_2 \in \ker(T)$.

For $v_1 \in \ker(T)$, $\lambda \in F$, we have $T(\lambda v_1) = \lambda T(v_1) = \lambda \cdot 0_w = 0_w$.

$\Rightarrow \lambda v_1 \in \ker(T)$. △

We show similar properties for the image $\mathrm{im}(T)$. □

example(s): Let $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ be defined by $T(x, y, z) = (2x+y, 3x-z)$. Then $T$ is linear.

Let $D: \mathbb{C}[x] \longrightarrow \mathbb{C}[x]$, $p(x) \longmapsto p'(x)$ where $\mathbb{C}[x]$ is a $\mathbb{C}$-vector space. Then $D$ is a linear map.

Proof: Say $p(x), \bar{p}(x) \in \mathbb{C}[x]$. We have $D(p(x) + \bar{p}(x)) = (p(x) + \bar{p}(x))'$

$= p'(x) + \bar{p}'(x) = D(p(x)) + D(\bar{p}(x))$. △

Let $\lambda \in \mathbb{C}$. We have $D(\lambda p(x)) = (\lambda p(x))' = \lambda p'(x) = \lambda D(p(x))$. □

Consider $\mathbb{F}_{p^n}$ a field of characteristic $p$, where $p$ is prime. Then $\mathbb{F}_p \subseteq \mathbb{F}_{p^n}$ and $\mathbb{F}_{p^n}$ is an $\mathbb{F}_p$-vector space.

Define $\varphi: \mathbb{F}_{p^n} \longrightarrow \mathbb{F}_{p^n}$, $x \longmapsto x^p$ the Frobenius endomorphism. Then $\varphi$ is an $\mathbb{F}_p$-linear map.

Proof: Say $\alpha, \beta \in \mathbb{F}_{p^n}$. Then $\varphi(\alpha+\beta) = (\alpha+\beta)^p = \alpha^p + \beta^p = \varphi(\alpha) + \varphi(\beta)$. $(\mathrm{char}(\mathbb{F}_{p^n}) = p)$. △

Let $\lambda \in \mathbb{F}_p$. Then $\varphi(\lambda\alpha) = (\lambda\alpha)^p = \lambda^p \alpha^p = \lambda \alpha^p = \lambda \varphi(\alpha)$.

(since $\lambda \in \mathbb{F}_p \Rightarrow o(\lambda) = p-1$, or $\lambda^{p-1} = 1 \in \mathbb{F}_p$.) △

**Prop.:** All the automorphisms $\sigma \in \mathrm{Aut}_{\mathbb{F}_p}(\mathbb{F}_{p^n})$ are $\mathbb{F}_p$-linear maps.

Proof: Notice that $\mathrm{Aut}_{\mathbb{F}_p}(\mathbb{F}_{p^n}) = \langle \varphi \rangle$ where $\varphi(x) = x^p$ is the Frobenius map, i.e. $\varphi$ is a generator of $\mathrm{Aut}_{\mathbb{F}_p}(\mathbb{F}_{p^n})$.

We know $\varphi$ is $\mathbb{F}_p$-linear from above. It suffices to show that $\varphi \circ \varphi$ is also $\mathbb{F}_p$-linear.

**Lemma (Composition):** Let $U$, $V$ and $W$ be $F$-vector spaces, where $F$ is a field. Then:

1) If $T: V \to W$ and $S: W \to U$ are $F$-linear, then so is $S \circ T: V \to U$.

2) If $T: V \to W$ is $F$-linear and bijective, then the inverse $T^{-1}: W \to V$ is also $F$-linear.

3) The identity map $\mathrm{id}_V: V \to V$, $x \mapsto x$ is $F$-linear.

**Proof (1):** Let $v_1, v_2 \in V$. Then we have:

$$S \circ T(v_1 + v_2) = S(T(v_1 + v_2)) = S(T(v_1) + T(v_2)) \quad \text{by linearity of } T;$$

$$= S(T(v_1)) + S(T(v_2)) = S \circ T(v_1) + S \circ T(v_2) \quad \text{by linearity of } S.$$

Say $\lambda \in F$. By linearity, $S \circ T(\lambda v_1) = S(T(\lambda v_1)) = S(\lambda T(v_1)) = \lambda S(T(v_1)) = \lambda S \circ T(v_1)$. △

**(2):** Consider $v_1, v_2 \in W$. We show that $T^{-1}(v_1 + v_2) = T^{-1}(v_1) + T^{-1}(v_2)$:

$$T(T^{-1}(v_1 + v_2)) = v_1 + v_2 = T(T^{-1}(v_1)) + T(T^{-1}(v_2)) \overset{\text{linearity}}{=} T(T^{-1}(v_1) + T^{-1}(v_2)); \text{ since } T \text{ bijective: done.}$$

Now take $\lambda \in F$. By the same argument:

$$T(T^{-1}(\lambda v_1)) = \lambda v_1 = \lambda T(T^{-1}(v_1)) \overset{\text{linearity}}{=} T(\lambda T^{-1}(v_1)); \text{ done.}$$ △

**(3):** Trivially. △

Notice that the composition of linear maps may not be commutative.

Consider $T: \mathbb{C}[x] \to \mathbb{C}[x]$, $p(x) \mapsto p'(x)$ and $S: \mathbb{C}[x] \to \mathbb{C}[x]$, $p(x) \mapsto x\,p(x)$. We show they are $\mathbb{C}$-linear but $S \circ T \neq T \circ S$:

**$\mathbb{C}$-linearity:** $T$: See example on previous page.

$S$: We want to show, for $p(x), \bar{p}(x) \in \mathbb{C}[x]$, that additive linearity holds:

$$S(p(x) + \bar{p}(x)) = x(p(x) + \bar{p}(x)) = x p(x) + x\bar{p}(x) = S(p(x)) + S(\bar{p}(x)) \quad \text{by distributivity in } \mathbb{C}[x].$$

For $\lambda \in \mathbb{C}$, we have $S(\lambda p(x)) = x(\lambda p(x)) = \lambda(x p(x)) = \lambda S(p(x))$ by commutativity in $\mathbb{C}[x]$. △

**$S \circ T \neq T \circ S$:** Say $p(x) \in \mathbb{C}[x]$. We see that $S \circ T(p(x)) = S(T(p(x))) = S(p'(x)) = x p'(x)$, but

$$T \circ S(p(x)) = T(S(p(x))) = T(x p(x)) = x p'(x) + p(x) \neq x p'(x).$$ △

Remark: Notice that $S \circ T - T \circ S = \mathrm{id}_{\mathbb{C}[x]}$ the identity map.

We say a linear map $T: V \to W$ is an isomorphism between vector spaces if $T$ is bijective.

We say $V \cong W$ if an isomorphism exists between $V$ and $W$. Notice that
1) $V \cong V$;
2) If $V \cong W$, then $W \cong V$;
3) If $V \cong W \wedge W \cong U$, then $V \cong U$.

i.e. $\cong$ forms an equivalence relation of classes.

example: Consider $F^{n+1}$ and $F[x]_{\leq n}$ as $F$-vector spaces. Then $F^{n+1} \cong F[x]_{\leq n}$.

Proof: We have $F^{n+1} = \{(a_0, \ldots, a_n) : a_i \in F\}$ and $F[x]_{\leq n} = \{a_0 + a_1 x + \ldots + a_n x^n : a_i \in F\}$.

Define $\varphi : F[x]_{\leq n} \longrightarrow F^{n+1}$, $a_0 + \ldots + a_n x^n \longmapsto (a_0, \ldots, a_n)$ a map. We want to show $\varphi$ is $F$-linear and bijective.

F-linearity: Say $p(x), \bar{p}(x) \in F[x]_{\leq n}$, and $\lambda \in F$. Let $p(x) = a_0 + \ldots + a_n x^n$ and $\bar{p}(x) = \bar{a}_0 + \ldots + \bar{a}_n x^n$.

$$\begin{aligned}\varphi(p(x) + \bar{p}(x)) &= \varphi(a_0 + \ldots + a_n x^n + \bar{a}_0 + \ldots + \bar{a}_n x^n) \\ &= \varphi((a_0 + \bar{a}_0) + \ldots + (a_n + \bar{a}_n) x^n) = (a_0 + \bar{a}_0, \ldots, a_n + \bar{a}_n) = (a_0, \ldots, a_n) + (\bar{a}_0, \ldots, \bar{a}_n) \\ &= \varphi(p(x)) + \varphi(\bar{p}(x)).\end{aligned}$$

$$\begin{aligned}\varphi(\lambda p(x)) &= \varphi(\lambda(a_0 + \ldots + a_n x^n)) = \varphi(\lambda a_0 + \ldots + \lambda a_n x^n) \\ &= (\lambda a_0, \ldots, \lambda a_n) = \lambda \cdot (a_0, \ldots, a_n) = \lambda \varphi(p(x)). \quad \triangle\end{aligned}$$

Bijective: left as an exercise. $\square$

Let $V$ be an $F$-vector space with subspaces $W_1, W_2, \ldots, W_n \leq V$.

Define the sum of subspaces $W_i$ to be $W_1 + W_2 + \ldots + W_n = \{w_1 + w_2 + \ldots + w_n : w_i \in W_i\} \subseteq V$.

Theorem: Let $V$ be a vector space with subspaces $W_1, \ldots, W_n$, over some field $F$. Then $W_1 + \ldots + W_n \subseteq V$ is a subspace of $V$.

Remark: for arbitrarily chosen $W_1, \ldots, W_n$.

Proof: Let $S = W_1 + \ldots + W_n \subseteq V$.

1) For $v, \bar{v} \in S$, we have $v = w_1 + \ldots + w_n$ and $\bar{v} = \bar{w}_1 + \ldots + \bar{w}_n$ for $w_i, \bar{w}_i \in W_i$.

Then $v + \bar{v} = \underbrace{(w_1 + \bar{w}_1)}_{\in W_1} + \ldots + \underbrace{(w_n + \bar{w}_n)}_{\in W_n}$ since addition is closed in each $W_i$, so $v + \bar{v} \in S$. $\triangle$

2) For $v \in S$, $c \in F$, we have $v = w_1 + \ldots + w_n \Rightarrow cv = c(w_1 + \ldots + w_n)$

$= \underbrace{cw_1}_{\in W_1} + \ldots + \underbrace{cw_n}_{\in W_n}$ so $cv \in S$. $\square$

Consider $W_1, \ldots, W_n \leq V$ and the sum $W_1 + \ldots + W_n = \{w_1 + \ldots + w_n : w_i \in W_i\}$. If they are independent subspaces, there is only one solution to $w_1 + \ldots + w_n = 0$ which is the trivial solution, i.e. $w_1 = \ldots = w_n = 0 \in V$.

Then if $v \in W_1 + \ldots + W_n$, there is only one sum $w_1 + \ldots + w_n$ with $w_i \in W_i$ such that $v = w_1 + \ldots + w_n$.

Lemma: Let $W_1, \ldots, W_n \leq V$ be independent subspaces. Then, for $v \in W_1 + \ldots + W_n$, there is only one combination of $w_i \in W_i$ such that $v = w_1 + \ldots + w_n$.

Proof: By contradiction. Assume distinct solutions exist, then find a non-trivial solution to $w_1 + \ldots + w_n = 0$. Exercise (easy). $\square$

Let $W_1, W_2 \subseteq V$ be subspaces.

Lemma: $W_1$ and $W_2$ are independent $\iff W_1 \cap W_2 = \{0\}$.

Proof: ($\Leftarrow$) Let $W_1, W_2 \subseteq V$ such that $W_1 \cap W_2 = \{0\}$. Then for $\omega_1 + \omega_2$, we need $\omega_1 + \omega_2 = 0 \Rightarrow \omega_1 = \omega_2 = 0$.

If $\omega_1 = 0$, then $\omega_1 + \omega_2 \neq 0$, so say $\omega_1 \neq 0$, $\omega_2 \neq 0$. Then $\omega_1 + \omega_2 = 0 \Rightarrow \omega_1 = -\omega_2$ contradiction, since it implies $\omega_2 \in W_1$. $\omega_2 \in W_2$ $\lightning$

($\Rightarrow$): Say $W_1$ and $W_2$ are independent. Then $\omega_1 + \omega_2 = 0 \Rightarrow \omega_1 + \omega_2 = 0$.

Suppose for a contradiction that $v \neq 0$, such that $v \in W_1 \cap W_2$. Then $(-v) \in W_1 \cap W_2$. Construct $v + (-v) = 0$ non-trivial. $\lightning$

If $W_1, \dots, W_n$ are independent subspaces, we write the direct sum $W_1 \oplus \dots \oplus W_n$ using the $\oplus$ symbol.

We formally define the direct sum of spaces $V$ and $W$ to be $V \oplus W = \{(v, w) : v \in V, w \in W\}$.

To continue and explain some prior results, we need the following:

1) (Axiom) Every vector space $V$ has a basis. ($\iff$ Axiom of Choice)

2) If $V$ is a vector space, either all bases for $V$ are infinite, or $\exists n \in \mathbb{Z}_{\geq 0}$ such that all bases have size $n$.

We define the notion(s) of partially ordered sets $(X, \leq)$, chains, extremal/extremum elements, etc.

From Zermelo-Fraenkel Theory,

Zorn's Lemma ($\iff$ Axiom of Choice): A partially ordered set containing upper bounds for every chain (i.e. for every totally ordered subset), necessarily contains at least one maximal element.

Equivalencies: Hahn-Banach Theorem (functional analysis)

Tychonoff's Theorem (topology): every product of compact spaces is compact.

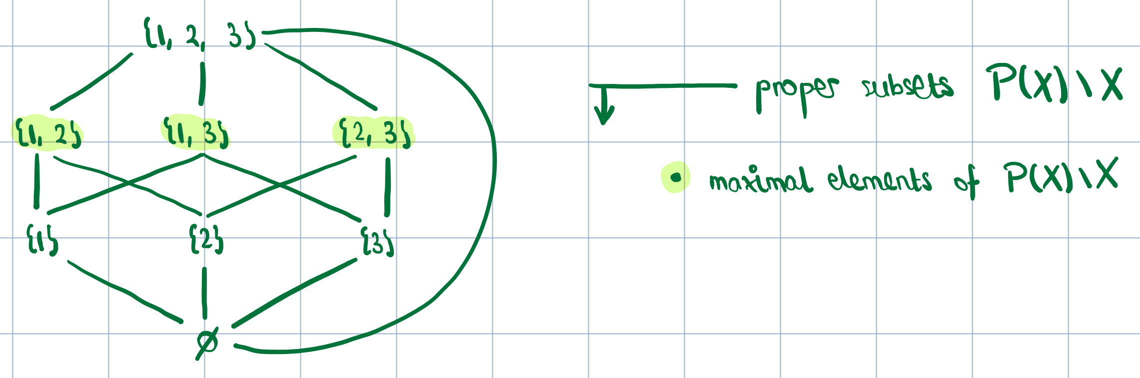
(abstract algebra): in a ring, every proper ideal is contained in a maximal ideal.

every field has an algebraic closure.

Theorem: Let $V$ be a vector space, $S \subseteq V$ be linearly independent, and $T \subseteq V$ such that $\mathrm{Span}(T) = V$. Then there exists a basis $S \subseteq B \subseteq T$. In particular, every linearly independent $S \subseteq V$ can be extended; every spanning set $T \subseteq V$ can be refined; to form a basis for $V$.

Remark: Consider the set $X = \{1, 2, 3\}$. Then $P(X)$ the power set is a partially ordered set.

Graphically:



**Proof (Existence of a Basis):** Let $\mathcal{U}$ be the set of all linearly independent subsets $U \subseteq V$ with $S \subseteq U \subseteq T$.

$\mathcal{U} = \{U \subseteq V : S \subseteq U \subseteq T, \; U \text{ linearly independent}\}$.

We can view $\mathcal{U}$ as a poset $(\mathcal{U}, \subseteq)$ with containment.

1 Every chain in $\mathcal{U}$ has an upper bound.

Notice that a chain of $\mathcal{U}$ is just a collection of sets $\{U_\lambda\}_{\lambda \in X}$ where $X$ is an index set, such that:

a) $S \subseteq U_\lambda \subseteq T, \quad \forall \lambda \in X$;

b) $U_\lambda$ is linearly independent, $\forall \lambda \in X$;

c) $\forall \alpha, \beta \in X$, either $U_\alpha \subseteq U_\beta$ or $U_\beta \subseteq U_\alpha$.

1.1 Let $U = \bigcup_{\lambda \in X} U_\lambda$. Then $U \in \mathcal{U}$ and $U$ is an upper bound for our chain.

It is clear, by definition, that $U$ is an upper bound for $\{U_\lambda\}_{\lambda \in X}$. We show now that $U \in \mathcal{U}$.

a) $U \subseteq T$: We have each $U_\lambda \subseteq T$, so $\bigcup_{\lambda \in X} U_\lambda = U \subseteq T \Rightarrow S \subseteq U \subseteq T$.

b) $U$ is linearly independent:

★ If a subset $Y \subseteq V$ is linearly dependent, then $\exists Y_0 \subseteq Y$ with $|Y_0| < \infty$ that is linearly dependent.

With ★, we show that $U$ is linearly independent.

Suppose for a contradiction that $U$ is linearly dependent. By ★, $\exists \{v_1, \dots, v_n\} \subseteq V$ that is linearly dependent.

But since $U = \bigcup_{\lambda \in X} U_\lambda$, we have some $\lambda_i$ for each $i = 1, 2, \dots, n$ such that $v_i \in U_{\lambda_i}$.

Then we have $v_1 \in U_{\lambda_1}, \dots, v_n \in U_{\lambda_n}$ but since $\{U_\lambda\}_{\lambda \in X}$ is a chain, $\exists i$ such that $U_{\lambda_j} \subseteq U_{\lambda_i}$ for all $j = 1, 2, \dots, n$.

$\Rightarrow \{v_1, \dots, v_n\} \subseteq U_{\lambda_i}$ for some $i$, so $U_{\lambda_i}$ is linearly dependent. Contradiction, since $U_{\lambda_i} \in \mathcal{U}$. $\lightning$

We have shown that every chain of our poset $(\mathcal{U}, \subseteq)$, namely the chain(s) given by $\{U_\lambda\}_{\lambda \in X}$, has an upper bound.

By Zorn's Lemma, our poset $(\mathcal{U}, \subseteq)$ contains a maximal element. Call this element $B$.

2 $B$ is a basis of $V$.

Since $B \in \mathcal{U}$, we have $S \subseteq B \subseteq T$ and $B$ linearly independent. It suffices to show, then, that:

a) $\mathrm{Span}(B) = V$.

Say $\mathrm{Span}(B) \neq V$. Then $T \nsubseteq \mathrm{Span}(B)$ so $\exists t \in T$ such that $t \notin \mathrm{Span}(B)$.

But then notice $B \cup \{t\}$ is still linearly independent and $B \cup \{t\} \subseteq T$ $(\Rightarrow B \cup \{t\} \in \mathcal{U})$:

Then $B \subseteq B \cup \{t\} \in \mathcal{U}$ contradicts the maximality of $B$. $\lightning$

We have $B \in \mathcal{U}$ such that $\mathrm{Span}(B) = V$, so $B$ is a basis of $V$. $\square$

Remarks:

- if we take $S = \emptyset$, and $\mathrm{Span}(T) = V$, we get the refinement of $T$ to form $B$.
- if we take $S$ linearly independent and $T = V$, we get the extension of $S$ to form $B$.

Given a linear (endomorphic) map $T: V \to V$ from a vector space $V$ to itself, let $T^n : V \to V$ denote the $n$-fold composition of $T$. We say a linear map $T$ is nilpotent if there is some $n \geq 1$ such that $T^n(v) = 0_V$, $\forall v \in V$.

Show by induction that differentiation is a nilpotent linear map from $\mathbb{R}[x]_{\leq n}$ to itself.

Proof : Since $\{\deg(f) : f \in \mathbb{R}[x]_{\leq n}\}$ has a minimum element $0 \in \mathbb{Z}$, we can use induction.

Base case ($\deg(f) = 0$): we see $f = a_0 \in \mathbb{R}$ and $\frac{df}{dx} = 0$.

IH — Suppose $\frac{d^{k+1}}{dx^{k+1}}(f_k) = 0$ for some $k \in \mathbb{N} \cup \{0\}$ where $\deg(f_k) \leq k \leq n$ and $\frac{d^{k+1}}{dx^{k+1}} = T^{k+1}$ $(k+1)$-fold composition.

Consider $f_{k+1}$ with $\deg(f_{k+1}) \leq k+1$. If $\deg(f_{k+1}) < k+1$, we are done by the IH.

Say $\deg(f_{k+1}) = k+1$. Write $f_{k+1} = a_0 + a_1 x + \ldots + a_{k+1} x^{k+1}$ for $a_i \in \mathbb{R}$.

Then $\frac{df_{k+1}}{dx} = a_1 + 2a_2 x + \ldots + a_{k+1}(k+1)x^k$ with degree $k$, so by the IH:

$\frac{d^{k+1}}{dx^{k+1}}\left(\frac{df_{k+1}}{dx}\right) = 0$, but here we have the $(k+2)$-fold composition of $\frac{d}{dx}$. We are done. $\square$

We say a linear map $T$ is locally nilpotent if for every $v \in V$, there is some $n \geq 1$ (possibly dependent on $v$) such that $T^n(v) = 0_V$.

Show that differentiation is a locally nilpotent linear map from $\mathbb{R}[x]$ to itself, but is not nilpotent.

Proof (Sketch) : Choose $n+1$ for a given $\deg(f) = n$ for $f \in \mathbb{R}[x]$. However, a larger degree polynomial exists. $\square$

A linear map $T: V \to W$ is called invertible if it is bijective, so $T^{-1}: W \to V$ exists.

Theorem: Let $W_1, \ldots, W_n \leq V$ with bases $B_1, \ldots, B_n$ be independent subspaces of a vector space $V$. Then $B = \bigcup_{i=1}^{n} B_i$ is a basis for $V$.

Proof : Exercise.

## Linear Transformations

A linear map $T$ on a finite-dimensional vector space $V$ can be entirely characterised by knowing the transformation of the basis.

Let $V$ be a finite-dimensional vector space with basis $B = \{v_1, \ldots, v_n\}$. Let $T: V \to W$ be a linear map.

Then for any $v \in V$, we have $v = c_1 v_1 + \ldots + c_n v_n \longmapsto T(c_1 v_1 + \ldots + c_n v_n) = T(c_1 v_1) + \ldots + T(c_n v_n)$

$= c_1 T(v_1) + \ldots + c_n T(v_n)$ fully characterized.

$\in W$

Let $F$ be a field and $E$ be a field extension. If $E/F$ is a finite extension, then every $\alpha \in E$ is algebraic over $F$.

Suppose $d = [E:F]$ and $\alpha \neq 0$. Then $\{1, \alpha, \ldots, \alpha^d\}$ is a set of size $d+1$ and so is not linearly independent.

Then $c_0 + c_1\alpha + \ldots + c_d\alpha^d = 0$ for some non-trivial combination for $c_i \in F$. Then $\alpha$ is a root of the non-zero polynomial

$c_0 + c_1x + \ldots + c_dx^d$. $\square$

Let $T: V \to W$ be a linear map. Recall that $\mathrm{im}(T) \subseteq W$ and $\ker(T) \subseteq V$ are subspaces.

We define $\mathrm{rank}(T) = \dim(\mathrm{im}(T))$ and $\mathrm{nullity}(T) = \dim(\ker(T))$.

**Theorem (Rank + Nullity):**

(not necessary in general)

Let $V, W$ be finite-dimensional $F$-vector spaces and let $T: V \to W$ be linear.

Then $\dim(V) = \mathrm{rank}(T) + \mathrm{nullity}(T)$ $(= \dim(\mathrm{im}(T)) + \dim(\ker(T))$ by definition$)$.

**Proof:** Let $d = \mathrm{nullity}(T) = \dim(\ker(T))$ and let $\{v_1, \ldots, v_d\} \subseteq \ker(T)$ be a basis.

$e = \mathrm{rank}(T) = \dim(\mathrm{im}(T))$ and let $\{w_1, \ldots, w_e\} \subseteq \mathrm{im}(T)$ be a basis.

Claim: $S = \{v_1, \ldots, v_d, u_1, \ldots, u_e\} \subseteq V$ is a basis for $V$, for some vectors $u_i$.

(i) $S$ spans $V$.

Let $v \in V$ be arbitrary. Notice that if $S$ spans, then $v = \alpha_1 v_1 + \ldots + \alpha_d v_d + \beta_1 u_1 + \ldots + \beta_e u_e$

$$\Rightarrow T(v) = \alpha_1 \cancelto{0}{T(v_1)} + \ldots + \alpha_d \cancelto{0}{T(v_d)} + \beta_1 T(u_1) + \ldots + \beta_e T(u_e).$$

$$= \beta_1 T(u_1) + \ldots + \beta_e T(u_e),$$

so $\{T(u_1), \ldots, T(u_e)\}$ forms a basis for $\mathrm{im}(T)$.

Since $w_1, \ldots, w_e$ span $\mathrm{im}(T)$, we have $T(v) = \beta_1 w_1 + \ldots + \beta_e w_e$ for all $v \in V$, where $\beta_i \in F$.

$\Rightarrow v - \beta_1 u_1 - \ldots - \beta_e u_e = 0 \in \ker(T)$. Since $\{v_1, \ldots, v_d\}$ is a basis for $\ker(T)$,

$\exists \alpha_1, \ldots, \alpha_d \in F$ with $v - \beta_1 u_1 - \ldots - \beta_e u_e = \alpha_1 v_1 + \ldots + \alpha_d v_d \Rightarrow v \in \mathrm{Span}(\{u_1, \ldots, u_e, v_1, \ldots, v_d\})$. $\triangle$

(ii) $S$ is linearly independent.

Consider a linear relation $\alpha_1 v_1 + \ldots + \alpha_d v_d + \beta_1 u_1 + \ldots + \beta_e u_e = 0$.

Then $T(\alpha_1 v_1 + \ldots + \alpha_d v_d + \beta_1 u_1 + \ldots + \beta_e u_e) = T(0) = 0$.

$\Rightarrow \alpha_1 \cancelto{0}{T(v_1)} + \ldots + \alpha_d \cancelto{0}{T(v_d)} + \beta_1 w_1 + \ldots + \beta_e w_e = 0$ since $v_i \in \ker(T)$.

$\Rightarrow \beta_1 w_1 + \ldots + \beta_e w_e = 0 \Rightarrow \beta_i = 0$ since $\{w_1, \ldots, w_e\}$ is linearly independent. $\triangle$

Therefore $\{v_1, \ldots, v_d, u_1, \ldots, u_e\}$ is a basis for $V$.

$\Rightarrow \dim(V) = \mathrm{rank}(T) + \mathrm{nullity}(T)$. $\square$

Let $F$ be a field. Given $m, n \in \mathbb{N}$, we let $M_{m\times n}(F)$ denote the set of rectangular $m\times n$ arrays of the form

$$A = \underbrace{\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}}_{n \text{ columns}} \Big\} m \text{ rows} \qquad \text{with each entry } a_{ij} \in F.$$

We call $a_{ij}$ the $(i,j)$-entry of $A$ and we call $A$ an $m\times n$ matrix with entries in $F$. We write $a_{ij} = A(i,j)$.

For instance, $\begin{pmatrix} 2 & \pi & i \\ e & 3 & -6 \end{pmatrix} \in M_{2\times 3}(\mathbb{C})$.

Notice that $M_{m\times n}(F)$ is an $F$-vector space with

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & & \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{pmatrix}$$

i.e. $A(i,j) + B(i,j) = (A+B)(i,j)$ and $\lambda \cdot A(i,j) = (\lambda A)(i,j)$.

and $O_{m\times n} \in M_{m\times n}(F)$ being $\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$. We can form the standard basis as usual, with matrices.

Remark: Notation gets heavy. Maybe see Kronecker-delta.

$$\text{Denote } E_{ij} = \begin{pmatrix} 0 & 0 & \cdots & & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & & 0 \end{pmatrix} \quad (1 \text{ in the } i^{th} \text{ row, } j^{th} \text{ column}).$$

**Lemma:** $\{E_{ij} : 1 \le i \le m,\ 1 \le j \le n\}$ is a basis for $M_{m\times n}(F)$.

*Proof:* Check definitions.

**Corollary:** $M_{m\times n}(F) \cong F^{mn}$ are isomorphic as $F$-vector spaces.

*Proof:* $M_{m\times n}(F)$ has a basis of size $mn$. $\square$

When $m=n$, we can give $M_{n\times n}(F)$ a ring structure (non-abelian) with the multiplication defined below.

Let $m, n, p \in \mathbb{N}$.

We have a multiplication $\bullet$ given by $\bullet : M_{m\times n}(F) \times M_{n\times p}(F) \longrightarrow M_{m\times p}(F)$ multiplying rows:

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix} = \begin{pmatrix} \square & \square & \square \\ \square & \square & \square \end{pmatrix} \qquad = (b_1\ b_2\ b_3)\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \text{ dot product.}$$

Let $A \in M_{2\times 2}(\mathbb{R})$. Show that the sum $\exp(A) := I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots \ (= e^A)$ converges.

Given an $n \times n$ matrix $A$ with $(i,j)$-entry $a_{ij}$, we define the trace of $A$ to be the sum of the diagonal terms

$$Tr(A) = a_{11} + a_{22} + \ldots + a_{nn}$$

Let $N \in M_{n\times n}(F)$ be a matrix whose $(i,j)$-entry is $0$ whenever $i \geq j$. Show that $N^n = 0$.

Proof: (Sketch) Consider $s \in \mathbb{N}$. We show that $N^s$ has $(i,j)$-entry $0$ whenever $i \geq j+1-s$.

Base case $(s=1)$: We have $N^1 = N$ and $(i,j)$-entry $0$ when $i \geq j$ by definition.

Trial case: Let $s=2$. Then $N^2 = N \cdot N = \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ \vdots & 0 & & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix}^2$

$(1,1)$-entry $\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ a_{12} \\ \vdots \\ a_{1n} \end{pmatrix}$ $(1,2)$-entry $\begin{pmatrix} a_{12} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ a_{2n} \end{pmatrix}$ $\ldots$ we see that every entry has each product containing zero.

Suppose for some $k \in \mathbb{N}$, we have $N^k$ with $(i,j)$-entry $0$ whenever $i \geq j+1-k$.

Consider $N^{k+1}$.

We have $N^{k+1} = N \cdot N^k = \begin{pmatrix} 0 & & \\ \vdots & 0 & \\ 0 & \cdots & 0 \end{pmatrix} N^k$. nonzero elements We want $(i,j)$-entry zero whenever $i \geq j-k$.

Suppose for a contradiction that $\exists \alpha \in F$ non-zero for some $i \geq j-k$.

We have $\alpha = row_i(N) \cdot col_j(N^k) = (0 \; 0 \; \cdots \; \cancel{a_j}^{0} \cdots a_n) \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ \cancel{b_{j-1+k}}^{0} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ by the hypothesis.

So we take our product by definition and apply our hypothesis:

$$= 0 \cdot b_1 + \ldots + 0 \cdot b_j + 0 \cdot 0 + \ldots + a_j \cdot 0 + \ldots + a_n \cdot 0 = \alpha$$

but $F$ is a field $\Rightarrow$ $F$ is an integral domain $\Rightarrow \alpha = 0$. ↯

We apply inductively, and thus we are done. $\square$

Consider $\mathbb{F}_{p^d}$ and let $\mathcal{U}$ denote the set $\{A \in M_n(\mathbb{F}_{p^d}) : A(i,j) = 0$ whenever $i > j$, $A(i,i) = 1$ for $i \in \{1, 2, \ldots, n\}\}$.

Show that if $A \in \mathcal{U}$, then $A^{p^k} = I$ whenever $k \in \mathbb{N}$ such that $p^k > n$. Conclude that $\mathcal{U}$ is a finite group with size $p^{d \cdot \binom{n}{2}}$.

Suppose $A \in M_{m\times n}(F)$. For $j = 1, \ldots, n$, $Col_j(A)$ denotes the $j^{th}$ column of $A$. Thus, $Col_j(A) \in F^m$.

For $i = 1, \ldots, m$, $Row_i(A)$ denotes the $i^{th}$ row of $A$. Thus, $Row_i(A) \in F^n$.

We wish to explore the structure of matrices.

Let $\underset{R}{\cong}$ denote a ring isomorphism and $\underset{V}{\cong}$ denote a vector space isomorphism.

**Claim:** (Associativity) Consider $A \in M_{m\times n}(F)$, $B \in M_{n\times p}(F)$ and $C \in M_{p\times q}(F)$ where $m, n, p, q \in \mathbb{N}$.

Then $(A\cdot B)\cdot C = A\cdot(B\cdot C)$.

**Proof:** We show that all the entries are the same. Recall $(AB)(i,j) = \sum_{k=1}^{n} A(i,k)\, B(k,j)$.

We have

$$((A\cdot B)\cdot C)(i,j) = \sum_{k=1}^{p} \underbrace{(A\cdot B)(i,k)}_{\downarrow}\, C(k,j)$$

$$= \sum_{k=1}^{p}\left(\sum_{l=1}^{n} A(i,l)\, B(l,k)\right) C(k,j) = \sum_{k=1}^{p}\sum_{l=1}^{n} A(i,l)\, B(l,k)\, C(k,j)$$

$$(A\cdot(B\cdot C))(i,j) = \sum_{s=1}^{n} A(i,s)\cdot (B\cdot C)(s,j)$$

$\|$ equal up to arrangement

$$= \sum_{s=1}^{n} A(i,s)\left[\sum_{t=1}^{p} B(s,t)\, C(t,j)\right] = \sum_{s=1}^{n}\sum_{t=1}^{p} A(i,s)\, B(s,t)\, C(t,j).$$

Since each entry is equal, the matrices must be the same. $\square$

**Claim:** (Distributivity) •

**Theorem:** $M_n(F)$ is a ring.

**Proof:** Check definitions. $(M_n(F), +)$ satisfies group axioms since it is a vector space.

We show $(M_n(F), \cdot)$ satisfies associativity, distributivity and is closed (since $\cdot : M_n(F)\times M_n(F) \to M_n(F)$).

We choose $id_\cdot = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$ such that $id_\cdot(i,j) = 1$ when $i=j$, and $id_\cdot(i,j) = 0$ otherwise.

Then $id_\cdot \cdot A(i,j) = \sum_{k=1}^{n} id_\cdot(i,k)\, A(k,j)$ but $id_\cdot(i,k) = 0$ whenever $k \neq i$;

$= id_\cdot(i,i)\, A(i,j) = A(i,j)$. Notice also that $id_\cdot \cdot A = A \cdot id_\cdot = A$.

All the entries are the same, so the matrices are equal. $\square$

Remark: $M_1(F) \underset{R}{\cong} F$ are isomorphic, so $M_1(F)$ is a field.

$M_2(F)$ is non-commutative in general, so $M_2(F)$ is not a field.

Notice that $M_2(F)$ has many idempotent elements. If $F$ is not a finite field, then there are infinitely many.

eg. $\begin{pmatrix} 7 & e \\ -\frac{42}{e} & -6 \end{pmatrix} \in M_2(\mathbb{R})$ is idempotent.

A matrix $D \in M_n(F)$ is called diagonal if $D(i,j) = 0$ whenever $i \neq j$, so only diagonal entries can be nonzero.

$$D = \begin{pmatrix} a_{11} & \cdots & 0 \\ 0 & a_{22} & \vdots \\ \vdots & \ddots & 0 \\ 0 & \cdots & a_{nn} \end{pmatrix}.$$ Notice that multiplying diagonal matrices gives a diagonal matrix:

$$\begin{pmatrix} a_{11} & \cdots & 0 \\ 0 & a_{22} & 0 \\ \vdots & \ddots & \\ 0 & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & 0 \\ 0 & b_{22} & \vdots \\ \vdots & \ddots & 0 \\ 0 & \cdots & b_{nn} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} & \cdots & & 0 \\ 0 & a_{22}b_{22} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & & a_{nn}b_{nn} \end{pmatrix}$$ and notice this is commutative.

The set $D_n(F) := \{D \in M_n(F) : D \text{ is diagonal}\} \subseteq M_n(F)$ is a commutative subring. Notice this is not an integral domain.

In fact, $D_n(F) \underset{R}{\cong} \underbrace{F \times \ldots \times F}_{n}$.

A matrix $U \in M_n(F)$ is upper triangular if it is of the form $U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & u_{nn} \end{pmatrix}$.

all zero

It is a subring of $M_n(F)$.

**Theorem:** For $E_{a,b}, E_{c,d} \in M_n(F)$, we have $E_{a,b} \cdot E_{c,d} = \begin{cases} 0 & \text{if } b \neq c; \\ E_{a,d} & \text{if } b = c. \end{cases}$

**Proof:** Look at the $(i,j)$-entry of $E_{a,b} \cdot E_{c,d}$. $(E_{a,b} \cdot E_{c,d})(i,j) = \sum_{k=1}^{n} \underbrace{E_{a,b}(i,k)}_{=1 \Leftrightarrow \substack{a=i \\ b=k}} \underbrace{E_{c,d}(k,j)}_{=1 \Leftrightarrow \substack{c=k \\ d=j}}$

If $b \neq c$, then either $E_{a,b}(i,k) = 0$ or $E_{c,d}(k,j) = 0$.

If $b = c$, then $\sum_{k=1}^{\infty} E_{a,b}(i,k) E_{c,d}(k,j)$ with only contribution when $k = b$.

$= E_{a,b}(i,b) E_{b,d}(b,j)$ since $b = c$, and when $k = b$,

$\Rightarrow E_{a,b}(i,b) E_{b,d}(b,j) = \begin{cases} 1 & \text{when } i = a \text{ and } j = d; \\ 0 & \text{otherwise.} \end{cases}$ $\square$

*Remark:* What is $E_{i,j}^2$? We have $E_{i,j} \cdot E_{i,j} = 0$ when $i \neq j$, (nilpotent)

$= E_{i,j}$ when $i = j$. (idempotent)

## $GL_n(F)$ (general linear group over $F$).

**Idea:** $M_n(F)$ forms a ring, so its units form a group. We let $GL_n(F)$ denote this group.

A matrix $A$ is called invertible (or non-singular) if it is in $GL_n(F)$ and non-invertible (singular) otherwise.

Theorem: $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(F) \Leftrightarrow ad-bc \neq 0.$

Proof ($\Rightarrow$) Suppose for a contradiction that $ad-bc=0$ and $A \in GL_2(F)$ ($\Rightarrow A$ is invertible).

Then $\exists A^{-1} \in GL_2(F)$ such that $A^{-1}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} A^{-1} = I.$

Consider $\begin{bmatrix} b \\ -a \end{bmatrix} \in M_{2\times 1}(F).$

$$\text{We have } A^{-1}\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{bmatrix} b \\ -a \end{bmatrix}\right) \overset{\text{associative}}{=} \left(A^{-1}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)\begin{bmatrix} b \\ -a \end{bmatrix} = I \cdot \begin{bmatrix} b \\ -a \end{bmatrix} = \begin{bmatrix} b \\ -a \end{bmatrix}.$$

$$\parallel$$

$$A^{-1}\begin{bmatrix} 0 \\ cb-ad \end{bmatrix} = A^{-1}\begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0 \quad \text{since } ad-bc=0$$

$$\Rightarrow a=b=0.$$

Similarly we have $c=d=0$.

$$\Rightarrow A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \notin GL_2(F). \quad ⚡$$

($\Leftarrow$): If $ad-bc \neq 0$, consider $A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}.$

Then $A^{-1}A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ (check by calculations). $\square$

For what follows, we will think of $F^n$ as $M_{n\times 1}(F)$. Notice $M_{n\times m}(F) \cong F^{nm} \Rightarrow M_{n\times 1}(F) \cong F^n$.

Also notice if $A \in M_{m\times n}(F)$ and $v \in M_{n\times 1}(F)$, then

$$\underset{m\times n}{A} \cdot \underset{n\times 1}{v} \in M_{m\times 1}(F) \cong F^m.$$

If $A \in M_{m\times n}(F)$, we can think of $A$ as being a map $T_A : \underset{\overset{\cong}{M_{n\times 1}(F)}}{F^n} \longrightarrow \underset{\overset{\cong}{M_{m\times 1}(F)}}{F^m}$ given by

$$T_A\left(\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}\right) = A \cdot \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$ which is a linear transformation.
exercise

In particular, $T_A(v+w) = A(v+w) = Av + Aw = T_A(v) + T_A(w)$;

$\lambda T_A(v) = \lambda A(v) = (\lambda A)(v) = A(\lambda v)$.

Universal property

Let $V, W$ be vector spaces and let $B$ be a basis for $V$. We want to understand linear maps from $V$ to $W$.

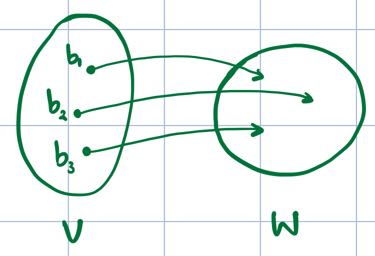
- UP: Let $f: B \to W$ be a set map.

  Then $\exists!$ linear map $T: V \to W$ such that $T|_B = f$.

Let $V$ and $W$ be vector spaces. Let $T: V \to W$ be linear. We say $T$ is a linear transformation.

When $V = W$, we say $T$ is a linear operator.

IDEA: We want to construct a linear transformation $T$ by analyzing maps from $B$ to $W$.




i.e. fix a basis $B$ for $V$ and specify what $T$ does to $B$. Extend this (uniquely) to a linear transformation.

Consider $\mathbb{R}^2 \longrightarrow \mathbb{R}^3$. Then $T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = T\left(a\begin{bmatrix} 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)$

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 5 \\ -\pi \\ \sqrt{e} \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 5 \\ -\pi \\ \sqrt{e} \end{bmatrix}$$

$$= aT\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) + bT\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = a\begin{bmatrix} 5 \\ -\pi \\ \sqrt{e} \end{bmatrix} + b\begin{bmatrix} 5 \\ -\pi \\ \sqrt{e} \end{bmatrix}.$$

i.e. we have uniquely determined our transformation.

**Theorem:** Let $V$ and $W$ be vector spaces, and let $B$ be a basis for $V$. If $f: B \to W$ is a set map, then $\exists!\ T: V \to W$ linear, such that $T|_B = f$.

uniqueness (i.e. if $S: V \to W$ linear and $S|_B = f$, then $T = S$).

**Proof:** Recall that since $B$ is a basis for $V$, if $v \in V$, we have $v = \sum_{b \in B} \lambda_b B$ uniquely, where $\lambda_b \in F$ and all but finitely many $\lambda_b$ are zero.

1 linearity $\Rightarrow$ uniqueness

Suppose $T: V \to W$ is linear and $T|_B = f$. Then $T\left(\sum_{b \in B} \lambda_b b\right) = \sum_{b \in B} \lambda_b T(b) = \sum_{b \in B} \lambda_b f(b)$ since $T|_B = f$.

Therefore $T$ linear $\Rightarrow$ $T$ unique.

2 $T$ is linear

It suffices to prove that $T(v + \lambda w) = T(v) + \lambda T(w)$, for $v, w \in V$ and $\lambda \in F$.

Let $v = \sum_{b \in B} \alpha_b b \in V$ and $w = \sum_{b \in B} \beta_b b \in V$. Then $v + \lambda w = \sum_{b \in B} (\alpha_b + \lambda \beta_b) b$.

$$\Rightarrow T(v + \lambda w) = T\left(\sum_{b \in B} (\alpha_b + \lambda\beta_b) b\right) = \sum_{b \in B} (\alpha_b + \lambda\beta_b) f(b) = \sum_{b \in B} \alpha_b f(b) + \lambda \sum_{b \in B} \beta_b f(b). \qquad \square$$

Why does this work, though?

**Theorem:** Every linear transformation $T: F^n \to F^m$ is of the form $T(v) = Av \in F^m$, where $A \in M_{m\times n}(F)$, $v \in F^n \cong M_{n\times 1}(F)$.

**Proof:** Consider the standard basis for $F^n$ given by

$$e_1 = \begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}, \quad \dots, \quad e_n = \begin{bmatrix}0\\\vdots\\0\\1\end{bmatrix}. \quad \text{Let } T(e_j) = \begin{bmatrix}a_{1j}\\a_{2j}\\\vdots\\a_{mj}\end{bmatrix}$$

be the result of $T$ on some $e_j \in E_{F^n}$ ← standard basis. (where $a_{ij} \in F$).

$$\text{Let } A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix} = \begin{pmatrix} | & | & & | \\ T(e_1) & T(e_2) & \cdots & T(e_n) \\ | & | & & | \end{pmatrix}.$$

Then $Ae_j = A\begin{bmatrix}0\\\vdots\\1\\\vdots\\0\end{bmatrix} = T(e_j)$. Therefore $(v \mapsto Av)\big|_{E_{F^n}} = T\big|_B$, i.e. both maps agree on a basis.

Therefore, $T = (v \mapsto Av)$.

Find the matrix of the linear operator $T: \mathbb{R}^2 \to \mathbb{R}^2$ with $T$ rotating points anti-clockwise by $\frac{\pi}{4}$ about $(0,0)$.

Well, all we need to know is what $T$ does to a basis!

Choose the standard basis $\left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\right\}$. We see $T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{bmatrix}$ and $T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}-\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{bmatrix}$.

where $\begin{bmatrix}1\\0\end{bmatrix}$ goes

Then we define $A := \begin{pmatrix}\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix}$. Generally, $Av = \begin{pmatrix}a & 0\\0 & b\end{pmatrix}\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}ax\\by\end{bmatrix}$.

where $\begin{bmatrix}0\\1\end{bmatrix}$ goes

We say that the matrix $A = \begin{pmatrix} | & & | \\ T(e_1) & \cdots & T(e_n) \\ | & & | \end{pmatrix}$ is the matrix of $T$ with respect to the standard basis.

We'd like to extend this idea to understand linear transformations between abstract finite-dimensional vector spaces.

Let $V = \mathbb{R}[x]_{\leq 2}$ and $W = \mathbb{R}[x]_{\leq 3}$. Let $T: V \longrightarrow W$, $p(x) \longmapsto \int_0^x p(t)\,dt$ be a linear transformation.

Notice $V \cong \mathbb{R}^3$ and $W \cong \mathbb{R}^4$. Let $B = (1, x, x^2) \subseteq V$ be a basis. Take these as ordered lists.

$\mathcal{C} = (1, x, x^2, x^3) \subseteq W$ be a basis.

So we can use $B$ and $\mathcal{C}$ to give elements of $V$ and $W$ coordinates.

Let $\mathrm{Aff}(\mathbb{R})$ be the set of maps $f: \mathbb{R} \to \mathbb{R}$, $x \longmapsto ax+b$ $(a \neq 0)$. Show that $\mathrm{Aff}(\mathbb{R})$ is a group, and is isomorphic to a subgroup of $GL_2(\mathbb{R})$.

Let $\underset{\substack{\dim(V) \\ \| \\ n}}{T : V} \longrightarrow \underset{\substack{\dim(W) \\ \| \\ m}}{W}$ be a linear map.

An ordered basis $\mathcal{B}$ of $V$ is a list $(v_1, \dots, v_n) \subseteq V$ of distinct vectors such that $\{v_1, \dots, v_n\} \subseteq V$ is a basis.

We can use ordered bases to endow vector spaces with coordinates.

Given $v \in V$, write $v = c_1 v_1 + \dots + c_n v_n$ uniquely using our basis, with $c_i \in F$.

Then make coordinate vector $[v]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in F^n$ where order matters.

**Theorem (Coordinate Vectors):** Let $V$ be a finite-dimensional vector space with $\dim(V) = n$. Let $\mathcal{B}$ be an ordered basis of $V$.

Then the map $\begin{aligned} \varphi : V &\longrightarrow F^n \\ v &\longmapsto [v]_{\mathcal{B}} \end{aligned}$ is an isomorphism.

**Proof:** Trivial by Universal Property, just analyze what $\varphi$ does to $\mathcal{B}$.

Let us proceed manually as an exercise. Let $\mathcal{B} = (v_1, \dots, v_n)$.

1 Linearity: $\varphi(v + \lambda w) = \varphi(v) + \lambda \varphi(w)$.

Let $v = c_1 v_1 + \dots + c_n v_n \in V$ and $w = d_1 v_1 + \dots + d_n c_n \in V$.

Then $[v]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ and $[w]_{\mathcal{B}} = \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} \Rightarrow [v]_{\mathcal{B}} + \lambda [w]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \lambda \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}$.

We also have $v + \lambda w = (c_1 + \lambda d_1) v_1 + \dots + (c_n + \lambda d_n) v_n$

So $[v + \lambda w]_{\mathcal{B}} = \underbrace{\begin{bmatrix} c_1 + \lambda d_1 \\ \vdots \\ c_n + \lambda d_n \end{bmatrix}}_{\in F^n} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \lambda \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}$ as required. $\square$

2 Isomorphism: Use Rank+Nullity Theorem (i.e. First Isomorphism Theorem).

Suffices to show $\ker(\varphi) = \{0\}$. Trivial. $\square$ $\square$

Let $V$ and $W$ be vector spaces. We can draw a commutative diagram as shown below, with $\mathcal{B} \subseteq V$ and $\mathcal{C} \subseteq W$ bases.
$\dim(V) = n$ $\dim(W) = m$

$$\begin{tikzcd}
\underset{\text{ordered basis}}{\mathcal{B} \subseteq} \overset{n\text{-dim}}{V} \arrow[r, "T"] \arrow[d, "\wr", "v \mapsto {[v]_{\mathcal{B}}}"'] & \overset{m\text{-dim}}{W} \underset{\text{ordered basis}}{\supseteq \mathcal{C}} \arrow[d, "\wr"', "w \mapsto {[w]_{\mathcal{C}}}"] \\
F^n \arrow[r, "A"'] & F^m
\end{tikzcd}$$

$\underline{A \in M_{n \times m}(F)}$
unique matrix

Commuting relates to the red and blue paths being identical, i.e.

$[T(v)]_{\mathcal{C}} = A \cdot [v]_{\mathcal{B}}, \ \forall v \in V.$)

Let $V = \mathbb{R}[x]_{\le 2}$ and $W = \mathbb{R}[x]_{\le 3}$. Let $T : V \longrightarrow W$, $p(x) \longmapsto \int_0^x p(t)\,dt$ be a linear transformation.

Notice $V \cong \mathbb{R}^3$ and $W \cong \mathbb{R}^4$. Let $\mathcal{B} = (1, x, x^2) \subseteq V$ be a basis. Take these as <u>ordered lists.</u>

$\mathcal{C} = (1, x, x^2, x^3) \subseteq W$ be a basis.

So we can use $\mathcal{B}$ and $\mathcal{C}$ to give elements of $V$ and $W$ coordinates. What is the matrix that makes the diagram commute?

$$\begin{tikzcd}
v = a + bx + cx^2 \arrow[r, mapsto, "T"] \arrow[d, rightsquigarrow] & 0\cdot 1 + ax + \frac{b}{2}x^2 + \frac{c}{3}x^3 = w \arrow[d, rightsquigarrow] \\
(\mathbb{R}^3) \quad \begin{bmatrix} a \\ b \\ c \end{bmatrix} \arrow[r, mapsto, "A \in M_{4\times 3}(\mathbb{R})"'] & \begin{bmatrix} 0 \\ a \\ b/2 \\ c/3 \end{bmatrix} \quad (\mathbb{R}^4)
\end{tikzcd}$$

We simple analyse what $A$ does to our basis in $\mathbb{R}^3$: $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

We have
$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ 0 \\ 1/2 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1/3 \end{bmatrix} \quad \text{so} \quad A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{pmatrix} \in M_{4\times 3}(\mathbb{R}).$$

In general, the matrix making our diagram commute is called the **matrix of $T$** with respect to the bases $\mathcal{B}$ and $\mathcal{C}$,

and we write $A = [T]_{\mathcal{B},\mathcal{C}}$ (converting $\mathcal{B}$-coords into $\mathcal{C}$-coords).

If $\mathcal{B} = (b_1, \dots, b_n) \subseteq V$ and $\mathcal{C} = (c_1, \dots, c_m) \subseteq W$, then

$$A = [T]_{\mathcal{B},\mathcal{C}} = \underbrace{\begin{pmatrix} | & & | \\ [T(b_1)]_{\mathcal{B},\mathcal{C}} & \cdots & [T(b_n)]_{\mathcal{B},\mathcal{C}} \\ | & & | \end{pmatrix}}_{n \text{ columns}} \Bigg\} \, m \text{ rows}$$

**Theorem:** Let $\underset{n\text{-dim}}{V}$ and $\underset{m\text{-dim}}{W}$ be vector spaces with $\mathcal{B} = (b_1, \dots, b_n) \subseteq V$ and $\mathcal{C} = (c_1, \dots, c_m) \subseteq W$ ordered bases.

Let $T : V \to W$ be a linear transformation. Then $\forall v \in V$, $[T]_{\mathcal{B},\mathcal{C}} [v]_{\mathcal{B}} = [T(v)]_{\mathcal{C}}$ (see commutative diagram).

<u>Proof</u>: It suffices to show, by the Universal Property, that $[T]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}}$ and $[T(v)]_{\mathcal{C}}$ agree on a basis.

Consider some $b_j \in \mathcal{B}$. We have $[T]_{\mathcal{B},\mathcal{C}} [b_j]_{\mathcal{B}}$ where $b_j = 0\cdot b_1 + \dots + 0 \cdot b_{j-1} + 1 \cdot b_j + 0 \cdot b_{j+1} + \dots + 0 \cdot b_n$

$$= [T]_{\mathcal{B},\mathcal{C}} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow j^{\text{th}} \text{ row}$$

$= j^{\text{th}}$ column of $[T]_{\mathcal{B},\mathcal{C}} = [T(b_j)]_{\mathcal{C}}$ by our definition above.

Since $b_j \in \mathcal{B}$ arbitrary, we have shown that both maps agree on a basis. $\square$

## Composition

Let $\underset{\substack{n\text{-dim} \\ \mathcal{B}=(b_1,\dots,b_n)}}{V}$ $\quad \underset{\substack{m\text{-dim} \\ \zeta=(c_1,\dots,c_m)}}{W}$ $\quad \underset{\substack{p\text{-dim} \\ \mathcal{D}=(d_1,\dots,d_p)}}{U}$ and let $T: V \to W$ and $S: W \to U$ be linear transformations.

so $S \circ T : V \to U$.

Q: What's the relationship between $[S\circ T]_{\mathcal{B},\mathcal{D}}$ and $[T]_{\mathcal{B},\zeta}$ and $[S]_{\zeta,\mathcal{D}}$ ?

**Theorem:** Let the above be our setup. Then $\underset{p\times n}{[S\circ T]_{\mathcal{B},\mathcal{D}}} = \underset{p\times m}{[S]_{\zeta,\mathcal{D}}}\ \underset{m\times n}{[T]_{\mathcal{B},\zeta}}$

**Proof:** (sketch) Look at the commutative diagram!

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, leftrightarrow] \arrow[dr, phantom, "\circlearrowright"] & W \arrow[d, leftrightarrow] \\
F^n \arrow[r, "{[T]_{\mathcal{B},\zeta}}"'] & F^m
\end{tikzcd}
\qquad
\begin{tikzcd}
W \arrow[r, "S"] \arrow[d, leftrightarrow] \arrow[dr, phantom, "\circlearrowright"] & U \arrow[d, leftrightarrow] \\
F^m \arrow[r, "{[S]_{\zeta,\mathcal{D}}}"'] & F^p
\end{tikzcd}$$

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[rr, bend left, "S\circ T"] \arrow[d, leftrightarrow] \arrow[dr, phantom, "\circlearrowright"] & W \arrow[r, "S"] \arrow[d, leftrightarrow] \arrow[dr, phantom, "\circlearrowright"] & U \arrow[d, leftrightarrow] \\
F^n \arrow[r, "{[T]_{\mathcal{B},\zeta}}"'] & F^m \arrow[r, "{[S]_{\zeta,\mathcal{D}}}"'] & F^p
\end{tikzcd}$$

and... we are done! $\square$

Remark: If $A_1, A_2 \in M_{m\times n}(F)$ and $A_1 v = A_2 v$, $\forall v \in F^n$, then $A_1 = A_2$.

This is because $A_1 e_i = A_2 e_i$ for all $e_i \in E$ standard basis, or since $\Rightarrow$ $\left.\begin{array}{l} j^{th} \text{ col. of } A_1 \\ = j^{th} \text{ col of } A_2 \end{array}\right\}$ for all $j$.

**Proof:** Let $[v]_{\mathcal{B}} \in F^n$. Consider $[S]_{\zeta,\mathcal{D}} \cdot [T]_{\mathcal{B},\zeta} \cdot [v]_{\mathcal{B}}$

$$= [S]_{\zeta,\mathcal{D}} \cdot \left([T]_{\mathcal{B},\zeta} \cdot [v]_{\mathcal{B}}\right) \quad \text{(by associativity);}$$

$$= \underset{\text{converts } F^m \to F^p}{[S]_{\zeta,\mathcal{D}}} \cdot \big(\underbrace{[T(v)]_{\zeta}}_{\in F^m}\big) \quad \text{from previously;}$$

$$= [S(T(v))]_{\mathcal{D}} = [(S\circ T)(v)]_{\mathcal{D}} = [S\circ T]_{\mathcal{B},\mathcal{D}}\,[v]_{\mathcal{B}} \quad \text{for all } v \in V. \quad \square$$

Remarks: We used the following key properties: (1) $[T]_{\mathcal{B},\zeta}[v]_{\mathcal{B}} = [T(v)]_{\zeta}$ ; (2) $[S]_{\zeta,\mathcal{D}}[w]_{\zeta} = [S(w)]_{\mathcal{D}}$.

We have a special case when $V = W$.

Let $\underset{\text{basis } \mathcal{B}=(b_1,\dots,b_n)}{V}$ be a vector space and $T: \underset{\mathcal{B}}{V} \to \underset{\mathcal{B}}{V}$ be a linear operator. Let $[T]_{\mathcal{B}}$ denote $[T]_{\mathcal{B},\mathcal{B}}$.

Often, we may have another ordered basis $\zeta \subseteq V$, and we'd like to know the relationship between $[T]_{\mathcal{B}}$ and $[T]_{\zeta}$.

Consider $T: \mathbb{R}^3 \to \mathbb{R}^3$ given by $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} x+2y+3z \\ x+y+z \\ 2x+y+z \end{pmatrix}$. Let

$$\mathcal{B} = \left(\begin{array}{ccc} b_1 & b_2 & b_3 \\ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, & \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, & \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \end{array}\right),$$

$$\zeta = \left(\begin{array}{ccc} c_1 & c_2 & c_3 \\ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, & \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, & \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \end{array}\right).$$

a) Find $[T]_{\mathcal{B},\mathcal{B}}$ and $[T]_{\zeta,\zeta}$.

We want to see what $T$ does to elements of our bases $\mathcal{B}$ and $\zeta$.

$$T(b_1) = T\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right) = \begin{bmatrix}3\\1\\1\end{bmatrix} = 1\cdot b_1 + 1\cdot b_2 + 3\cdot b_3 \rightsquigarrow \begin{bmatrix}1\\1\\3\end{bmatrix}$$

$$T(b_2) = T\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right) = \begin{bmatrix}2\\1\\1\end{bmatrix} = 1\cdot b_1 + 1\cdot b_2 + 2\cdot b_3 \rightsquigarrow \begin{bmatrix}1\\1\\2\end{bmatrix}$$

$$T(b_3) = T\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = \begin{bmatrix}1\\1\\2\end{bmatrix} = 2\cdot b_1 + 1\cdot b_2 + 1\cdot b_3 \rightsquigarrow \begin{bmatrix}2\\1\\1\end{bmatrix}$$

$$\text{So } [T]_{\mathcal{B},\mathcal{B}} = \begin{pmatrix}1 & 1 & 2\\ 1 & 1 & 1\\ 3 & 2 & 1\end{pmatrix}.$$

We repeat the same for $[T]_{\mathcal{C},\mathcal{C}}$.

Let $S: V \longrightarrow V$ be the linear operator with the property $S[v]_{\mathcal{B}} = [v]_{\mathcal{C}}$ for all $v \in V$.

We call the matrix associated with $S$ the change of basis matrix from $\mathcal{B}$ to $\mathcal{C}$, and we write $S_{\mathcal{B}\to\mathcal{C}}$.

Lemma: Let $V$ be a vector space and $S$ be a change of basis matrix $\mathcal{B}\to\mathcal{C}$. Then $S_{\mathcal{B}\to\mathcal{C}} \cdot S_{\mathcal{C}\to\mathcal{B}} = S_{\mathcal{C}\to\mathcal{B}} \cdot S_{\mathcal{B}\to\mathcal{C}} = I$.

Proof: Consider $[v]_{\mathcal{B}}$ for $v \in V$. We have

$$\begin{aligned}(S_{\mathcal{B}\to\mathcal{C}} \cdot S_{\mathcal{C}\to\mathcal{B}})[v]_{\mathcal{C}} &= S_{\mathcal{B}\to\mathcal{C}}(S_{\mathcal{C}\to\mathcal{B}} \cdot [v]_{\mathcal{C}}) \quad \text{by associativity;}\\ &= S_{\mathcal{B}\to\mathcal{C}}[v]_{\mathcal{B}} \quad \text{by definition;}\\ &= [v]_{\mathcal{C}}.\end{aligned}$$

and $(S_{\mathcal{C}\to\mathcal{B}} \cdot S_{\mathcal{B}\to\mathcal{C}})[v]_{\mathcal{B}} = S_{\mathcal{C}\to\mathcal{B}}(S_{\mathcal{B}\to\mathcal{C}} \cdot [v]_{\mathcal{B}}) = S_{\mathcal{C}\to\mathcal{B}} \cdot [v]_{\mathcal{C}} = [v]_{\mathcal{B}}$.

Since the above holds for all $v \in V$, we are done. $\square$

We have $S_{\mathcal{B}\to\mathcal{C}}[v]_{\mathcal{B}} = [v]_{\mathcal{C}}$ so $\underbrace{S_{\mathcal{B}\to\mathcal{C}}[b_j]_{\mathcal{B}}}_{j^{\text{th}} \text{ column of } S_{\mathcal{B}\to\mathcal{C}}} = [b_j]_{\mathcal{C}} \Rightarrow S_{\mathcal{B}\to\mathcal{C}} = \begin{pmatrix} | & & | \\ [b_1]_{\mathcal{C}} & \cdots & [b_n]_{\mathcal{C}} \\ | & & | \end{pmatrix}$.

Notice by a commutative diagram,

$$S_{\mathcal{B}\to\mathcal{C}}[T]_{\mathcal{B}}[v]_{\mathcal{B}} = [T]_{\mathcal{C}} S_{\mathcal{B}\to\mathcal{C}}[v]_{\mathcal{B}} \Rightarrow S_{\mathcal{B}\to\mathcal{C}}[T]_{\mathcal{B}} = [T]_{\mathcal{C}} S_{\mathcal{B}\to\mathcal{C}} \quad \text{so} \quad S_{\mathcal{B}\to\mathcal{C}}[T]_{\mathcal{B}} S_{\mathcal{C}\to\mathcal{B}} = [T]_{\mathcal{C}} \quad \text{(by right-multiplying).}$$

Let $A, B \in M_n(F)$. We say that $B$ is similar to $A$ if $\exists S \in GL_n(F)$ such that $B = S^{-1}AS$.

Proposition: Similarity is an equivalence relation.

Proof: Reflexivity: $A \sim A$ by $S = I \in GL_n(F)$.

Symmetric: $B \sim A \Rightarrow B = S^{-1}AS \Rightarrow (S^{-1})^{-1}BS^{-1} = A$, so $A \sim B$.

Transitivity: $A \sim B \wedge B \sim C \Rightarrow A = S^{-1}BS = S^{-1}(T^{-1}CT)S = (S^{-1}T^{-1})C(TS)$

$$= (TS)^{-1}C(TS) \Rightarrow A \sim C. \quad \square$$

If $A \sim B$, we say that $A$ and $B$ are similar.

Given $A \in M_n(F)$, say $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$, we define the trace $\mathrm{Tr}(A) := \sum_{i=1}^{n} a_{ii}$ to be the sum of the diagonal entries.

**Theorem (Trace):** Let $A, B \in M_n(F)$. Then $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

**Proof:** By computation. We have
$$\begin{aligned}
\mathrm{Tr}(AB) &= \sum_{i=1}^{n} (AB)(i,i) \\
&= \sum_{i=1}^{n} \Big( \sum_{k=1}^{n} A(i,k)\, B(k,i) \Big) \\
&= \sum_{i=1}^{n} \sum_{k=1}^{n} A(i,k)\, B(k,i) \\
&= \sum_{k=1}^{n} \sum_{i=1}^{n} A(k,i)\, B(i,k) && \text{by switching indices;} \\
&= \sum_{k=1}^{n} \sum_{i=1}^{n} B(i,k)\, A(i,k) && \text{since entries are in } F \Rightarrow \text{commutativity;} \\
&= \sum_{k=1}^{n} \Big( \sum_{i=1}^{n} B(i,k)\, A(k,i) \Big) = \sum_{i=1}^{n} (BA)(i,i) = \mathrm{Tr}(BA). \qquad \square
\end{aligned}$$

**Corollary:** Let $A, B \in M_n(F)$ be such that $A \sim B$. Then $\mathrm{Tr}(A) = \mathrm{Tr}(B)$.

**Proof:** We have $B = SAS^{-1}$ for some $S \in GL_n(F)$.

Then $\mathrm{Tr}(B) = \mathrm{Tr}(SAS^{-1}) = \mathrm{Tr}(S(AS^{-1}))$ and by the Theorem and associativity,

$$= \mathrm{Tr}((AS^{-1})S) = \mathrm{Tr}(A(SS^{-1})) = \mathrm{Tr}(A). \qquad \square$$

Remark: Since $[T]_B \sim [T]_\zeta$, we have that $\mathrm{Tr}([T]_B) = \mathrm{Tr}([T]_\zeta)$.

Let $V$ and $W$ be $F$-vector spaces. We write $\mathrm{Hom}_F(V, W) := \{\sigma \mid \sigma : V \to W \text{ linear}\}$ for the set of $F$-linear maps from $V$ to $W$.

Q: Given $\sigma \in \mathrm{Hom}_F(V, W)$, we can find a unique matrix $[\sigma]$. Is the converse true?

**Theorem:** Let $V$ and $W$ be $F$-vector spaces. Then $\mathrm{Hom}_F(V, W) \underset{V}{\cong} M_{m \times n}(F)$. In particular, $\dim(\mathrm{Hom}_F(V, W)) = \dim(M_{m \times n}(F))$.
$\dim(V) = n \quad \dim(W) = m$

**Proposition:** $\mathrm{Hom}_F(V, W)$ is an $F$-vector space.

**Proof (Prop.):** Let $S, T \in \mathrm{Hom}_F(V, W)$. We need to show that $(\mathrm{Hom}_F(V, W), +)$ is an abelian group, and satisfies scalar multiplication.

Define $+$ with $(S+T)(v) = S(v) + T(v) \in W$.

Additive closure: We show that $(S+T)$ is linear: $(S+T)(v_1 + \lambda v_2) = S(v_1 + \lambda v_2) + T(v_1 + \lambda v_2)$

$$= S(v_1) + T(v_1) + \lambda S(v_2) + \lambda T(v_2) = (S+T)(v_1) + \lambda (S+T)(v_2).$$

$+$ commutativity: Borrowed from $W$.

And scalar multiplication is defined by $\lambda T(v) = T(\lambda v)$ trivially. $\square$

**Proof (Theorem):** Here is our setup.

$B = (b_1, \dots, b_n)$.

$\zeta = (c_1, \dots, c_m)$.

$$\begin{tikzcd}
v \arrow[d, mapsto] & B \subseteq V \arrow[r, "T"] \arrow[d] \arrow[dr, phantom, "\circlearrowright"] & W \supseteq \zeta \arrow[d] & w \arrow[d, mapsto] \\
{[v]_B} & F^n \arrow[r, "{[T]_{B,\zeta}}"'] & F^m & {[w]_\zeta}
\end{tikzcd}$$

so these should be the same.

We want to show that there is a bijective linear map $\Psi: \mathrm{Hom}_F(V, W) \longrightarrow M_{m\times n}(F)$.

Define $\Psi: \mathrm{Hom}_F(V, W) \longrightarrow M_{m\times n}(F)$. We want to show this is linear and bijective.

$$T \longmapsto [T]_{B,\gamma}$$

**Linearity:** $\Psi(S+\lambda T) = \Psi(S) + \lambda\Psi(T)$, where $S, T \in \mathrm{Hom}_F(V, W)$ and $\lambda \in F$.

We have to show $[T+\lambda S]_{B,\gamma} = [T]_{B,\gamma} + \lambda[S]_{B,\gamma}$. Recall that we defined these matrices column-wise:

$j^{th}$ column of $[Y]_{B,\gamma} = [Y(b_j)]_\gamma$

$$\Rightarrow [Y]_{B,\gamma} = \begin{pmatrix} | & & | \\ [Y(b_1)]_\gamma & \cdots & [Y(b_n)]_\gamma \\ | & & | \end{pmatrix}.$$

It suffices to show that they are equal column-wise:

$$\underbrace{[(T+\lambda S)(b_j)]_\gamma}_{j^{th}\text{ column}} = [T(b_j) + \lambda S(b_j)]_\gamma \quad \text{by definition of addition}$$

$$= [T(b_j)]_\gamma + [\lambda S(b_j)]_\gamma$$

why can we do this?

$$= [T(b_j)]_\gamma + \underbrace{\lambda[S(b_j)]_\gamma}_{j^{th}\text{ column}}. \quad \triangle$$

1) Prove by definitions.

**Bijective:** There are two approaches.

2) Use Universal Property immediately.

**Injective:** Suffices to show that $\ker(\Psi) = \{0\}$. Note that this does not immediately prove surjectivity since dimensions have not been proven equal.

Suppose $T \in \ker(\Psi)$. Then $\Psi(T) = 0 \in M_{m\times n}(F)$ the zero-matrix $\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$.

$$\Rightarrow [T]_{B,\gamma} = 0$$

$$\Rightarrow j^{th} \text{ column of } [T]_{B,\gamma} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \text{ for all } j;$$

$\Rightarrow T(b_j) = 0, \quad \forall b_j \in B.$ Therefore $T$ kills every vector $v \in V$.

so then $T(v) = 0$ is the $0$-linear map. $\triangle$

**Surjective:** Let $A \in M_{m\times n}(F)$. We want to show $A$ is mapped onto.

Simply analyse what $A$ does to $B$ and extend this uniquely to a linear map

$T_A \in \mathrm{Hom}_F(V, W)$ by the Universal Property. $\triangle$ $\square$

Notice $\dim(M_{m\times n}(F)) = mn = \dim(M_{n\times m}(F))$, so $M_{m\times n}(F) \cong M_{n\times m}(F)$.

How can we define an isomorphism explicitly?

Define the transpose (map) $\square^t: M_{m\times n}(F) \longrightarrow M_{n\times m}(F)$ by

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \longmapsto \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

That is, $A(i,j) \mapsto A(j,i)$ for each entry.

**Proposition:** The transpose $\square^t$ is linear.

*Proof:* Exercise. $\square$ Also note that it is clearly bijective.

Notice that $\square^t$ is a vector space isomorphism fixing the diagonal of a matrix.

**Theorem:** Let $A \in M_{m\times n}(F)$ and $B \in M_{n\times p}(T)$. Then $(AB)^t = B^t A^t$.

Remark: Awfully similar to inverses in non-abelian groups, namely $(AB)^{-1} = B^{-1}A^{-1}$. Also notice that the dimensions above line up.

Proof: We will show that their $(i,j)$-entries are the same. We have

$$\begin{aligned}(AB)^t(i,j) &= (AB)(j,i) \quad \text{by definition}\\ &= \sum_{k=1}^{n} A(j,k)\,B(k,i)\\ &= \sum_{k=1}^{n} B(k,i)\,A(j,k) \quad \text{by commutativity in } F\\ &= \sum_{k=1}^{n} B^t(i,k)\,A^t(k,j) = (B^tA^t)(i,j) \quad \text{as required, for all entries } (i,j). \quad \square\end{aligned}$$

**Corollary:** $(A_1A_2\dots A_d)^t = A_d^t A_{d-1}^t \dots A_1^t$ whenever the product "makes sense".

Proof: Induction, exercise. $\square$

Let $D: \mathbb{R}[x]_{\le 2} \longrightarrow \mathbb{R}[x]_{\le 1}$, $p(x) \longmapsto p'(x)$ with $\mathcal{B} = (1, x, x^2) \subseteq \mathbb{R}[x]_{\le 2}$ and $\mathcal{C} = (1, x) \subseteq \mathbb{R}[x]_{\le 1}$.

Q: find a linear map $T: \mathbb{R}[x]_{\le 1} \longrightarrow \mathbb{R}[x]_{\le 2}$ such that $[T]_{\mathcal{C},\mathcal{B}} = [D]^t_{\mathcal{B},\mathcal{C}}$.

Recall $A \in M_{m\times n}(F)$ induces a linear map $T_A \in \mathrm{Hom}_F(\underset{F^n}{V}, \underset{F^m}{W})$ such that $T_A: F^n \to F^m$, $v \mapsto A\cdot v$.

We can then redefine $\mathrm{rank}(A) := \mathrm{rank}(T_A) = \dim(\mathrm{im}(T_A))$

$\mathrm{null}(A) := \mathrm{null}(T_A) = \dim(\ker(T_A))$ so our definitions extend to matrices.

When we have $R \in M_{p\times n}(F)$ and we left-multiply $R$ by $[x_1,\dots,x_p] \in M_{1\times p}(F)$, we get

$$[x_1,\dots,x_p]\,R = [x_1,\dots,x_p]\begin{pmatrix} - \vec{r}_1 - \\ - \vec{r}_2 - \\ \vdots \\ - \vec{r}_p - \end{pmatrix} = x_1\vec{r}_1 + x_2\vec{r}_2 + \dots + x_p\vec{r}_p.$$

these are all vectors

and right multiplication works similarly, where $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in M_{n\times 1}(F)$:

$$R\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{pmatrix} | & & | \\ \vec{c}_1 & \cdots & \vec{c}_n \\ | & & | \end{pmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{c}_1 + x_2\vec{c}_2 + \dots + x_n\vec{c}_n.$$

**Corollary:** Let $A \in M_{m\times n}(F)$. Then $\mathrm{rank}(A) = \dim(\mathrm{Span}(\{\text{columns of } A\}))$.

Proof: Let $A = \begin{pmatrix} | & & | \\ c_1 & \cdots & c_n \\ | & & | \end{pmatrix}$, where $c_i \in M_{m\times 1}(F)$. We show $\mathrm{rank}(A) = \dim(\mathrm{Span}\{c_1,\dots,c_n\})$.

Recall $\mathrm{rank}(A) = \dim(\mathrm{im}(T_A))$ where $T_A \in \mathrm{Hom}_F(\underbrace{F^n, F^m}_{\text{technically } M_{n\times 1}(F),\, M_{1\times m}(F)})$ is induced by $A$.

Consider $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in M_{n\times 1}(F)$. (for arbitrary $x_i \in F$)

Then $T_A\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = A\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{pmatrix} | & & | \\ c_1 & \cdots & c_n \\ | & & | \end{pmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{c}_1 + x_2\vec{c}_2 + \dots + x_n\vec{c}_n$ so $\mathrm{im}(T_A) \subseteq \mathrm{Span}\{c_1,\dots,c_n\}$.

However, $c_j \in \text{im}(T_A)$ since $c_j = Ae_j$ where $e_j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \in M_{n\times 1}(F)$, so $\text{Span}\{c_1, \dots, c_n\} \subseteq \underbrace{\text{im}(T_A)}_{\text{subspace containing all } c_j}$
(for each $j$)

since $\text{Span}\{c_1, \dots, c_n\}$ is the smallest subspace containing all $c_j$.

$\Rightarrow \text{im}(T_A) = \text{Span}\{c_1, \dots, c_n\}$

$\Rightarrow \dim(\text{im}(T_A)) = \text{rank}(A) = \dim(\text{Span}\{c_1, \dots, c_n\})$. $\square$

If $A = \begin{pmatrix} | & & | \\ c_1 & \cdots & c_n \\ | & & | \end{pmatrix} = \begin{pmatrix} — & r_1 & — \\ & \vdots & \\ — & r_m & — \end{pmatrix} \in M_{m\times n}(F)$, then we call $\dim(\text{Span}\{c_1, \dots, c_n\})$ the column rank.
$\dim(\text{Span}\{r_1, \dots, r_m\})$ the row rank.

**Theorem:** Let $A \in M_{m\times n}(F)$. Then $\text{rank}_{row}(A) = \text{rank}_{col}(A)$.

**Lemma:** Let $A \in M_{m\times n}(F)$ and let $c_1, \dots, c_p \in M_{m\times 1}(F)$ be a basis for $\text{Span}\{\text{colums of } A\}$.
Then $\exists R \in M_{p\times n}(F)$ such that $A = \underbrace{\begin{pmatrix} | & & | \\ c_1 & \cdots & c_p \\ | & & | \end{pmatrix}}_{m\times p} \underset{p\times n}{R}$. (Note: $A$ is not $\begin{pmatrix} | & & | \\ c_1 & \cdots & c_p \\ | & & | \end{pmatrix}$ here!).

**Proof:** Let $A = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{pmatrix}$ so that $u_j$ denotes the $j^{th}$ column of $A$, and $u_j \in M_{m\times 1}(F)$.

By assumption, each $u_j \in \text{Span}\{c_1, \dots, c_p\}$ since $\{c_1, \dots, c_p\}$ forms a basis for the column span.

In particular, $\exists r_{1j}, \dots, r_{pj} \in F$ such that $u_j = r_{1j}\vec{c}_1 + r_{2j}\vec{c}_2 + \dots + r_{pj}\vec{c}_p$.

Then $\begin{pmatrix} | & & | \\ c_1 & \cdots & c_p \\ | & & | \end{pmatrix} \begin{bmatrix} r_{1j} \\ \vdots \\ r_{pj} \end{bmatrix} = u_j$ for each $j$.

Choose $R = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & & & \\ r_{p1} & r_{p2} & \cdots & r_{pn} \end{pmatrix}$ so that $\begin{pmatrix} | & & | \\ c_1 & \cdots & c_p \\ | & & | \end{pmatrix} R = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{pmatrix} = A$. $\square$

# Determinant Functions

Recall that $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(F)$ is invertible $\Leftrightarrow$ $ad - bc \neq 0 \in F$. We will now generalize.

For what follows, let

$$A = \begin{pmatrix} - \vec{r}_1 - \\ - \vec{r}_2 - \\ \vdots \\ - \vec{r}_n - \end{pmatrix} \in M_n(F) \quad \text{where } \vec{r}_i \text{ is the } i^{th} \text{ row of } A, \text{ and we may write } A = (\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n).$$

We say a function $D : M_n(F) \to F$ is *n-linear* if, when we fix all rows except $\vec{r}_i$, letting $\vec{r}_i$ vary, we obtain a function linear in $\vec{r}_i$, for every $i$.

Let $D : M_n(F) \to F$ be n-linear. We say $D$ is *alternating* if the following hold:

- whenever $A \in M_n(F)$ has two equal rows, we have $D(A) = 0$.

when $\text{char}(F) \neq 2$ $\Rightarrow$ ( • if $A \xrightarrow[r_j \to r_i]{r_i \to r_j} B$ by a row interchange, then $D(A) = -D(B)$.)

Remark: When $\text{char}(F) = 2$, then first $\not\Rightarrow$ second.

**Lemma:** Let $D : M_n(F) \to F$ be n-linear and alternating. If $A \xrightarrow[\text{operations}]{\text{elementary}} B$, then:

(1) • $D(A) = D(B)$ if $r_i \to r_i + c r_j$ $(i \neq j)$;

(2) • $D(A) = -D(B)$ if $\begin{matrix} r_i \to r_j \\ r_j \to r_i \end{matrix}$ $(i \neq j)$;

(3) • $cD(A) = D(B)$ if $r_i \to c r_i$ for $c \in F \setminus \{0\}$.

**Proof:** (2) is true by definition of alternating.

(3) is true by n-linearity. Fix all rows but $\vec{r}_i$. Then $cD\begin{pmatrix} - * - \\ - \vec{r}_i - \\ - * - \end{pmatrix} = D\begin{pmatrix} - * - \\ - c\vec{r}_i - \\ - * - \end{pmatrix}$ by n-linearity.

(1) is slightly tricky. We have

$$D\begin{pmatrix} - \vec{r}_1 - \\ \vdots \\ - \vec{r}_i + c\vec{r}_j - \\ \vdots \\ - \vec{r}_n - \end{pmatrix} = D\begin{pmatrix} - \vec{r}_1 - \\ \vdots \\ - \vec{r}_i - \\ \vdots \\ - \vec{r}_n - \end{pmatrix} + c\,\underbrace{\cancel{D\begin{pmatrix} - \vec{r}_1 - \\ \vdots \\ - \vec{r}_j - \\ \vdots \\ - \vec{r}_j - \\ \vdots \\ - \vec{r}_n - \end{pmatrix}}}_{\substack{0 \\ \text{by alternating} \\ \text{(two equal rows)}}} \quad \text{by n-linearity.} \qquad \square$$

**Corollary:** If $D$ is n-linear and alternating, and $A \to B$ are row equivalent, then $D(A) = 0 \Leftrightarrow D(B) = 0$.

**Proof:** Use the above sequence of operations. Say $A \to A_1 \to \dots \to A_n \to B$ by elementary row operations.

Then notice $D(A_i) = 0 \Leftrightarrow D(A_{i-1}) = 0$. $\square$

**Corollary:** If $D$ is n-linear and alternating and $D(I) \neq 0$, then $D(A) = 0 \Leftrightarrow A \notin GL_n(F)$.

**Proof:** If $A \in GL_n(F)$, then $D(A) \neq 0$ since $A \to I$ are row equivalent and $D(I) \neq 0$.

If $D(A) = 0$, then $A \not\to I$ so $A \notin GL_n(F)$. $\square$

We say that a map $D: M_n(F) \to F$ is a determinant function if:

1) $D$ is $n$-linear;

2) $D$ is alternating;

3) $D(I) = 1$.

We'll see that for every $n \in \mathbb{N}$, there exists a unique determinant function. We will call this the determinant.

Theorem (Existence of Determinants): There exists a determinant function for every $n \in \mathbb{N}$.

Proof: We can produce determinants inductively.

Base cases: $n=1$ and $n=2$. Done.

Now suppose $d \geq 3$ and $\exists D: M_n(F) \to F$ determinant whenever $n<d$. We show existence for $n=d$.

Let $\det: M_{d-1}(F) \to F$ be a determinant function.

Given $A \in M_{n\times m}(F)$, we define $A(i|j)$ to be the matrix obtained by deleting the $i^{th}$ row and $j^{th}$ column of $A$.

Now define $E_j: M_d(F) \to F$ for $j=1,\dots,d$ as follows, where $A = ((a_{ij}))$:

$$E_j(A) = \sum_{i=1}^{d} a_{ij} (-1)^{i+j} \det(A(i|j)).$$

We show that $E_j: M_d(F) \to F$ defines a determinant.

1) $d$-linearity: Check manually.

2) $E_j(I) = 1$: Trivial to check. Computation.

3) alternating: It suffices to show that if $C \in M_d(F)$ with $\vec{r}_s = \vec{r}_t$ $(s \neq t)$ (consecutive rows) then $E_j(C) = 0$.

Notice that $E_j(C) = \sum_{i=1}^{d} a_{ij} (-1)^{i+j} \det(C(i|j))$

has at most two nonzero terms, since if $i \neq s,t$, then $C(i|j)$ has two equal rows

$\Rightarrow \det(C(i|j)) = 0$.

Therefore consider the two possibly nonzero terms, where $i=s,t$. Since it suffices to prove the case when $\vec{r}_s, \vec{r}_t$ are consecutive (any other arrangement gives some row equivalent matrix), say $t=s+1$.

Then $E_j = a_{sj}(-1)^{s+j} \det(C(s|j)) + a_{(s+1)j}(-1)^{s+j+1} \det(C(s+1|j))$

but $C(s|j)$ and $C(s+1|j)$ are equivalent, i.e. $C(s+1|j) = C(s|j)$.

$\Rightarrow E_j = \det(C(s|j))\left(a_{sj}(-1)^{s+j} - a_{(s+1)j}(-1)^{s+j}\right)$. But since $\vec{r}_s = \vec{r}_{s+1}$, we have $a_{sj} = a_{(s+1)j}$.

$= \det(C(s|j))(-1)^{s+j}(\underbrace{a_{sj} - a_{sj}}_{0}) = 0$.

By induction, determinants exist. $\square$

**Theorem (Uniqueness of Determinants):** For $n \in \mathbb{N}$, if $D: M_n(F) \to F$ is a determinant function, it is unique.

**Remark:** We will show that the general form for a determinant is very resticted.

**Proof:** 0 Setup.

let $e_1 = (1\ 0\ \dots\ 0)$ so $e_j = (0 \dots 0\ \overset{j^{th}\text{ position}}{1}\ 0 \dots 0)$.

$\vdots$

$e_n = (0 \dots 0\ 1)$

Let $A \in M_n(F)$ and write $A = ((a_{ij}))$. Denote the $i^{th}$ row by $\vec{r}_i = (a_{i1}\ a_{i2} \dots a_{in})$.

Write $A = \begin{pmatrix} a_{11}\vec{e}_1 + a_{12}\vec{e}_2 + \dots + a_{1n}\vec{e}_n \\ \text{---}\ \vec{r}_2\ \text{---} \\ \vdots \\ \text{---}\ \vec{r}_n\ \text{---} \end{pmatrix}$. Fix $\vec{r}_2, \dots, \vec{r}_n$ and look at $D$ as $n$-linear in $\vec{r}_1$.

Then $D(A) = a_{11} D\begin{pmatrix} \vec{e}_1 \\ -\vec{r}_2- \\ \vdots \\ -\vec{r}_n- \end{pmatrix} + \dots + a_{1n} D\begin{pmatrix} \vec{e}_n \\ -\vec{r}_2- \\ \vdots \\ -\vec{r}_n- \end{pmatrix} = \sum_{j=1}^{n} a_{1j} D\begin{pmatrix} \vec{e}_j \\ -\vec{r}_2- \\ \vdots \\ -\vec{r}_n- \end{pmatrix}$.

Repeat by fixing $\vec{r}_1, \dots, \vec{r}_{i-1}, \vec{r}_{i+1}, \dots, \vec{r}_n$ and letting $\vec{r}_i$ vary for each $i$.

Continuing in this manner, we see that

$$D(A) = \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} a_{1j_1} \cdots a_{nj_n} \underbrace{D\begin{pmatrix} \vec{e}_{j_1} \\ \vdots \\ \vec{e}_{j_n} \end{pmatrix}}_{\in F}.$$

Every $n$-linear function looks like this.

Now we use the fact that $D$ is alternating.

We have $D(A) = \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} \left( \prod_{i=1}^{n} a_{ij} \right) D\begin{pmatrix} \vec{e}_{j_1} \\ \vdots \\ \vec{e}_{j_n} \end{pmatrix}$. If $D$ alternates, then we only care about pairwise distinct terms, so $j_1, \dots, j_n$ is a permutation of $\{1, 2, \dots, n\}$ (by distinctness).

Therefore $f: \{1, \dots, n\} \to \{1, \dots, n\}$ defines a bijection.
$f(i) = j$

Hence $D\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \sum_{\sigma \in S_n} \left( \prod_{i=1}^{n} a_{i\sigma(i)} \right) D\begin{pmatrix} \vec{e}_{\sigma(1)} \\ \vdots \\ \vec{e}_{\sigma(n)} \end{pmatrix}$.

1 Let $\sigma \in S_n$. Then:

1.1 We can perform a series of row interchanges to transform $\begin{pmatrix} \vec{e}_{\sigma(1)} \\ \vdots \\ \vec{e}_{\sigma(n)} \end{pmatrix} \to I$.

Induction on $n$. When $n=1$, we are done since the matrix $[1] = I \in M_1(F)$.

Assume the statement holds for $n<d$ where $d \geq 2$. We will then show it holds for $n=d$.

We have $\sigma \in S_d \Rightarrow \sigma(1), \dots, \sigma(d)$ is some rearrangement of $1, \dots, d \Rightarrow \exists i : \sigma(i) = d$.

If $i=d$, then $\begin{pmatrix} \vec{e}_{\sigma(1)} \\ \vdots \\ \vec{e}_{\sigma(d)} \end{pmatrix} = \begin{pmatrix} \vec{e}_{\sigma(1)} \\ \vdots \\ 0 \dots 0\ 1 \end{pmatrix}$ so rearrange the first $d-1$ rows to form $I$ (by the hypothesis).

If $i \neq d$, then interchange $\vec{e}_{\sigma(i)}$ with $\vec{e}_{\sigma(d)}$. We get some new permutation $\tau \in S_{d-1}$ such that

$$\begin{pmatrix} \vec{e}_{\sigma(1)} \\ \vdots \\ \vec{e}_{\sigma(d)} \end{pmatrix} \longrightarrow \begin{pmatrix} \vec{e}_{\tau(1)} \\ \vdots \\ \vec{e}_{\tau(d-1)} \\ 0 \cdots 0 1 \end{pmatrix}$$ so we apply the inductive hypothesis.

$$\longrightarrow \begin{pmatrix} -\vec{e}_{\tau(1)}- 0 \\ \vdots \\ -\vec{e}_{\tau(d-1)}- 0 \\ 0 \cdots\cdots 0 1 \end{pmatrix} = I \in M_d(F).$$

1.2 If $d, e \in \mathbb{N}$ such that $\begin{pmatrix} \vec{e}_{\sigma(1)} \\ \vdots \\ \vec{e}_{\sigma(n)} \end{pmatrix} \rightarrow I$ using either $d$ or $e$ row interchanges, then $d \equiv e \pmod 2$.

Parity is nontrivial. Let $B = \begin{pmatrix} \vec{e}_{\sigma(1)} \\ \vdots \\ \vec{e}_{\sigma(n)} \end{pmatrix}$ we suppose there are two ways of row-interchanging to $I$:

$$\begin{tikzcd}[column sep=small, row sep=small] & B_1 \arrow[r] & \cdots \arrow[r] & B_{d-1} \arrow[dr] & \\ B \arrow[ur] \arrow[dr] & & & & I. \\ & B'_1 \arrow[r] & \cdots \arrow[r] & B'_{e-1} \arrow[ur] & \end{tikzcd}$$

Then top path: $D(B) = (-1)^d\, D(I) = (-1)^d$.

bottom path: $D(B) = (-1)^e\, D(I) = (-1)^e$.

Therefore $(-1)^e = (-1)^d \Rightarrow d \equiv e \pmod 2$.

We define $\operatorname{sgn}(\sigma) = (-1)^d$ for $\sigma \in S_n$ where we can row-reduce $\begin{pmatrix} \vec{e}_{\sigma(1)} \\ \vdots \\ \vec{e}_{\sigma(n)} \end{pmatrix} \xrightarrow[\text{interchanges}]{} I$ with $d$ row swaps.

Notice that $\operatorname{sgn}(\sigma)$ is well-defined for $\sigma$ by the congruence $(-1)^d = (-1)^e$.

2 If $D: M_n(F) \to F$ is a determinant function, then $D\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \sum_{\sigma \in S_n} \left( \prod_{i=1}^{n} a_{i\sigma(i)} \right) \operatorname{sgn}(\sigma)$.

We have $D\begin{pmatrix} \vec{e}_{\sigma(1)} \\ \vdots \\ \vec{e}_{\sigma(n)} \end{pmatrix} = (-1)^d = \operatorname{sgn}(\sigma)$ from above.

We also have an expression for $D(A)$. Thus

$$D\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \sum_{\sigma \in S_n} \left( \prod_{i=1}^{n} a_{i\sigma(i)} \right) D\begin{pmatrix} \vec{e}_{\sigma(1)} \\ \vdots \\ \vec{e}_{\sigma(n)} \end{pmatrix} = \underbrace{\sum_{\sigma \in S_n} \left( \prod_{i=1}^{n} a_{i\sigma(i)} \right)}_{\text{uniquely determined}} \overset{\text{well-defined}}{\operatorname{sgn}(\sigma)}.$$

Therefore, the function $D: M_n(F) \to F$ is unique. $\square$

Given $n \in \mathbb{N}$, we define the determinant $\det: M_n(F) \to F$ by $\det(A) = \sum_{\sigma \in S_n} \left( \prod_{i=1}^{n} a_{i\sigma(i)} \right)$.

**Theorem (Multiplicity of Determinants):** Let $A, B \in M_n(F)$. Then $\det(AB) = \det(A)\det(B)$.

**Proof:** Notice $AB \notin GL_n(F) \Leftrightarrow A \notin GL_n(F)$ or $B \notin GL_n(F)$. Therefore $\det(AB) = 0 \Leftrightarrow \det(A) = 0$ or $\det(B) = 0$.

Therefore, it suffices to consider the case when $A, B \in GL_n(F)$. Notice then $\det(A) \neq 0$ and $\det(B) \neq 0$.

If $C \in GL_n(F)$, we have $\det(C) = \sum_{\sigma \in S_n} \left( \prod_{i=1}^{n} c_{i\sigma(i)} \right)$ where $C = ([c_{ij}])$.

**Claim:** Let $D : M_n(F) \to F$ be given by $D(A) = \frac{\det(AB)}{\det(B)}$. Then $D$ defines a determinant function. (I mean this is just absolutely incredible).

1 $D(I) = 1$: We have $D(I) = \frac{\det(B)}{\det(B)} = 1$.

2 $n$-linearity: Write $A = \begin{pmatrix} \vec{r}_1 \\ \vdots \\ \vec{r}_n \end{pmatrix}$. Then $AB = \begin{pmatrix} \vec{r}_1 B \\ \vdots \\ \vec{r}_n B \end{pmatrix}$. Fix all rows but $\vec{r}_i$, letting $\vec{r}_i$ vary.

Then $D(A) = \frac{\det(AB)}{\det(B)} = \det \begin{pmatrix} \vec{r}_1 B \\ \vdots \\ \vec{r}_n B \end{pmatrix} / \det(B)$ where $\det(B) \in F \setminus \{0\}$.

We view this as doing three steps on some row $\vec{r}_i$:

$$\vec{r}_i \underset{\text{linear}}{\longmapsto} \vec{r}_i B \underset{n\text{-linear}}{\longmapsto} \det \begin{pmatrix} * \\ \vec{r}_i B \\ * \end{pmatrix} \underset{\text{linear}}{\longmapsto} \det \begin{pmatrix} * \\ \vec{r}_i B \\ * \end{pmatrix} / \det(B)$$

$\Rightarrow$ $D$ is formed by a composition of linear maps, so $D$ is $n$-linear.

3 alternation: Suppose $A$ has two equal rows, say $\vec{r}_p = \vec{r}_q$, where $p \neq q$.

Then $D(A) = \frac{\det(AB)}{\det(B)} = \det \begin{pmatrix} * \\ \vec{r}_p B \\ \vec{r}_q B \\ * \end{pmatrix} / \det(B)$ but $\vec{r}_p = \vec{r}_q \Rightarrow \vec{r}_p B = \vec{r}_q B \Rightarrow \det \begin{pmatrix} * \\ \vec{r}_p B \\ \vec{r}_q B \\ * \end{pmatrix} = 0 \Rightarrow D(A) = 0$.

Therefore, $D$ is a determinant function. By uniqueness, we then have $D(A) = \det(A) = \frac{\det(AB)}{\det(B)}$ so $\det(AB) = \det(A)\det(B)$. $\square$

**Corollary:** If $\sigma, \tau \in S_n$, then $\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)$.

**Proof:** Exercise.

We define the special linear group $SL_n(F) := \{A \in GL_n(F) : \det(A) = 1\}$. This forms a subgroup of $GL_n(F)$.

Show that any $A \in GL_n(\mathbb{C})$ can be expressed in the form $\alpha A_s$, where $A_s \in SL_n(\mathbb{C})$ and $\alpha \in \mathbb{C} \setminus \{0\}$.

Let $A \in \mathcal{U}_n(F) := \{A \in M_n(F) : A \text{ is upper triangular}\}$ be written as $A = ([a_{ij}])$. Show that $\det(A) = \prod_{i=1}^{n} a_{ii}$.

Notice $sgn: \underset{\text{group}}{S_n} \longrightarrow \underset{\text{group}}{\mathbb{Z}^\times}$ satisfies $sgn(\sigma \circ \tau) = sgn(\sigma)\, sgn(\tau)$ and $sgn(id_n) = 1$, so sgn defines a group homomorphism.

By definition, $ker(sgn) = \{\sigma \in S_n : sgn(\sigma) = id_{\mathbb{Z}^\times} = 1\}$ which is a subgroup of $S_n$ under composition.

We define the alternating group $A_n := (ker(sgn), \circ)$ which is closed, contains inverses and has the identity.

Show $|A_n| = \frac{|S_n|}{2} = \frac{n!}{2}$.

**Theorem (Determinant of Transpose):** Let $A \in M_n(F)$. Then $det(A^t) = det(A)$.

**Proof:** Write $A = ((a_{ij}))$. Then $det(A) = \sum_{\sigma \in S_n} sgn(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}$ and $det(A^t) = \sum_{\sigma \in S_n} sgn(\sigma) \prod_{i=1}^{n} a_{\sigma(i) i}$.

Notice $\sigma \in S_n \Rightarrow \sigma^{-1} \in S_n$. Let $\sigma^{-1}$ such that $\sigma^{-1}(j) = i \Leftrightarrow j = \sigma(i)$. Then $a_{\sigma(i) i} \rightsquigarrow a_{j \sigma^{-1}(j)}$.

Also note that $sgn(\sigma)\, sgn(\sigma^{-1}) = sgn(id_n) = 1$ so $sgn(\sigma) = sgn(\sigma^{-1})$.

Thus we have $det(A^t) = \sum_{\sigma \in S_n} sgn(\sigma^{-1}) \prod_{j=1}^{n} a_{j\sigma^{-1}(j)}$

$$\overset{\text{relabelling}}{=} \sum_{\tau \in S_n} sgn(\tau) \prod_{j=1}^{n} a_{j\tau(j)} = det(A)$$ so we are done. □

Recall that in Theorem (Existence of Determinants) we showed that $E_j(A) = \sum_{i=1}^{n} a_{ij} (-1)^{i+j} det(A(i|j))$ is a determinant.

By uniqueness, we have $det(A) = \sum_{i=1}^{n} a_{ij} (-1)^{i+j} det(A(i|j))$ giving a recursive formula.

This process for computing determinants is called cofactor expansion.

**Theorem (Classical Adjoint):** Let $A \in GL_n(F)$ so $det(A) \neq 0$. Then $A^{-1} = \frac{1}{det(A)} \left( (-1)^{i+j} det(A(i|j)) \right)_{1 \le i,j \le n}$.

**Proof:** We consider $B = \frac{1}{det(A)} \left( (-1)^{i+j} det(A(i|j)) \right)$ and look at the $(i,j)$-entry of $AB$.

Case 1: $i = j$. We show $(AB)(i,i) = 1$.

We have $(AB)(i,i) = \sum_{k=1}^{n} A(i,k)\, B(k,i)$

$$= \sum_{k=1}^{n} A(i,k) \frac{1}{det(A)} (-1)^{i+k} det(A(i|k)).$$

We leave the rest as an exercise. □

If $dim(V) = n$ and $T: V \to V$ is a linear operator, we define

$det(T) := det([T]_{B \to B})$ where $B$ is any ordered basis of $V$. Note this is well-defined regardless of $B$. why?

Let $T: V \to V$ be a linear operator. We say that $c \in F$ is an eigenvalue of $T$ if $\exists \vec{v} \in V \setminus \{\vec{0}\}$ so $T(\vec{v}) = c\vec{v}$.

Conversely, if $T(\vec{v}) = c\vec{v}$ for $\vec{v} \in V \setminus \{\vec{0}\}$, we say $\vec{v} \in V \setminus \{\vec{0}\}$ is an eigenvector of $T$.

Remark: If $v \in V \setminus \{0\}$ is an eigenvector of $T$, then $\lambda \vec{v} \in V \setminus \{\vec{0}\}$ is an eigenvector too, $\forall \lambda \in F \setminus \{0\}$.

For $A \in M_n(F)$ and $c \in F$, we have $c \in F$ an eigenvalue of $A$ if $\exists \vec{v} \in F^n \setminus \{\vec{0}\}$ such that $A\vec{v} = c\vec{v}$.

eigenvector

Consider $V = \mathbb{R}[x]$, and $T: V \longrightarrow V$, $p(x) \longmapsto \int_0^x p(t)\,dt$ and $S: V \longrightarrow V$, $p(x) \longmapsto p'(x)$ as linear operators.

Q: What are the eigenvalues and eigenvectors of $T$ and $S$?

Notice $\deg(p) \underset{T}{\longmapsto} \deg(p)+1$ unless $p(x)=0$ which cannot be an eigenvector;

and $\deg(p) \underset{T}{\longmapsto} \deg(p)-1$ (or $-\infty$ since $\deg(0) := -\infty$).

Therefore if $T(p(x)) = cp(x) = \int_0^x p(t)\,dt$

$\Rightarrow cp'(x) = p(x)$ has no nonzero solution $p(x) \in \mathbb{R}[x]$. So $T$ has no eigenvectors $\Rightarrow$ no eigenvalues.

However if $S(p(x)) = cp(x) = p'(x)$ then $p(x) = \alpha \neq 0$ (constant) satisfies the equation with $c = 0 \in \mathbb{R}$

since $0 \cdot \alpha = (\alpha)' = 0$. Therefore $0 \in \mathbb{R}$ an eigenvalue and $p(x) = \alpha \in \mathbb{R}[x] \setminus \{0\}$ are all eigenvectors.

Let $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$. Then $A$ has no real eigenvalues, but does have $\mathbb{C}$-eigenvalues.

Suppose $c \in \mathbb{C}$ and $\begin{bmatrix} a \\ b \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix} \in M_2(\mathbb{R})$ such that $A\begin{bmatrix} a \\ b \end{bmatrix} = c\begin{bmatrix} a \\ b \end{bmatrix}$

$\Rightarrow \begin{bmatrix} b \\ -a \end{bmatrix} = \begin{bmatrix} ca \\ cb \end{bmatrix}$ but $a, b \neq 0 \Rightarrow c \neq 0$

$\Rightarrow -ba = (c-a)(c-b) \Rightarrow c^2 = -1 \Rightarrow c = i.$ $\square$

Let $A = I_n \in M_n(F)$. What are the eigenvalues/eigenvectors of $A$?

Notice $I_n\vec{v} = \vec{v}$ so $1 \in F$ eigenvalue and all $\vec{v} \in F^n \setminus \{\vec{0}\}$ eigenvector.

Let $T: V \longrightarrow V$ be an operator (notice this includes matrices).

Then if $c \in F$, we let $W_c := \{\vec{v} \in V : T(\vec{v}) = c\vec{v}\}$ and we call $W_c$ the eigenspace of $T$ associated to $c$.

**Theorem ($W_c$ subspace):** Let $c \in F$. Then $W_c$ is a subspace of $V$.

*Proof:* Check that $0 \in W_c$ and $W_c$ is closed under $+$ and $\cdot$.

Notice that $W_c$ counts $v = 0 \in V$ so $T(\vec{0}) = c\vec{0}$ for all $c \in F \Rightarrow \vec{0} \in W_c, \forall c \in F \Rightarrow W_c \neq \emptyset$.

If $v_1, v_2 \in W_c$ then $T(v_1+v_2) \overset{\text{linearity}}{=} T(v_1) + T(v_2) = cv_1 + cv_2 = c(v_1+v_2) \Rightarrow v_1 + v_2 \in W_c$.

If $\lambda \in F$ then $T(\lambda v) \overset{\text{homogeneous}}{=} \lambda T(v) = \lambda c v = c(\lambda v) \Rightarrow \lambda v \in W_c$. $\square$

*Remark:* $W_c \subseteq V$ is always a subspace, but not all $c \in F$ are necessarily eigenvalues.

Notice $c \in F$ an eigenvalue $\Leftrightarrow W_c \neq \{\vec{0}\}$ by definition;

$\Leftrightarrow \dim(W_c) \geq 1$ since $\dim(\{\vec{0}\}) = 0$.

**Theorem (Eigenspaces):** Let $T: V \to V$ be a linear operator where $\dim(V) < \infty$. Let $c \in F$. Then the following are equivalent:

(i) $c \in F$ is an eigenvalue of $T$;

(ii) The eigenspace $W_c \neq \{\vec{0}\}$;

(iii) $T - cI \in \mathrm{Hom}_F(V, V)$ is not invertible;

(iv) $\det(T - cI) = 0$.

**Proof:** (i) $\Leftrightarrow$ (ii) Already done. $\triangle$

(ii) $\Rightarrow$ (iii) Suppose $W_c \neq \{\vec{0}\}$. Then $\exists \vec{v} \neq 0$ such that $T(\vec{v}) = c\vec{v} \Leftrightarrow T(\vec{v}) - c\vec{v} = \vec{0} \Leftrightarrow (T - cI)(\vec{v}) = \vec{0}$

$\Leftrightarrow \ker(T - cI) \neq \{\vec{0}\}$

$\Leftrightarrow T - cI$ not invertible. $\triangle$

(iii) $\Rightarrow$ (iv) By definition.

(iv) $\Rightarrow$ (i) $\det(T - cI) = 0 \Rightarrow T - cI$ not invertible

$\Rightarrow \exists \vec{v} \in \ker(T - cI)$ such that $\vec{v} \neq \vec{0}$

$\Rightarrow (T - cI)(\vec{v}) = \vec{0} \Rightarrow T(\vec{v}) = c\vec{v}$ so $c \in F$ is an eigenvalue of $T$. $\triangle$ $\square$

Let $A \in M_n(F)$. We define the characteristic polynomial of $A$ by $p_A(x) = \det(xI - A)$

**Proposition:** For $A \in M_n(F)$, the polynomial $p_A(x) = \det(xI - A) \in F[x]$ with $\deg(p_A) = n$.

Furthermore, the roots of $p_A(x)$ in $F$ are precisely the eigenvalues of $A$.

Consider $A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} \in M_2(\mathbb{R})$. Then $xI - A = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} x-1 & -2 \\ 1 & x-4 \end{pmatrix}$.

$\Rightarrow \det(xI - A) = (x-1)(x-4) - (-2)(1) = x^2 - 5x + 6 = (x-3)(x-2) = p_A(x)$. Notice $\{2, 3\}$ are roots of $p_A$.

Thus $A$ has eigenvalues $2$ and $3$ in $\mathbb{R}$.

**Remark:** Notice in general, a matrix $A \in M_n(\mathbb{R})$ may have a characteristic polynomial $p_A(x)$ with no real roots (at least when $n$ even).

However, we will always have $n$ eigenvalues in $\mathbb{C}$ (by Fundamental Theorem of Algebra).

In general, we can consider the splitting field of $p_A(x)$ to have this property.

**Theorem (Characteristic Polynomial):** Let $p_A(x) = \det(xI - A)$ for $A \in M_n(F)$. Then $p_A(x) \in F[x]$ monic with $\deg(p_A) = n$.

**Proof:** Let $A = ((a_{ij}))$. Then

$$xI - A = \begin{bmatrix} x - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & x - a_{22} & \ddots & \vdots \\ \vdots & & & \\ -a_{n1} & \cdots & & x - a_{nn} \end{bmatrix} =: B. \quad \text{Write } b_{ij} = x\delta_{ij} - a_{ij} \text{ where } \delta_{ij} = \begin{cases} 1 & \text{when } i = j; \\ 0 & \text{otherwise.} \end{cases}$$

Then $\det(xI - A) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \left( \prod_{i=1}^{n} (x\delta_{i\sigma(i)} - a_{i\sigma(i)}) \right)$ and we see that this is a polynomial.

$$= \sum_{\sigma \in S_n} \Big( \mathrm{sgn}(\sigma)\, x^n \underbrace{\left( \prod_{i=1}^{n} \delta_{i\sigma(i)} \right)}_{\text{this product } = 1 \text{ because the highest degree } x \text{ comes from the diagonal terms multiplied}} + (\text{lower degree terms}) \Big) \quad \square$$

$$\sum_{\sigma \in S_n} \left( \prod_{i=1}^{n} \delta_{i\sigma(i)} \right) \mathrm{sgn}(\sigma) = \det \begin{pmatrix} \delta_{11} & \delta_{12} & \cdots & \delta_{1n} \\ \vdots & & & \vdots \\ \delta_{n1} & \cdots & \cdots & \delta_{nn} \end{pmatrix} = \det(I) = 1.$$

**Theorem (Similar Characteristic Polynomials):**

Let $A, B \in M_n(F)$ be similar (so $\exists S \in GL_n(F)$ such that $S^{-1}AS = B$). Then $p_A(x) = p_B(x)$.

In particular, $A$ and $B$ have the same eigenvalues.

**Proof:** We have $B = S^{-1}AS$ for some $S \in GL_n(F)$ so $p_B(x) = \det(xI - B) = \det(xI - S^{-1}AS)$.

Notice $xI - S^{-1}AS = S^{-1}(xI - A)S = (S^{-1}xI - S^{-1}A)S = \underline{S^{-1}xI}S - S^{-1}AS$

$= S^{-1}xS = xS^{-1}S = xI$ ← why?

so we have $\det(xI - S^{-1}AS) = \det(S^{-1}(xI - A)S)$

$= \det(S^{-1})\det(xI - A)\det(S) = \det(S^{-1})\det(S)\det(xI - A)$ (commutativity in $F[x]$)

$= \det(S^{-1}S)\, p_A(x) = p_A(x)$. □

**Remark:** $p_A(0) = \det(-A) = \det(-I)\det(A) = (-1)^n \det(A)$ where $A \in M_n(F)$.

**Lemma (Triangular Characteristic Polynomials):** Let $\mathcal{U}_n$ denote the set of $n \times n$ upper triangular matrices.

If $A = ((a_{ij})) = \begin{pmatrix} a_{11} & & * \\ & a_{22} & \\ 0 & & \ddots \; a_{nn} \end{pmatrix} \in \mathcal{U}_n$, then $p_A(x) = \prod_{i=1}^{n}(x - a_{ii})$ so $A$ has eigenvalues $a_{11}, a_{22}, \ldots, a_{nn} \in F$.

**Proof:** Recall $B \in \mathcal{U}_n \Rightarrow \det(B) = \prod_{i=1}^{n} b_{ii}$ where $B = ((b_{ij}))$. Notice that $xI - A \in \mathcal{U}_n$ for each $x \in F$.

Thus $\det(xI - A) = \prod_{i=1}^{n}(x - a_{ii})$. □

If $c$ is an eigenvalue of $A$, then we can write $p_A(x) = (x-c)^m q(x)$ where $q(c) \neq 0$.

($m$:) multiplicity of the eigenvalue $c$, and $\dim W_c \leq m$.

**Lemma:** Let $p(x) \in F[x]$ be monic with $\deg(p) = n$. Then $\exists C \in M_n(F)$ such that $p(x) = p_C(x) = \det(xI - C)$.

**Proof:** Let $p(x) = x^n + c_{n-1}x^{n-1} + \ldots + c_1 x + c_0 \in F[x]$.

Choose $C = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & 0 & \cdots & 0 & -c_2 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -c_{n-1} \end{pmatrix}$ the companion matrix of $p(x)$.

We can prove by induction on $n$ that $\det(xI - C) = p(x)$. Use cofactor expansion for the inductive step. Exercise.

**Lemma:** Let $A \in \mathcal{U}_n(F)$ be upper-triangular with diagonal entries $\gamma_1, \ldots, \gamma_n$. Then $p_A(x) = \prod_i (x - \gamma_i)$.

**Proof:** Follows from $xI - U \in \mathcal{U}_n(F) \Rightarrow \det(xI - U) =$ product of diagonal entries. □

**Q:** Let $\Phi_m(x) \in \mathbb{Z}[x] \subseteq \mathbb{Q}[x]$ denote the $m^{th}$ cyclotomic polynomial. We have shown this is monic.

What is the companion matrix for $\Phi_m(x)$ in $M_m(\mathbb{Q})$?

Let $W_1, W_2$ be subspaces of $V$.

We can define $W_1 + W_2 = \{w + \bar{w} \mid w \in W_1,\ \bar{w} \in W_2\} \subseteq V$ the smallest subspace containing $W_1$ and $W_2$.

and $W_1 \cap W_2 = \{w \mid w \in W_1 \cap W_2\} \supseteq W_1, W_2$ the largest subspace contained in both $W_1$ and $W_2$.

We say $W_1$ and $W_2$ are complementary if $W_1 + W_2 = V$ and $W_1 \cap W_2 = \{0\}$.

We define quotient spaces as we always do in algebra.

Let $A \in M_n(F)$. We say $A$ is diagonalizable (over $F$) if there exists $S \in GL_n(F)$ and $D$ diagonal $(n \times n)$ such that $D = S^{-1}AS$ (i.e. $A$ is similar to a diagonal matrix).

"Similarity" defines a ring isomorphism $\varphi : M_n(F) \longrightarrow M_n(F)$.
$X \longmapsto S^{-1}XS$

Notice $\varphi(A)\varphi(B) = (S^{-1}AS)(S^{-1}BS) = S^{-1}(AB)S = \varphi(AB)$;

$\varphi(A) + \varphi(B) = S^{-1}AS + S^{-1}BS = S^{-1}(A+B)S = \varphi(A+B)$;

$\varphi^{-1}(A) := SAS^{-1}$ so $\varphi^{-1}(\varphi(A)) = S(S^{-1}AS)S^{-1} = A$.

<u>Lemma:</u> Let $A \in M_n(F)$ such that $F^n$ has a basis $\vec{v}_1, \dots, \vec{v}_n \in F^n$ consisting of eigenvectors of $A$.
Then $A$ is diagonalizable, and if $c_1, \dots, c_n$ are eigenvalues of $A$, then $\underbrace{\begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix}}_{\text{diagonal}} = S^{-1}AS$ where $S = \begin{pmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{pmatrix}$.

invertible since $\{\vec{v}_1, \dots, \vec{v}_n\}$ basis $\Rightarrow \operatorname{rank}(S) = n$.

<u>Proof:</u> By computation. Let $S = \begin{pmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{pmatrix} \in GL_n(F)$, so $\det(S) \neq 0$. Note $S^{-1}S = \begin{pmatrix} | & & | \\ S^{-1}\vec{v}_1 & \cdots & S^{-1}\vec{v}_n \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ \vec{e}_1 & \cdots & \vec{e}_n \\ | & & | \end{pmatrix}$.

Then $AS = A\begin{pmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ A\vec{v}_1 & \cdots & A\vec{v}_n \\ | & & | \end{pmatrix}$ since $A\vec{v}_i = c_i\vec{v}_i$ for each of these eigenvectors.

- how are we doing this step?

$$= \begin{pmatrix} | & & | \\ c_1\vec{v}_1 & \cdots & c_n\vec{v}_n \\ | & & | \end{pmatrix} \quad \text{so} \quad S^{-1}AS = S^{-1}\begin{pmatrix} | & & | \\ c_1\vec{v}_1 & \cdots & c_n\vec{v}_n \\ | & & | \end{pmatrix}$$

$$= \begin{pmatrix} | & & | \\ S^{-1}c_1\vec{v}_1 & \cdots & S^{-1}c_n\vec{v}_n \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ c_1S^{-1}\vec{v}_1 & \cdots & c_nS^{-1}\vec{v}_n \\ | & & | \end{pmatrix}$$

$$= \begin{pmatrix} | & & | \\ c_1\vec{e}_1 & \cdots & c_n\vec{e}_n \\ | & & | \end{pmatrix}.$$ We are done. $\square$

<u>Theorem (Diagonalizability and Eigenvectors):</u> Let $A \in M_n(F)$. Then $A$ is diagonalizable over $F$
$\Leftrightarrow$ there is a basis for $F^n$ consisting of eigenvectors of $A$.

<u>Proof:</u> $(\Rightarrow)$ Suppose $A$ is diagonalizable, so $D = S^{-1}AS$ for some $S \in GL_n(F)$. Since $S$ is invertible, its columns form a basis for $F^n$.

Let $S = \begin{pmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{pmatrix}$. Then $D = \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix}$ and notice that $D\vec{e}_j = c_j\vec{e}_j$.

Thus $D\vec{e}_j = S^{-1}AS\,\vec{e}_j = c_j\vec{e}_j$

$\Rightarrow S(S^{-1}AS\vec{e}_j) = AS\vec{e}_j = Sc_j\vec{e}_j = c_j(S\vec{e}_j)$. ($j^{\text{th}}$ column of $S$)

$\Rightarrow A\vec{v}_j = c_j\vec{v}_j$ so each $\vec{v}_j$ is an eigenvector of $A$.

$\Rightarrow F^n$ has a basis of eigenvectors of $A$. $\triangle$

$(\Leftarrow)$ See above Lemma. $\square$

Theorem (Roots and Diagonalizability): Let $A \in M_n(F)$. If the characteristic polynomial $p_A(x) \in F[x]$ has $n$ distinct roots $c_1, \dots, c_n \in F$, then $A$ is diagonalizable over $F$.

Proof: Let $\vec{v}_1, \dots, \vec{v}_n$ be the corresponding eigenvectors to these roots.

Claim: $\vec{v}_1, \dots, \vec{v}_n$ forms a basis for $F^n$

It suffices to show that the vectors are linearly independent, since there are $n$ of them.

Suppose for a contradiction that $\{\vec{v}_1, \dots, \vec{v}_n\}$ is linearly dependent.

Then there is a minimal linearly dependent subset $\{\vec{v}_{i_1}, \dots, \vec{v}_{i_k}\}$ with $k \geq 2$ since all these vectors are nonzero.

$\Rightarrow$ there exists a nontrivial relation $\sum_{j=1}^{k} \lambda_j \vec{v}_{i_j} = 0$

$$\Rightarrow A \sum_{j=1}^{k} \lambda_j \vec{v}_{i_j} = \sum_{j=1}^{k} \lambda_j A \vec{v}_{i_j} = \sum_{j=1}^{k} \lambda_j c_{i_j} \vec{v}_{i_j} = 0 \qquad \text{(left-multiplying by } A\text{);}$$

multiplying first relation by scalar $c_{i_k}$

so we have
$$c_{i_k} \sum_{j=1}^{k} \lambda_j \vec{v}_{i_j} - \sum_{j=1}^{k} \lambda_j c_{i_j} \vec{v}_{i_j} = 0$$
$$\parallel$$
$$\sum_{j=1}^{k} \lambda_j (c_{i_k} - c_{i_j}) \vec{v}_{i_j} = 0.$$

By minimality of our dependent subset, we must have $\lambda_j (c_{i_k} - c_{i_j}) = 0$ for $j \in \{1, 2, \dots, k-1\}$.

Since the $c_i$ are distinct, we have $c_{i_k} - c_{i_j} \neq 0$ for $j \in \{1, 2, \dots, k-1\}$ so $\lambda_{i_j} = 0$ for these $j$.

$$\Rightarrow \sum_{j=1}^{k} \lambda_j \vec{v}_{i_j} = 0 = 0 + \lambda_k \vec{v}_{i_k} \quad \text{so } \lambda_k = 0 \text{ as well.}$$

Therefore our linear relation was trivial. Contradiction. ↯

Corollary: Let $F$ be an algebraically closed field. Let $A \in M_n(F)$. If $\gcd(p_A(x), p_A'(x)) = 1$, then $A$ is diagonalizable.

Proof: In algebraically closed fields, we have distinct roots when $\gcd(p_A, p_A') = 1$. Namely, let $p_A(x) = (x-c_1)^{m} \dots (x-c_n)^{m}$.

We have $(x-c_k)^{m-1} \mid p_A'(x) \Leftrightarrow c_k$ is a zero of multiplicity $m$ for $p_A(x)$. $\square$

We say $A \in M_n(F)$ is triangularizable if $A$ is similar to a triangular matrix (so $A = S^{-1}TS$ for some $S \in GL_n(F)$, and $T$ upper-triangular).

Theorem (Triangularization): Let $F$ be algebraically closed and let $A \in M_n(F)$. Then $A$ is triangularizable.

Proof: By induction on $n$. When $n=1$, this is obviously true (every $M_1(F) = \mathcal{U}_1(F)$).

Suppose any $A \in M_{n-1}(F)$ is triangularizable for some $n \geq 2$. Let $A \in M_n(F)$.

$\Rightarrow p_A(x) \in F[x]$ is monic, $\deg(p_A) = n$, and splits completely over $F$ by algebraic closure, so let $c \in F$ be a root of $p_A$.

We have $c$ an eigenvalue of $A$, so let $v_1 \in F^n$ be the corresponding eigenvector. Extend $v_1$ to an ordered basis $(v_1, \dots, v_n) \subseteq F^n$.

Let $S = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}$. Since the columns of $S$ form a basis for $F^n$, we have $S \in GL_n(F)$.

We check $S^{-1}AS$ is of the form

$$S^{-1}AS = \left(\begin{array}{c|c} 1 & \boxed{w} \\ \hline \boxed{0} & B \end{array}\right) \quad \text{for some } B \in M_{n-1}(F) \text{ and row-vector } w \in F^{n-1}.$$

$T^{-1}BT = U \in \mathcal{U}_{n-1}$

induction hypothesis

$$\Rightarrow \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & T^{-1} \end{pmatrix}}\begin{pmatrix} 1 & w \\ 0 & B \end{pmatrix}\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & T \end{pmatrix}}_{\in GL_n(F)} = \begin{pmatrix} 1 & w' \\ 0 & T^{-1}BT \end{pmatrix} = \begin{pmatrix} 1 & w' \\ 0 & U \end{pmatrix} \in \mathcal{U}_n(F).$$

$$\left(\begin{array}{c|c} 1 & \boxed{0} \\ \hline \boxed{0} & T^{-1} \end{array}\right) \in GL_n(F)$$

$T^{-1} \in GL_{n-1}(F)$

Therefore $A$ is triangularizable. $\square$

**Corollary:** Let $A \in M_n(F)$ (algebraically closed) with $p_A(x) = (x-\lambda_1)\cdots(x-\lambda_n)$. Then $Tr(A) = \sum \lambda_i$ and $\det(A) = \prod \lambda_i$.

**Proof:** Since $F$ algebraically closed, we have $A$ triangularizable by $U \in \mathcal{U}_n(F)$. Let $\gamma_1, \dots, \gamma_n$ denote the diagonal elements of $U$.

We have $xI - U \in \mathcal{U}_n(F) \Rightarrow \det(xI-U) = \prod_i (x-\gamma_i)$. Since $A \sim U$, we have $p_A(x) = p_U(x)$

$$\Rightarrow \prod_i (x-\lambda_i) = \prod_i (x-\gamma_i)$$ so the $\gamma_i$ are a rearrangement of the $\lambda_i$.

Also $A \sim U \Rightarrow Tr(A) = Tr(U) = \sum_i \gamma_i = \sum_i \lambda_i$. $\square$

Notice if $A \in M_n(F)$ diagonalizable, then $A = SDS^{-1} \Rightarrow A^k = SD^kS^{-1}$ $(= (SD\cancel{S^{-1}})(\cancel{S}D\cancel{S^{-1}})\cdots(\cancel{S}DS^{-1}))$.

An $F$-valued (field) sequence $f_0, f_1, \dots = \{f(n)\}_{n \geq 0}$ satisfies a *linear recurrence* if there exists some $d \geq 1$ and $c_1, \dots, c_d \in F$ (constants) such that

$f(n) = c_1 f(n-1) + \dots + c_d f(n-d)$ for all $n \geq d$. We say $f(0), \dots, f(d-1)$ are the *initial values* of the sequence $f(n)$.

The Fibonacci numbers $f(0) = 0$, $f(1) = 1$, and $f(n) = f(n-1) + f(n-2)$ for $n \geq 2$ satisfies a linear recurrence.

## Solving (most) Linear Recurrences

0 Let $v = \begin{bmatrix} f(0) \\ f(1) \\ \vdots \\ f(d-1) \end{bmatrix} \in F^d$ be an *initial value vector*. Let $w = [1\ 0 \cdots 0]$ and let $A = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ c_d & c_{d-1} & \cdots & c_3 & c_2 & c_1 \end{pmatrix}$.

1 **Proposition:** For all $n \geq 0$ we have $A^n v = \begin{pmatrix} f(n) \\ f(n+1) \\ \vdots \\ f(n+d-1) \end{pmatrix}$.

**Proof:** Induction on $n$. When $n=0$, then $A^0 v = Iv = v = \begin{bmatrix} f(0) \\ f(1) \\ \vdots \\ f(d-1) \end{bmatrix}$.

Suppose the claim holds true for some $k \geq 0$, and consider the case when $n = k+1$.

Then $A^{k+1}v = A(A^k v) = A\begin{pmatrix} f(k) \\ \vdots \\ f(k+d-1) \end{pmatrix}$

$$= \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ c_d & c_{d-1} & \cdots & c_3 & c_2 & c_1 \end{pmatrix}\begin{pmatrix} f(k) \\ f(k+1) \\ \vdots \\ f(k+d-1) \end{pmatrix} = \begin{pmatrix} f(k+1) \\ f(k+2) \\ \vdots \\ f(k+d-1) \\ c_d f(k) + \dots + c_1 f(k+d-1) \end{pmatrix}$$

$$= \begin{pmatrix} f(k+1) \\ \vdots \\ f(k+d) \end{pmatrix}.$$ Done by induction. $\square$

2 If $A$ is diagonalizable, write $D = S^{-1}AS$ and compute $w_0 := wS$ and $v_0 := S^{-1}v$.

Notice $A^n = SD^nS^{-1}$ so $wA^n v = [1\ 0 \cdots 0]\begin{bmatrix} f(n) \\ \vdots \\ f(n+d-1) \end{bmatrix} = f(n)$ (one-hot encoding!)

$\Rightarrow f(n) = wSD^nS^{-1}v = (wS)D^n(S^{-1}v) = w_0 D^n v_0$.

3 Write $D = \operatorname{diag}(d_1, \dots, d_n)$ and $w_0 = (a_1\ a_2 \dots a_d)$ and $v_0 = \begin{pmatrix} b_1 \\ \vdots \\ b_d \end{pmatrix}$. Then $f(n) = w_0 D^n v_0 = \sum_{i=1}^{d} a_i b_i d_i^n$ for $n \geq 0$.

Fibonacci closed form. Notice the linear recurrence is defined by $f(n) = f(n-1) + f(n-2)$ for $n \geq 2$.

We have $w = [1\ 0]$ and initial value vector $v = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. We also have $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$.

Notice $p_A(x) = x^2 - x - 1$ which has distinct roots $\frac{\sqrt{5}-1}{2}, \frac{\sqrt{5}+1}{2} \in \mathbb{R}$ so we have $A$ diagonalizable.

$\Rightarrow D = S^{-1}AS$ where $S$ has columns that are eigenvectors corresponding to those eigenvalues. By computation we find.

$$S = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix} \xrightarrow{\text{adjoint}} S^{-1} = \sqrt{5}^{-1} \begin{pmatrix} 1 & -\frac{1-\sqrt{5}}{2} \\ -1 & \frac{1+\sqrt{5}}{2} \end{pmatrix} \xrightarrow{\text{computation}} D = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}.$$

By computing $w_0, v_0$, we find that $f(n) = \sqrt{5}^{-1}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right)$.

<u>Proposition</u>: $f(n) = \sqrt{5}^{-1}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right) \in \mathbb{Z}$ is an integer for $n \geq 2$.

<u>Proof</u>: Involved Galois theory proof. $\square$

<u>Theorem (Triangularisable in Algebraic Closure)</u>: Let $A \in M_n(F)$. There exists a field extension $K \supseteq F$ such that $A \in M_n(K)$ is triangularisable. Moreover, if $A$ has distinct eigenvalues in $K$, then $A$ is diagonalisable.

<u>Proof</u>: Choose either $K = F[\underbrace{c_1, \dots, c_n}_{\text{eigenvalues}}]$ the splitting field of $p_A(x)$, or choose $K$ to be the algebraic closure of $F$. $\square$

# A09 Warm-Up Q12

Let $A \in M_n(\mathbb{C})$ and let $p_A(t) = t^n + c_{n-1}t^{n-1} + \ldots + c_0$ where $c_{n-1} = -\mathrm{Tr}(A)$ and $c_0 = (-1)^n \det(A)$.

Show that the sum of the eigenvalues of $A$ (with multiplicity) $= \mathrm{Tr}(A)$ and product of eigenvalues $= \det(A)$.

Proof: Product of eigenvalues obvious since $p_A(t) = (t-\lambda_1)^{m_1} \cdots (t-\lambda_r)^{m_r} \in \mathbb{C}[t]$ splits completely. Thus $p_A(0) = (-1)^n \lambda_1 \cdots \lambda_r$.

We show now that the sum of our eigenvalues is $\mathrm{Tr}(A)$ by induction. We show $n=1,2,3$ as an exercise.

Base case: $n=1$: $A = [a] \in M_1(\mathbb{C})$ so $p_A(t) = t-a \Rightarrow a$ eigenvalue. But $a = \mathrm{Tr}(A)$ too. $\triangle$

$n=2$: $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in M_2(\mathbb{C})$ so $\det(t\mathbb{1}-A) = \det\left(\begin{pmatrix} t-a_{11} & -a_{12} \\ -a_{21} & t-a_{22} \end{pmatrix}\right)$

$$= (t-a_{11})(t-a_{22}) - a_{12}a_{21}$$

second coefficient of polynomial is sum of roots by Vieta's relations $\Rightarrow$ sum of eigenvalues

$$= t^2 - (a_{11}+a_{22})t - a_{12}a_{21} = t^2 - \mathrm{Tr}(A)t - a_{12}a_{21}. \quad \triangle$$

$n=3$: $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \in M_3(\mathbb{C}) \Rightarrow \det(t\mathbb{1}-A) = \det\begin{pmatrix} t-a_{11} & -a_{12} & -a_{13} \\ -a_{21} & t-a_{22} & -a_{23} \\ -a_{31} & -a_{32} & t-a_{33} \end{pmatrix}$

cofactor expansion along first row

$$= (t-a_{11})\begin{vmatrix} t-a_{22} & -a_{23} \\ -a_{32} & t-a_{33} \end{vmatrix} \underbrace{- (a_{12})\begin{vmatrix} -a_{21} & -a_{23} \\ -a_{31} & t-a_{33} \end{vmatrix} + (-a_{13})\begin{vmatrix} -a_{21} & t-a_{22} \\ -a_{31} & -a_{32} \end{vmatrix}}_{\text{terms of degree } \leq n-2 \text{ ignored}}$$

$$\Rightarrow (t-a_{11})\begin{vmatrix} t-a_{22} & -a_{23} \\ -a_{32} & t-a_{33} \end{vmatrix} = (t-a_{11}) \det\left(t\mathbb{1} - \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}\right)$$

$$= (t-a_{11})\left(t^2 - \mathrm{Tr}\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} t + c_0\right) \quad \text{by } n=2 \text{ case;}$$

$$= t^3 - (a_{11}+a_{22}+a_{33})t^2 + \ldots$$

$$= t^3 - \underbrace{\mathrm{Tr}(A)}_{\text{sum of eigenvalues by Vieta's}} t^2 + \ldots \quad \triangle$$

Consider the statement true for all $n \leq k-1$ where $k \geq 2$. Consider $n=k$. We have:

$$\underset{((a_{ij}))_{i,j\leq k}}{A_k} \in M_k(\mathbb{C}) \Rightarrow \det(t\mathbb{1}-A) = \det\begin{pmatrix} t-a_{11} & -a_{12} & \cdots & -a_{1k} \\ -a_{21} & t-a_{22} & \cdots & -a_{2k} \\ \vdots & & & \\ -a_{k1} & \cdot & \cdot & t-a_{kk} \end{pmatrix}$$

$$= (t-a_{11})\det\begin{pmatrix} t-a_{22} & \cdots & -a_{2k} \\ \vdots & \ddots & \\ -a_{k2} & & t-a_{kk} \end{pmatrix} + \overbrace{\sum_{i=2}^{k} (-1)^{i-1}(-a_{1i})A(1|i)}^{\text{terms of degree } \leq k-2 \Rightarrow \text{ignore}}.$$

inductive hypothesis

$$= (t-a_{11})(t^{k-1} - \mathrm{Tr}(A(1|1))t^{k-2} + \ldots + c_0) + \mathcal{O}(t^{k-2})$$

$$= t^k - \underbrace{\mathrm{Tr}(A)}_{\text{sum of eigenvalues by Vieta's}} t^{k-1} + \mathcal{O}(t^{k-2}).$$ Thus true for $n=k \Rightarrow$ done by induction. $\square$

We define a Vandermonde matrix $V_n(\lambda_1, \ldots, \lambda_n)$ to be of the form $\begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-1} \end{pmatrix}$ so $\det(V_n(\lambda_1, \ldots, \lambda_n)) = \prod_{\substack{j \neq i \\ j > i}} (\lambda_j - \lambda_i)$.

Exercise: Let $A \in M_n(\mathbb{C})$ such that $\mathrm{Tr}(A^i) = 0$ for each $i \leq n$. Then $A^n = 0$.

Proof: $A \in M_n(\mathbb{C}) \Rightarrow A$ triangularisable $\Rightarrow U = S^{-1}AS$ for $S \in GL_n(\mathbb{C})$ and $U \in \mathcal{U}_n(\mathbb{C})$. Note that $\mathrm{Tr}(A) \overset{\text{similar matrices have same trace}}{=} \mathrm{Tr}(U)$.

Thus $A^i \sim U^i \Rightarrow \mathrm{Tr}(A^i) = \mathrm{Tr}(U^i) = \sum$ (diagonal entries of $U$).

However, the diagonal entries of $\mathcal{U}$ correspond to eigenvalues of $A$.

all nonzero

Let $\lambda_1, \dots, \lambda_r$ denote distinct nonzero eigenvalues of $A$ with multiplicities $m_1, \dots, m_r \in \mathbb{N}$. Notice $r \leq n$.

$$\Rightarrow \operatorname{Tr}(A^i) = \operatorname{Tr}(\mathcal{U}^i) = \sum_{k=1}^{r} \lambda_k^i m_k = \lambda_1^i m_1 + \dots + \lambda_r^i m_r.$$

We use Vandermonde matrices to show that we cannot have such eigenvalues, thus arriving at a contradiction.

Define $V = \begin{pmatrix} \lambda_1^1 & \lambda_2^1 & \cdots & \lambda_r^1 \\ \vdots & \vdots & & \vdots \\ \lambda_1^r & \lambda_2^r & \cdots & \lambda_r^r \end{pmatrix}$ and notice that $\begin{pmatrix} \lambda_1^1 & \lambda_2^1 & \cdots & \lambda_r^1 \\ \vdots & \vdots & & \vdots \\ \lambda_1^r & \lambda_2^r & \cdots & \lambda_r^r \end{pmatrix} \begin{bmatrix} m_1 \\ \vdots \\ m_r \end{bmatrix} = \begin{pmatrix} \overbrace{m_1\lambda_1 + \dots + m_r\lambda_r}^{\operatorname{Tr}(\mathcal{U}^1)} \\ \vdots \\ \underbrace{m_1\lambda_1^r + \dots + m_r\lambda_r^r}_{\operatorname{Tr}(\mathcal{U}^r)} \end{pmatrix} = 0$ by $\operatorname{Tr}(\mathcal{U}^i) = 0$.
(and $r \leq n$)

But $\begin{bmatrix} m_1 \\ \vdots \\ m_r \end{bmatrix} \neq 0 \Rightarrow \det(V) = 0$ since $\ker(V) \neq \{0\}$. Recall since determinant cofactor expansion can be along either rows or columns, we have $\det(V) = \det(V^t)$.

Since determinant is $n$-linear, we have $\det\begin{pmatrix} -\alpha r_i - \\ * \end{pmatrix} = \alpha \det\begin{pmatrix} -r_i- \\ * \end{pmatrix}$. Scaling a row by $\alpha \in F$ has this effect.

Notice $V^t = \begin{pmatrix} \lambda_1^1 & \lambda_1^2 & \cdots & \lambda_1^r \\ \lambda_2^1 & \lambda_2^2 & \cdots & \lambda_2^r \\ \vdots & \vdots & & \vdots \\ \lambda_r & \lambda_r^2 & \cdots & \lambda_r^r \end{pmatrix} = \begin{pmatrix} \lambda_1 & \\ \vdots & \boxed{V_r(\lambda_1, \dots, \lambda_r)} \\ \lambda_r & \end{pmatrix} \Rightarrow \det(V) = \det(V^t) = \lambda_1 \cdots \lambda_r \det(V_r(\lambda_1, \dots, \lambda_r))$

rescaling each row

$$= \lambda_1 \cdots \lambda_r \prod_{\substack{i \neq j \\ i < j}} (\lambda_i - \lambda_j) = 0.$$

integral domain

But $\mathbb{C} = F$ and each $\lambda_i$ distinct and nonzero. Contradiction. ↯ They must all be zero.

$\Rightarrow A = S\mathcal{U}S^{-1}$ where $\mathcal{U}$ strictly upper triangular, so $\mathcal{U}^n$ vanishes.

$\Rightarrow A^n = S\mathcal{U}^n S^{-1} = S \cdot 0 \cdot S^{-1} = 0$. Done. $\square$

<u>Remark</u>: The above does not hold for $\operatorname{char}(F) \neq 0$. Why not?

any field $F$ with $\operatorname{char}(F)=0$

<u>Exercise (Herstein)</u>: Suppose $A, B \in M_n(\mathbb{C})$ and $A$ commutes with $AB-BA$ (so $A(AB-BA) = (AB-BA)A$).

Show $AB-BA$ is nilpotent.

<u>Proof</u>: Denote $AB-BA = C \in M_n(F)$. Claim: $\operatorname{Tr}(C^i) = 0$ for $i \geq 1$. Then we use our previous exercise's result on $C$.

For $i=1$, we see that $\operatorname{Tr}(AB) = \operatorname{Tr}(BA) \Rightarrow \operatorname{Tr}(AB-BA) = \operatorname{Tr}(AB) - \operatorname{Tr}(BA) = 0$.

Suppose $\operatorname{Tr}(C^{k-1}) = 0$ for some $k \geq 2$. Consider $\operatorname{Tr}(C^k)$. Notice:

$$A(AB-BA)^{k-1} = (AB-BA)^{k-1}A \Rightarrow (AB-BA)^k = (AB-BA)^{k-1}(AB-BA)$$

$$= (AB-BA)^{k-1}AB - (AB-BA)^{k-1}BA$$

commute

$$= \underbrace{AC^{k-1}}_{N}B - \underbrace{C^{k-1}B}_{N}A = AN - NA \rightsquigarrow \operatorname{Tr}(AN-NA) = 0$$

$$\Rightarrow \operatorname{Tr}(C^k) = 0.$$

Thus the result follows from the previous exercise. $\square$

Exercise: The set $\{e^{\lambda x} : \lambda \in \mathbb{R}\}$ is linearly independent in $C^1(\mathbb{R})$.

Proof: Contradiction. Suppose for a contradiction $S = \{e^{\lambda x} : \lambda \in \mathbb{R}\}$ is linearly dependent. Thus $\exists \overset{\text{not all zero}}{c_1, \dots, c_n} : c_1 e^{\lambda_1 x} + \dots + c_n e^{\lambda_n x} = 0$.

Notice that all the $\lambda_i$ are distinct, and $0 \in C^1(\mathbb{R})$ refers to the zero function. We apply repeated differentiation to both sides:

$$c_1 e^{\lambda_1 x} + \dots + c_n e^{\lambda_n x} = 0$$

$$\Rightarrow \begin{aligned} c_1 \lambda_1 e^{\lambda_1 x} + \dots + c_n \lambda_n e^{\lambda_n x} &= 0 \\ c_1 \lambda_1^2 e^{\lambda_1 x} + \dots + c_n \lambda_n^2 e^{\lambda_n x} &= 0 \\ \vdots \qquad\qquad & \ \ \vdots \\ c_1 \lambda_1^{n-1} e^{\lambda_1 x} + \dots + c_n \lambda_n^{n-1} e^{\lambda_n x} &= 0. \end{aligned}$$

Set up a matrix $V = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & & & \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix}$ and notice $V \cdot \overset{\text{nonzero matrix}}{\begin{bmatrix} c_1 e^{\lambda_1 x} \\ \vdots \\ c_n e^{\lambda_n x} \end{bmatrix}} = 0$.

$$\Rightarrow \det(V) = 0 \text{ since } \ker(V) \neq \{0\}.$$

We have $\det(V) = \det(V^t)$ and $V^t = V_n(\lambda_1, \dots, \lambda_n) \Rightarrow \det(V^t) = \prod_{\substack{i \neq j \\ i < j}} (\lambda_i - \lambda_j) = 0$. But each $\lambda_i \neq \lambda_j$ by distinction.

Contradiction. $\lightning$

Thus our set $S$ must be linearly independent. $\square$