MATH 115A - SPRING 2022

Pablo S. Ocal

1. Fields and vector spaces

For the non-mathematician, linear algebra is the study of linear equations and linear transformations. For us, linear algebra will be the study of linear maps between vector spaces. We should think of vector spaces as abstract objects with special structure that behaves nicely with respect to scalars, and linear maps are functions that preserve this special structure.

Definition: A field $\mathbb{F}$ is a set with two operations

$$+ : \mathbb{F} \times \mathbb{F} \longrightarrow \mathbb{F} \qquad \text{and} \qquad \cdot : \mathbb{F} \times \mathbb{F} \longrightarrow \mathbb{F}$$
$$(a,b) \longmapsto a+b \qquad\qquad (a,b) \longmapsto a \cdot b$$

called sum and product respectively, such that for all $a, b, c \in \mathbb{F}$ we have:

(1) Commutativity: $a+b = b+a$ and $a \cdot b = b \cdot a$.

(2) Associativity: $(a+b)+c = a+(b+c)$ and $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

(3) Identity: there exist $0, 1 \in \mathbb{F}$ with $a+0 = a$ and $a \cdot 1 = a$.

(4) Inverses: when $a \neq 0$ there exist $-a, a^{-1} \in \mathbb{F}$ with $a+(-a) = 0$ and $a a^{-1} = 1$.

(5) Distributivity: $a \cdot (b+c) = a \cdot b + a \cdot c$.

The elements of a field are called scalars.

Example:

1. Some number sets are fields: $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.

2. Some number sets are not fields: $\mathbb{N}$, $\mathbb{Z}$. (why?)

3. There are weird fields:

$$\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$$

$\mathbb{Z}_2 = \{[0], [1]\}$ is the field of integers mod 2, having:

$$0+1=1 \quad , \quad 0+0=0 \quad , \quad 1+1=0 \quad , \quad 0 \cdot 1 = 0 \quad , \quad 1 \cdot 1 = 1.$$

We can think of $\mathbb{Z}_2$ as $\mathbb{Z}$ where we have declared that all even numbers are the same, and also that all odd numbers are the same:

$$2k \equiv 0 \quad \text{and} \quad 2k+1 \equiv 0 \quad \text{for all } k \in \mathbb{Z}.$$

$\mathbb{Z}_p = \{[0], [1], \cdots, [p-1]\}$ for $p \in \mathbb{N}$ prime is the field of integers mod p.

We can think of $\mathbb{Z}_p$ as $\mathbb{Z}$ where we declare that two numbers are equal if and only if they have the same remainder when divided by p:

$$p \cdot k + j \equiv j \text{ for all } 0 \leq j < p \text{ and all } k \in \mathbb{Z}, \text{ so}$$

$$[j] = \{\text{integers with remainder } j \text{ upon division by } p\}.$$

Definition: A vector space V over a field $\mathbb{F}$ is a set with two operations:

$+ : V \times V \longrightarrow V$ and $\cdot : \mathbb{F} \times V \longrightarrow V$

$(x,y) \longmapsto x+y$ $\quad$ $(a,x) \longmapsto a \cdot x$

called addition and scalar multiplication respectively, such that for all $x,y,z \in V$ and $a,b \in \mathbb{F}$:

(1) Commutativity of addition: $x+y = y+x$.

(2) Associativity of addition: $(x+y)+z = x+(y+z)$.

(3) Identity in V: there exists $\vec{0} \in V$ with $x+\vec{0} = \vec{0}$.

(4) Inverses in V: there exists $-x \in V$ with $x+(-x) = \vec{0}$.

(5) Scalar identity: $1 \cdot x = x$. (do we have multiplicative inverses x^{-1} in V?)

(6) Associativity of scalar multiplication: $a \cdot (b \cdot x) = (a \cdot b) \cdot x$.

(7) Distributivity of scalar multiplication over addition: $a \cdot (x+y) = a \cdot x + a \cdot y$.

(8) Distributivity of the sum over scalar multiplication: $(a+b) \cdot x = a \cdot x + b \cdot x$.

These properties are saying that the addition and multiplication by scalars in V behave well with respect to the sum and product in $\mathbb{F}$.

Remark: Alternatively, we could say that a vector space is a commutative group under addition with associative and distributive scalar multiplication.

In particular, vector spaces are closed under finite sums and scalar multiplication: if $x_1, \dots, x_n \in V$ and $a_1, \dots, a_n \in \mathbb{F}$, then $a_1 x_1 + \cdots + a_n x_n \in V$.

When $\mathbb{F} = \mathbb{R}$, we say that V is a real vector space. When $\mathbb{F} = \mathbb{C}$ we say that V is a complex vector space.

Examples:

1. $\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n}$ is called the real n-space.

   The elements in $\mathbb{R}^n$ are n-tuples $(r_1, \ldots, r_n)$ with $r_1, \ldots, r_n \in \mathbb{R}$.

   The vector addition is done componentwise:

   $$(r_1, \ldots, r_n) + (s_1, \ldots, s_n) = (r_1 + s_1, \ldots, r_n + s_n)$$

   The scalar multiplication is done componentwise:

   $$a \cdot (r_1, \ldots, r_n) = (a \cdot r_1, \ldots, a \cdot r_n)$$

   Remark: Here we could replace the field $\mathbb{R}$ by $\mathbb{Q}$, and everything would still make sense. It is important to specify over which field we are working.

   In fact, if we replace $\mathbb{R}$ by $\mathbb{Z}$, things still make sense. When we work over a ring instead of a field, we generalize vector spaces to the notion of modules.

2. Let $\mathbb{F}$ be a field, let S be a set, let V be the set of functions from S to $\mathbb{F}$.

   Namely elements $f \in V$ are functions of sets $f : S \to \mathbb{F}$.

   The scalar multiplication $a \cdot f$ is the function satisfying $(a \cdot f)(x) = a \cdot f(x)$.

The addition $f+g$ is the function satisfying $(f+g)(x) = f(x) + g(x)$.

$$a \cdot f : S \longrightarrow \mathbb{F} \qquad\qquad f+g : S \longrightarrow \mathbb{F}$$
$$x \longmapsto a \cdot f(x) \qquad\qquad x \longmapsto f(x) + g(x)$$

Many important examples arise in this way.

2.1. Let V be the set of continuous functions over $\mathbb{R}$ or over $\mathbb{C}$, denoted $\mathcal{C}(\mathbb{R})$ or $\mathcal{C}(\mathbb{C})$.

2.2. Let V be the set of polynomials with coefficients in $\mathbb{F}$, denoted $\mathbb{F}[x]$. Recall that $p(x) \in \mathbb{F}[x]$ has the form $p(x) = a_n x^n + \cdots + a_1 x + a_0$ for $a_n, \ldots, a_0 \in \mathbb{F}$.

2.3. Let V be the set of symmetric polynomials in n-variables, denoted $\mathrm{Sym}_n(\mathbb{F})$.

The elements are polynomials in the variables $x_1, \ldots, x_n$ such that:

$$p(x_1, \ldots, x_i, \ldots, x_j \ldots, x_n) = p(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n) \quad \text{for all } i,j \in \{1, \ldots, n\}.$$

That is, exchanging two variables does not change the polynomial.

Fix $n=3$, then:

$p(x_1, x_2, x_3) = x_1 + x_2 + x_3$ is symmetric,

$q(x_1, x_2, x_3) = x_1 + x_2$ is not symmetric since $q(x_1, x_3, x_2) = x_1 + x_3 \neq q(x_1, x_2, x_3)$.

$r(x_1, x_2, x_3) = x_1 x_2 + 2x_1 x_3 + x_2 x_3$ is not symmetric,

$S(x_1, x_2, x_3) = x_1x_2 + x_1x_3 + x_2x_3$ is symmetric.

3. Let $\mathbb{F}$ be a field, let V be the set of $n \times m$ matrices with entries in $\mathbb{F}$, denoted $M_{n\times m}(\mathbb{F})$. The matrix addition and scalar multiplication are both defined componentwise.

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nm} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1m}+b_{1m} \\ \vdots & & \vdots \\ a_{n1}+b_{n1} & \cdots & a_{nm}+b_{nm} \end{bmatrix}$$

$$a \cdot \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} = \begin{bmatrix} a \cdot a_{11} & \cdots & a \cdot a_{1m} \\ \vdots & & \vdots \\ a \cdot a_{n1} & \cdots & a \cdot a_{nm} \end{bmatrix}$$

The zero vector is the zero matrix.

$$\begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$$

In general, $M_{n\times m}(\mathbb{F})$ is not a field since we cannot multiply two $n \times m$ matrices.

4. Let V be the field of rational functions over $\mathbb{F}$, denoted $\mathbb{F}(x)$. Elements in $\mathbb{F}(x)$ are fractions of polynomials, namely $\frac{p(x)}{q(x)}$ with $p(x), q(x) \in \mathbb{F}[x]$. Now $\mathbb{F}[x]$ is a vector space over $\mathbb{F}$, and $\mathbb{F}[x]$ is also a field on its own.

The vector addition is:

$$\frac{p(x)}{q(x)} + \frac{r(x)}{s(x)} = \frac{p(x)s(x) + q(x)r(x)}{q(x)s(x)}$$

$$\frac{p(x)}{q(x)} + \frac{s(x)}{s(x)} = \frac{\quad}{q(x)\,s(x)}.$$

The scalar multiplication is:

$$a \cdot \frac{p(x)}{q(x)} = \frac{a \cdot p(x)}{q(x)}.$$

With these two operations, $\mathbb{F}(x)$ is a vector space over $\mathbb{F}$. Consider the sum:

$$\frac{p(x)}{q(x)} + \frac{r(x)}{s(x)} = \frac{p(x)\,s(x) + q(x)\,r(x)}{q(x)\,s(x)}$$

and the product:

$$\frac{p(x)}{q(x)} \cdot \frac{r(x)}{s(x)} = \frac{p(x)\,r(x)}{q(x)\,s(x)}.$$

With these two operations, $\mathbb{F}(x)$ is a field. The identities of $\mathbb{F}(x)$ are:

$z(x) = 0$ the zero, and $e(x) = 1$ the one.

Note that $\mathbb{F}(x)$ is closed since the sum and product give rational functions.

In fact, $\mathbb{F}(x)$ is also a vector space over $\mathbb{F}(x)$. The difference is that here the scalars have changed from $\mathbb{F}$ to $\mathbb{F}(x)$.

* Small aside on proof techniques:

We will mainly be using four techniques:

1. Induction: this is useful when proving something for all natural numbers.

Example: Prove that $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.

This is true for $n=1$ since $1 = \frac{1\cdot(1+1)}{2}$.

Suppose this is true for n. We now prove it for n:

$$\sum_{i=1}^{n+1} i = \sum_{i=1}^{n} i + (n+1) = \frac{n\cdot(n+1)}{2} + (n+1) = \frac{n\cdot(n+1)+2\cdot(n+1)}{2} =$$

$$= \frac{(n+1)\cdot(n+2)}{2} = \frac{(n+1)\cdot((n+1)+1)}{2}. \qquad \square.$$

2. Using the definition: this is useful when we do not know much.

Example: Prove that $\mathbb{Z}_2$ is a field. We only have to check that all the axioms hold true.

3. Using theorems and other results: this is useful when we know a lot.

Example: Prove that every $p(x) \in \mathbb{C}[x]$ factors into linear terms.

Let n be the degree of $p(x)$. If $n = 0, 1$ we are done. If $n \neq 0, 1$, by the Fundamental Theorem of Algebra $p(x)$ has one root $a_1 \in \mathbb{C}$. Then $x - a_1$ divides $p(x)$ so $p(x) = (x-a_1)\cdot q(x)$ with $q(x)$ of degree $n-1$. If $n-1 = 1$ then $q(x) = x - a_2$ so $p(x) = (x-a_1)(x-a_2)$ and we are done. If $n-1 \neq 1$, apply the Fundamental Theorem of Algebra again. Since at every step we are lowering the degree by 1, we will repeat this exactly n times, so we will have $p(x) = (x-a_1)\cdots(x-a_n)$. Thus $p(x)$ factors

into linear terms, as desired. □.

4. Follow your nose: when we are given several hypothesis and we are asked to verify that a statement holds, often we have to "put all the hypothesis in a box, shake it up a bit, and our desired conclusion will fall out."

Example: Prove that $\sqrt{2}$ is not a rational number.

Suppose that $\sqrt{2}$ is rational. We could then write $\sqrt{2} = \frac{a}{b}$ for $a, b \in \mathbb{Z}$.

Then squaring this we obtain $2 = \frac{a^2}{b^2}$ so $2b^2 = a^2$. Thus a^2 is even.

Since even times even is even, and odd times odd is odd, a is even.

Hence there exists $k \in \mathbb{Z}$ with $a = 2k$, so $2b^2 = (2k)^2 = 4k^2$. This means that $b^2 = 2k^2$, so as before b is even. All in all, we proved that if $\sqrt{2} = \frac{a}{b}$ then a and b are both even. However, rational numbers can be written in an irreducible way, that is, if $\sqrt{2} \in \mathbb{Q}$ then there are $p, q \in \mathbb{Z}$ such that p and q do not share any divisors and $\sqrt{2} = \frac{p}{q}$. This is a contradiction with what we just proved: p and q both should be divisible by 2. Thus $\sqrt{2} \notin \mathbb{Q}$. □.

End of the aside. *

We now prove some properties of vector spaces.

Theorem 1: Let V be a vector space. If $x, y, z \in V$ and $x+z = y+z$ then $x = y$.

Proof: Since $z \in V$, by axiom 4 there is $-z \in V$ with $z + (-z) = 0$. Hence:

$$x+z = y+z \Rightarrow (x+z)+(-z) = (y+z)+(-z)$$

Associativity (2) $\Rightarrow$ $x+(z+(-z)) = y+(z+(-z))$

Inverses (4) $\Rightarrow$ $x+0 = y+0$

Identity (3) $\Rightarrow$ $x = y$. $\square$.

Corollary 2: Let V be a vector space. The vector $\vec{0} \in V$ is unique.

Proof: Suppose that there is a vector $\vec{0}' \in V$ such that $z + \vec{0}' = z$ for all $z \in V$.

Now $z + \vec{0} = z = z + \vec{0}'$ so by Theorem 1 we have $\vec{0} = \vec{0}'$. Every vector in V that satisfies axiom 3 is equal to $\vec{0}$, so $\vec{0}$ is unique. $\square$.

Corollary 3: Let V be a vector space, fix $x \in V$. Then $-x \in V$ is unique.

Proof: Analogous to the above.

Given a mathematical object with structure, we always look at how that structure reappears in mathematical objects inside the original one. For sets, these are subsets. For

vector spaces, we look at vector subspaces.

Definition: Let V be a vector space over $\mathbb{F}$. A vector subspace W of V is a subset of V that is also a vector space with the addition and multiplication by scalars inherited from V.

Example:

1. Over $\mathbb{Q}$ we have $\mathbb{Q}^n \subsetneq \mathbb{R}^n \subsetneq \mathbb{C}^n$ are all subspaces of $\mathbb{C}^n$.
2. Over $\mathbb{Q}$ we have $\mathbb{Q}[x] \subsetneq \mathbb{R}[x] \subsetneq \mathbb{C}[x]$ are all subspaces of $\mathbb{C}[x]$.
3. Over $\mathbb{Q}$ we have $M_{n\times m}(\mathbb{Q}) \subsetneq M_{n\times m}(\mathbb{R}) \subsetneq M_{n\times m}(\mathbb{C})$ are all subspaces of $M_{n\times m}(\mathbb{C})$.
4. The symmetric polynomials in n variables $\mathrm{Sym}_n(\mathbb{F})$ are a subspace of the vector space of all the polynomials in n variables $\mathbb{F}[x_1, \cdots, x_n]$.

Theorem 4: Let V be a vector space. A subset W of V is a subspace of V if and only if all the following hold:

(1) $\vec{0} \in W$.

(2) $x+y \in W$ for all $x, y \in W$.

(3) $a \cdot x \in W$ for all $a \in \mathbb{F}$.

Proof: ($\Rightarrow$) Suppose that W is a subspace of V. We want to show that (1), (2), (3) hold. Since W is a vector space, there is $\vec{0}' \in W$ such that $w + \vec{0}' = w$ for all $w \in W$. Since W is a subset of V then $w \in V$ so $w + \vec{0} = w$. Now by Theorem 1 we have $\vec{0}' = \vec{0}$, so (1) holds. Since W is a vector space, (2) and (3) hold by definition.

($\Leftarrow$) Suppose that (1), (2), (3) hold. We have to verify that the following hold.

1. Commutativity of addition: it already is, $+$ is closed by (2).

2. Associativity of addition: it already is.

3. Identity in W: $\vec{0} \in W$ by (1).

4. Inverses in W: let $w \in W$, now $(-1) \cdot w \in W$ by (3), and $-w = (-1) \cdot w$. why? Prove it!

5. Scalar identity: let $w \in W$, since $w \in V$ then $1 \cdot w = w$, and by (3) we have $1 \cdot w \in W$, so $1 \cdot w = w$ is an equality in W.

6. Associativity of scalar multiplication: it already is.

7. Distributivity of scalar multiplication over addition: it already is.

8. Distributivity of the sum over scalar multiplication: it already is. $\square$.

This emphasizes that W needs to be a subset of V that is closed under the same addition

and multiplication by scalars as V.

A particularly important vector space is $M_{n\times m}(\mathbb{F})$. We now recall some definitions:

1. Column vectors: $\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{F}^n$ with $a_1, \dots, a_n \in \mathbb{F}$.

2. Row vectors: $(a_1, \dots, a_n) \in \mathbb{F}^n$ with $a_1, \dots, a_n \in \mathbb{F}$.

3. Square matrices: $M_{n\times n}(\mathbb{F})$, also denoted $M_n(\mathbb{F})$.

4. Diagonal matrices: $\begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix} \in M_{n\times n}(\mathbb{F})$ with $a_{ij} = 0$ for $i \neq j$.

5. Upper triangular matrices: $\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & & a_{nn} \end{bmatrix} \in M_{n\times n}(\mathbb{F})$ with $a_{ij} = 0$ for $i > j$.

   Similarly, we have lower triangular matrices.

5.1. Strictly upper triangular matrices: we instead require $a_{ij} = 0$ for $i \geq j$, so the diagonal entries are also zero.

   Similarly, we have strictly lower triangular matrices.

6. Zero matrix: $\begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \in M_{n\times m}(\mathbb{F})$ with $a_{ij} = 0$ for all i, j.

7. Identity matrix: $I = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \in M_{n\times n}(\mathbb{F})$ the diagonal matrix with $a_{ii} = 1$ for all i, and $a_{ij} = 0$ for $i \neq j$.

8. Transpose of a matrix: let $M = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \end{bmatrix} \in M_{n\times m}(\mathbb{F})$, its transpose is

8. Transpose of a matrix: let $M = \begin{bmatrix} \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in M_{mxn}(\mathbb{F})$ its <u>transpose</u> is $M^t = \begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix} \in M_{nxm}(\mathbb{F})$ the matrix with a_{ji} in the i-th row and j-th column.

9. Trace of a matrix: let $M = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \in M_{nxn}(\mathbb{F})$ its <u>trace</u> is:

$$tr(M) = \sum_{i=1}^{n} a_{ii}, \quad \text{the sum of the diagonal entries.}$$

<u>Examples:</u>

1. Symmetric matrices are a subspace of square matrices.

Recall that a matrix is symmetric when it is equal to its transpose.

<u>Proof:</u> We show that $\{A \in M_n(\mathbb{F}) \mid A^t = A\}$ is a subspace of $M_n(\mathbb{F})$.

(1) $0 = 0^t$ so 0 is symmetric.

(2) If A, B are symmetric then:

$$(A+B)^t = A^t + B^t = A + B \quad \text{so } A+B \text{ is symmetric.}$$

(3) If A is symmetric and $c \in \mathbb{F}$ then:

$$(cA)^t = c(A^t) = cA \quad \text{so } cA \text{ is symmetric.} \qquad \square.$$

2. Diagonal matrices are a subspace of square matrices.

3. Upper triangular matrices are a subspace of square matrices.

4. The set of the zero matrix $\{0\}$ is a subspace of $M_{n\times m}(\mathbb{F})$. We call it the zero subspace, or the trivial subspace, of $M_{n\times m}(\mathbb{F})$.

5. The set of traceless matrices, namely matrices whose trace is zero, is a subspace of square matrices.

6. The set of polynomials with degree less than or equal to n, denoted $P_n(\mathbb{F})$ or $\mathbb{F}_n[x]$, is a subspace of the vector space of polynomials $\mathbb{F}[x]$. Here we note that the zero polynomial has any degree we want.

7. We have the inclusions of subspaces:

$$\mathbb{R} \subseteq \mathbb{R}_n[x] \subseteq \mathbb{R}[x] \subseteq \mathcal{C}(\mathbb{R}) \subseteq \mathcal{F}(\mathbb{R}, \mathbb{R})$$

Now that we know many examples of vector spaces, we would like to construct new vector subspaces from old ones.

<u>Theorem 5</u>: Let V be a vector space, let U and W be subspaces of V. Then $U \cap W$ is a subspace of V.

<u>Proof</u>: We show that $U \cap W$ satisfies conditions (1), (2), (3) of Theorem 4.

(1) We want to show $\vec{0} \in U \cap W$. Since U and W are subspaces of V, then $\vec{0} \in U$ and $\vec{0} \in W$, so $\vec{0} \in U \cap W$.

(2) We want to show $x+y \in U \cap W$ when $x, y \in U \cap W$. Suppose $x, y \in U \cap W$, then $x, y \in U$ and $x, y \in W$. Since U and W are subspaces then $x+y \in U$ and $x+y \in W$, so $x+y \in U \cap W$.

(3) We want to show $c \cdot x \in U \cap W$ when $x \in U \cap W$ and $c \in \mathbb{F}$. Suppose $x \in U \cap W$ and $c \in \mathbb{F}$. Then $x \in U$ and $x \in W$. Since U and W are subspaces then $c \cdot x \in U$ and $c \cdot x \in W$, so $c \cdot x \in U \cap W$. □.

However, $U \cup W$ is almost never a subspace of V. The correct way of "putting two vector subspaces together" is by addition.

Definition: Let V be a vector space, let U and W be subspaces of V. The internal sum or sum of U and W, denoted $U+W$, is:

$$U+W = \{u+w \mid u \in U, w \in W\}.$$

Theorem 6: Let V be a vector space, let U and W be subspaces of V. Then $U+W$ is a vector subspace of V.

Proof: We show that $U+W$ satisfies conditions (1), (2), (3) of Theorem 4.

(1) Since $\vec{0} \in U$ and $\vec{0} \in W$, then $\vec{0} = \vec{0} + \vec{0} \in U+W$.

(2) Given $u+w, u'+w' \in U+W$ then:

$$(u+w)+(u'+w') = (u+u')+(w+w') \in U+W$$

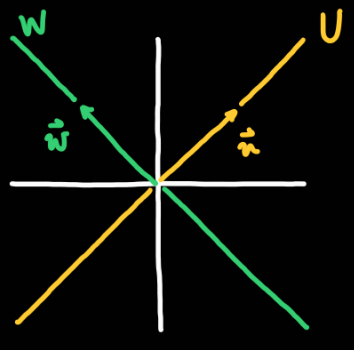
since $u+u' \in U$ and $w+w' \in W$.

(3) Given $u+w \in U+W$ and $c \in \mathbb{F}$ then:

$$c\cdot(u+w) = c\cdot u + c\cdot w \in U+W$$

since $c\cdot u \in U$ and $c\cdot w \in W$. $\square$.

Example: Let $V=\mathbb{R}^2$, let $\vec{u},\vec{w} \in V$ be non-zero and non-parallel. Set $U=\{r\cdot\vec{u} \mid r\in\mathbb{R}\}$ and $W=\{r\cdot\vec{w} \mid r\in\mathbb{R}\}$, these are vector subspaces of V. Now $U+W=V$.



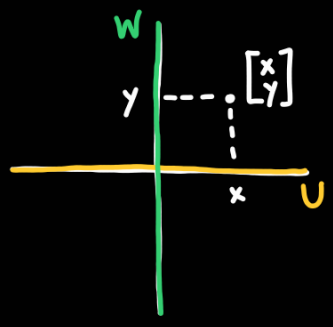
Definition: Let V be a vector space, let U and W be subspaces of V. We say that V is the direct sum of U and W, denoted $V=U\oplus W$, when the following hold.

(1) $V=U+W$.

(2) $U\cap W=\{\vec{0}\}$.

Example:

1. Let $V=\mathbb{R}^2$, let $U=\left\{\begin{bmatrix}x\\0\end{bmatrix} \mid x\in\mathbb{R}\right\}$, let $W=\left\{\begin{bmatrix}0\\y\end{bmatrix} \mid y\in\mathbb{R}\right\}$. Then $V=U\oplus W$.



2. We have: $M_n(\mathbb{F}) = \mathcal{U}_n(\mathbb{F}) \oplus \mathcal{D}_n(\mathbb{F}) \oplus \mathcal{L}_n(\mathbb{F})$ where:

$\mathcal{U}_n(\mathbb{F})$ are the strictly upper triangular matrices,

$\mathcal{D}_n(\mathbb{F})$ are the diagonal matrices,

$\mathcal{L}_n(\mathbb{F})$ are the strictly lower triangular matrices,

because any square matrix $A \in M_n(\mathbb{F})$ can be decomposed as:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2\,n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & a_{nn} \end{bmatrix} = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1\,n-1} & a_{1n} \\ 0 & 0 & \cdots & a_{2\,n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a_{n-1,n} \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} +$$

$$+ \begin{bmatrix} a_{11} & 0 & \cdots & 0 & 0 \\ 0 & a_{22} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{n-1,n-1} & 0 \\ 0 & 0 & \cdots & 0 & a_{nn} \end{bmatrix} + \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ a_{21} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & 0 & 0 \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & 0 \end{bmatrix}.$$

There is also a way of constructing a vector subspace from a collection of vectors.

<u>Definition:</u> Let V be a vector space. A vector $v \in V$ is said to be a <u>linear combination</u> of

the vectors $v_1, \dots, v_n \in V$ if there exist scalars $a_1, \dots, a_n \in \mathbb{F}$ such that:

$$v = a_1 v_1 + \cdots + a_n v_n.$$

The scalars $a_1, \dots, a_n \in \mathbb{F}$ are also called <u>coefficients</u>.

<u>Example:</u> Let $V = \mathbb{Q}[x]$. Since: $3x^5 + 2x - 1 = (5x^6 + 3x^5 - 2) + 2 \cdot (x^6 + x) - (7x^6 + 1)$

then $v = 3x^5 + 2x - 1$ is a linear combination of $v_1 = 5x^6 + 3x^5 - 2$, $v_2 = x^6 + x$,

$v_3 = 7x^6 - 1$ with coefficients $a_1 = 1$, $a_2 = 2$, $a_3 = -1$.

Also: $3x^5 + 2x + 1 = 3 \cdot (x^5) + 2 \cdot (x) - (1)$ so $v = 3x^5 + 2x - 1$ is also a linear

combination of $w_1 = x^5$, $w_2 = x$, $w_3 = 1$ with coefficients $b_1 = 3$, $b_2 = 2$, $a_3 = -1$.

<u>Definition:</u> Let V be a vector space, let $\{v_1, v_2, \dots\} \subseteq V$ be a (possibly infinite) subset of V.

The <u>span</u> of this subset, denoted $\text{Span}\{v_1, v_2, \dots\}$, is the set of all linear combinations

of $\{v_1, v_2, \dots\}$.

$$\text{Span}\{v_1, v_2, \dots\} = \{ a_{i_1} \cdot v_{i_1} + \cdots + a_{i_n} v_{i_n} \mid a_{i_1}, \dots, a_{i_n} \in \mathbb{F} \}.$$

<u>Theorem 7:</u> Let V be a vector space, let $\{v_1, v_2, \dots\} \subseteq V$ be a subset of V. Then

$\text{Span}\{v_1, v_2, \dots\}$ is a vector subspace of V.

<u>Proof:</u> We show that $\text{Span}\{v_1, v_2, \dots\}$ satisfies conditions (1), (2), (3) of Theorem 4.

(1) Note: $\vec{0} = 0 \cdot v_1 \in \text{Span}\{v_1, v_2, \dots\}$.

(2) Given $u, v \in \text{Span}\{v_1, v_2, \dots\}$ then there exist scalars $a_{i_1}, \dots, a_{i_n}, b_{j_1}, \dots, b_{j_m} \in \mathbb{F}$ such that $u = \sum_{k=1}^{n} a_{i_k} v_{i_k}$ and $v = \sum_{k=1}^{m} b_{j_k} \cdot v_{j_k}$. Suppose that the sets $\{v_{i_1}, \dots, v_{i_n}\}$ and $\{v_{j_1}, \dots, v_{j_m}\}$ have common elements, without loss of generality we may assume that $v_{i_1} = v_{j_1}, \dots, v_{i_\ell} = v_{j_\ell}$ for some ℓ.

Then:

$$u+v = \sum_{k=1}^{n} a_{i_k} v_{i_k} + \sum_{k=1}^{m} b_{j_k} \cdot v_{j_k} =$$

$$= \sum_{k=1}^{\ell} a_{i_k} v_{i_k} + \sum_{k=\ell+1}^{n} a_{i_k} v_{i_k} + \sum_{k=1}^{\ell} b_{j_k} v_{j_k} + \sum_{k=\ell+1}^{m} b_{j_k} v_{j_k} =$$

$$= \sum_{k=1}^{\ell} (a_{i_k} + b_{j_k}) \cdot v_{i_k} + \sum_{k=\ell+1}^{n} a_{i_k} v_{i_k} + \sum_{k=\ell+1}^{m} b_{j_k} v_{j_k}$$

is a linear combination of $v_{i_1}, \dots, v_{i_n}, v_{j_{\ell+1}}, \dots, v_{j_m}$, so $u+v \in \text{Span}\{v_1, v_2, \dots\}$.

(3) Given $u \in \text{Span}\{v_1, v_2, \dots\}$ and $c \in \mathbb{F}$ then there exist scalars $a_{i_1}, \dots, a_{i_n} \in \mathbb{F}$ such that $u = \sum_{k=1}^{n} a_{i_k} v_{i_k}$. Then:

$$c \cdot u = c \cdot \left(\sum_{k=1}^{n} a_{i_k} v_{i_k} \right) = \sum_{k=1}^{n} c \cdot (a_{i_k} \cdot v_{i_k}) = \sum_{k=1}^{n} (c \cdot a_{i_k}) \cdot v_{i_k}$$

is a linear combination of $v_{i_1}, \dots, v_{i_n}$, so $c \cdot u \in \text{Span}\{v_1, v_2, \dots\}$. □.

<u>Definition:</u> Let V be a vector space, let $\{v_1, v_2, \dots\} \subseteq V$ be a (possibly infinite) subset of V.

When $V = \text{Span}\{v_1, v_2, \dots\}$ we say that $\{v_1, v_2, \dots\}$ <u>generate</u> or <u>span</u> V. We call the

elements in $\{v_1, v_2, \dots\}$ the <u>generators</u> of V.

<u>Example:</u> Let $V = \mathbb{R}^3$ and $v_1 = \begin{bmatrix}1\\1\\0\end{bmatrix}$, $v_2 = \begin{bmatrix}1\\0\\1\end{bmatrix}$, $v_3 = \begin{bmatrix}0\\1\\1\end{bmatrix}$. Since any vector $\begin{bmatrix}a\\b\\c\end{bmatrix} \in \mathbb{R}^3$ can be expressed as:

$$\begin{bmatrix}a\\b\\c\end{bmatrix} = \tfrac{1}{2}(a+b-c)\begin{bmatrix}1\\1\\0\end{bmatrix} + \tfrac{1}{2}(a-b+c)\begin{bmatrix}1\\0\\1\end{bmatrix} + \tfrac{1}{2}(-a+b+c)\begin{bmatrix}0\\1\\1\end{bmatrix}$$

then $\mathbb{R}^3 = \text{Span}\left\{\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}\right\}$.

<u>Definition:</u> Let V be a vector space, let $v_1, \dots, v_n \in V$. We say that $v_1, \dots, v_n$ are <u>linearly dependent</u> if there exist scalars $a_1, \dots, a_n \in \mathbb{F}$, at least one of them non-zero, such that $a_1 v_1 + \cdots + a_n v_n = \vec{0}$.

<u>Example:</u> Let $V = \mathbb{R}^3$ and $v_1 = \begin{bmatrix}1\\1\\0\end{bmatrix}$, $v_2 = \begin{bmatrix}2\\3\\0\end{bmatrix}$, $v_3 = \begin{bmatrix}0\\\frac{1}{2}\\0\end{bmatrix}$. Now:

$$v_2 - 2v_1 = \begin{bmatrix}2\\3\\0\end{bmatrix} - 2\begin{bmatrix}1\\1\\0\end{bmatrix} = \begin{bmatrix}0\\1\\0\end{bmatrix} = 2v_3 \quad \text{so} \quad v_3 = \tfrac{1}{2}(v_2 - 2v_1) = \tfrac{1}{2}v_2 - v_1$$

and thus $-v_1 + \frac{1}{2}v_2 - v_3 = \vec{0}$, so v_1, v_2, v_3 are linearly dependent.

<u>Definition:</u> Let V be a vector space, let $v_1, \dots, v_n \in V$. We say that $v_1, \dots, v_n$ are <u>linearly independent</u> if they are not linearly dependent.

<u>Remark:</u> The vectors $v_1, \dots, v_n$ are linearly independent if and only if when a linear combination $a_1 v_1 + \cdots + a_n v_n = \vec{0}$ in V then $a_1 = \cdots = a_n = 0$ in $\mathbb{F}$.

<u>Example:</u> Let V be a vector space, let $v \in V$, let $a \in \mathbb{F}$.

1. The empty set is linearly independent.

2. A single vector $\{v\}$ is linearly independent if and only if $v \neq \vec{0}$.

3. A vector and a scalar multiple of that vector $\{v, av\}$ are linearly dependent.

Theorem 8: Let $v_1, \ldots, v_n \in \mathbb{R}^m$. The vectors $v_1, \ldots, v_n$ are linearly independent if and only if the matrix $\begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}$ can be row reduced to the identity matrix.

Theorem 9: Let V be a vector space, let $\{v_1, \ldots, v_n\} \subset V$ be a linearly independent subset, let $v_{n+1} \in V$. The set $\{v_1, \ldots, v_n, v_{n+1}\}$ is linearly independent if and only if $v_{n+1} \notin \text{Span}\{v_1, \ldots, v_n\}$.

Proof: ($\Rightarrow$) Suppose that $\{v_1, \ldots, v_n, v_{n+1}\}$ is linearly independent. Additionally, assume that $v_{n+1} \in \text{Span}\{v_1, \ldots, v_n\}$, we want to achieve a contradiction. Since $v_{n+1} \in \text{Span}\{v_1, \ldots, v_n\}$ then there are scalars $a_1, \ldots, a_n \in \mathbb{F}$, at least one of them not zero, such that $v_{n+1} = a_1 v_1 + \cdots + a_n v_n$. Thus $a_1 v_1 + \cdots + a_n v_n + (-1) \cdot v_{n+1} = \vec{0}$ is a linear combination of elements in $\{v_1, \ldots, v_{n+1}\}$ with at least one of the coefficients non-zero. This means that $\{v_1, \ldots, v_{n+1}\}$ is linearly dependent, a contradiction. Since we reached a contradiction, our additional assumption of

$v_{n+1} \in \text{Span}\{v_1, \ldots, v_n\}$ is false, and thus $v_{n+1} \notin \text{Span}\{v_1, \ldots, v_n\}$.

($\Leftarrow$) Suppose that $v_{n+1} \notin \text{Span}\{v_1, \ldots, v_n\}$. Additionally, assume that $\{v_1, \ldots, v_n, v_{n+1}\}$ is linearly dependent, we want to achieve a contradiction. Since $\{v_1, \ldots, v_n, v_{n+1}\}$ is linearly dependent, then there are scalars $a_1, \ldots, a_{n+1} \in \mathbb{F}$, at least one of them not zero, such that $a_1 v_1 + \cdots + a_n v_n + a_{n+1} v_{n+1} = \vec{0}$. If $a_{n+1} = 0$ then $a_1 v_1 + \cdots + a_n v_n = \vec{0}$ is a linear combination of elements in $\{v_1, \ldots, v_n\}$ with at least one non-zero coefficient, meaning that $\{v_1, \ldots, v_n\}$ is linearly dependent. However, $\{v_1, \ldots, v_n\}$ is linearly independent by hypothesis, yielding a contradiction. Thus $a_{n+1} = 0$ is not possible. If $a_{n+1} \neq 0$ then it has a multiplicative inverse $\frac{1}{a_{n+1}}$ in $\mathbb{F}$, so we can rewrite $a_1 v_1 + \cdots + a_n v_n + a_{n+1} v_{n+1} = \vec{0}$ as $a_{n+1} v_{n+1} = -a_1 v_1 - \cdots - a_n v_n$ and thus $v_{n+1} = \frac{-a_1}{a_{n+1}} v_1 + \cdots + \frac{-a_n}{a_{n+1}} v_n$ where at least one of $\frac{-a_1}{a_{n+1}}, \ldots, \frac{-a_n}{a_{n+1}}$ is not zero. Thus $v_{n+1} \in \text{Span}\{v_1, \ldots, v_n\}$, a contradiction. Since we reached a contradiction both when $a_{n+1} = 0$ and when $a_{n+1} \neq 0$, our original assumption of $\{v_1, \ldots, v_n, v_{n+1}\}$ being linearly dependent is false, and thus $\{v_1, \ldots, v_n, v_{n+1}\}$ is linearly independent. $\square$.

Corollary 10: Let V be a vector space, let $S \subset V$ be a (not necessarily finite) set of

linearly independent vectors, let $v \in V$. Then the set $S \cup \{v\}$ is linearly independent if and only if $v \notin \text{Span}(S)$.

Examples:

1. Let $V = \mathbb{R}^n$, let e_i be the vector with a 1 in the i-th entry and 0 in all other entries, so for $i = 1, \dots, n$ we have $e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, e_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$. Then $e_1, \dots, e_n$ are linearly independent and $\mathbb{R}^n = \text{Span}\{e_1, \dots, e_n\}$. These $e_1, \dots, e_n$ are known as the standard generators of $\mathbb{R}^n$.

2. Let $V = M_2(\mathbb{C})$, $M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $M_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Then M_1 and M_2 are linearly independent, since having scalars $a, b \in \mathbb{C}$ with $aM_1 + bM_2 = \vec{0}$ means:

$$a \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \text{ namely } \begin{bmatrix} a+b & 0 \\ 0 & a-b \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

hence $a + b = 0$ and $a - b = 0$, so $a = b = 0$.

3. Let $V = \mathbb{R}[x]$. Then $\{1, x, x^2, \dots\}$ is a linearly independent set. This can be proven by induction: suppose that there are exponents $n \in \mathbb{N}$ with x^n being a linear combination of monomials of lower degree, namely $x^n \in \text{Span}\{1, x, \dots, x^{n-1}\}$.

First, if $n = 1$, this would mean $x \in \text{Span}\{1\}$. Thus $x = a \cdot 1$ for some non-zero coefficient $a \in \mathbb{R}$. Evaluating at $x = 0$ this gives $0 = a$, a contradiction.

Hence $x \notin \text{Span}\{1\}$ and thus by Theorem 9 the set $\{1, x\}$ is linearly independent.

Suppose the statement is true for $n-1$, namely the set $\{1, x, \dots, x^{n-1}\}$ is linearly independent. We now prove the statement for n.

Suppose that $x^n \in \text{Span}\{1, x, \dots, x^{n-1}\}$. Then for some non-zero coefficients $a_0, \dots, a_{n-1} \in \mathbb{R}$ we have $x^n = a_0 \cdot 1 + a_1 \cdot x + \cdots + a_{n-1} \cdot x^{n-1}$. Evaluating at $x = 0$ this gives $0 = a_0$, so in fact we have $x^n = a_1 \cdot x + \cdots + a_{n-1} \cdot x^{n-1}$. This yields $x \cdot x^{n-1} = x \cdot (a_1 + \cdots + a_{n-1} \cdot x^{n-2})$ and thus $x^{n-1} = a_1 + \cdots + a_{n-1} \cdot x^{n-2}$. Hence $x^{n-1} \in \text{Span}\{1, x, \dots, x^{n-2}\}$, so by Theorem 9 the set $\{1, x, \dots, x^{n-1}\}$ is linearly dependent. This is a contradiction with the induction hypothesis, which was that $\{1, x, \dots, x^{n-1}\}$ is linearly independent. Thus our original assumption of $x^n \in \text{Span}\{1, x, \dots, x^{n-1}\}$ is false, meaning that $x^n \notin \text{Span}\{1, x, \dots, x^{n-1}\}$.

Hence by Theorem 9 the set $\{1, x, \dots, x^n\}$ is linearly independent.

To finish the reasoning, suppose that $\{1, x, x^2, \dots\}$ is linearly dependent, we want to achieve a contradiction. Then there are non-zero coefficients $a_{i_1}, \dots, a_{i_n} \in \mathbb{R}$ such that $a_{i_1} \cdot x^{i_1} + \cdots + a_{i_n} \cdot x^{i_n} = 0$, which can be rearranged

$x^{i_n} = -a_{i_1} \cdot x^{i_1} - \cdots - a_{i_{n-1}} \cdot x^{i_{n-1}}$. Hence $x^{i_n} \in \text{Span}\{x^{i_1}, \dots, x^{i_{n-1}}\}$

as $x^{i_n} = \frac{1}{a_{i_n}} \cdot x^{i_1} + \cdots + \frac{a_{i_{n-1}}}{a_{i_n}} \cdot x^{i_{n-1}}$. Hence $x^{i_n} \in \text{Span}\{x^{i_1}, \ldots, x^{i_{n-1}}\}$, in particular $x^{i_n} \in \text{Span}\{1, x, \ldots, x^{i_n - 1}\}$, so by Theorem 9 the set $\{1, x, \ldots, x^{i_n - 1}, x^{i_n}\}$ is linearly dependent, so the set $\{1, x, \ldots, x^{i_n - 1}, x^{i_n - 1 + 1}, \ldots, x^{i_n}\}$ is linearly dependent. This is a contradiction, since we just proved that the set $\{1, x, \ldots, x^n\}$ is linearly independent for all $n \in \mathbb{N}$. Thus our assumption that the set $\{1, x, x^2, \ldots\}$ is linearly dependent is false, meaning that the set $\{1, x, x^2, \ldots\}$ is linearly independent.

4. The vectors $\{(-2,0,3), (1,3,0), (2,4,-1)\}$ are linearly dependent. We consider the matrix with columns $(-2,0,3), (1,3,0), (2,4,-1)$, we row reduce, and we obtain:

$$\begin{bmatrix} -2 & 1 & 2 \\ 0 & 3 & 4 \\ 3 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 + R_3} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & 4 \\ 3 & 0 & -1 \end{bmatrix} \xrightarrow{R_3 - 3R_1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & 4 \\ 0 & -3 & -4 \end{bmatrix} \xrightarrow{R_3 + R_2} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & 4 \\ 0 & 0 & 0 \end{bmatrix}$$

This matrix in row-reduced echelon form is not the identity matrix, so the vectors $(-2,0,3), (1,3,0), (2,4,-1)$ are not linearly independent by Theorem 8.

5. Solve the system of linear equations:

$$\begin{aligned} 3x_1 - 7x_2 + 4x_3 &= 10 \\ x_1 - 2x_2 + x_3 &= 3 \\ 2x_1 - x_2 - 2x_3 &= 6. \end{aligned}$$

We can rewrite this system as an equation of matrices:

$$\begin{bmatrix}3\\1\\2\end{bmatrix}x_1+\begin{bmatrix}-7\\-1\\-2\end{bmatrix}x_2+\begin{bmatrix}4\\1\\-2\end{bmatrix}x_3=\begin{bmatrix}10\\3\\6\end{bmatrix} \quad \text{equivalently} \quad \begin{bmatrix}3&-7&4\\1&-2&1\\2&-1&-2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}10\\3\\6\end{bmatrix}.$$

This system will have a unique solution if and only if the vectors $(3,1,2)$, $(-7,-1,-2)$, $(4,1,2)$ are linearly independent. We use row reduction to find it.

$$\left[\begin{array}{ccc|c}3&-7&4&10\\1&-2&1&3\\2&-1&-2&6\end{array}\right] \xrightarrow[R_1-R_3]{R_3-2R_2} \left[\begin{array}{ccc|c}1&-6&6&4\\1&-2&1&3\\0&3&-4&0\end{array}\right] \xrightarrow{R_2-R_1} \left[\begin{array}{ccc|c}1&-6&6&4\\0&4&-5&-1\\0&3&-4&0\end{array}\right] \xrightarrow[R_2-R_3]{R_1+2R_3}$$

$$\left[\begin{array}{ccc|c}1&0&-2&4\\0&1&-1&-1\\0&3&-4&0\end{array}\right] \xrightarrow[-R_3]{R_3-3R_2} \left[\begin{array}{ccc|c}1&0&-2&4\\0&1&-1&-1\\0&0&1&-3\end{array}\right] \xrightarrow[R_2+R_3]{R_1+2R_3} \left[\begin{array}{ccc|c}1&0&0&-2\\0&1&0&-4\\0&0&1&-3\end{array}\right].$$

Hence $x_1=-2$, $x_2=-4$, $x_3=-3$.

In fact, we always have $V=\text{Span}(V)$ for all vector spaces V. The question to ask is whether there is a subset $S\subset V$ such that $V=\text{Span}(S)$ and such that S has the minimum possible number of elements. We call those subsets a basis of the vector space V.

Definition: Let V be a vector space. A subset $\{v_1,v_2,\dots\}\subset V$ is called a basis of V when $V=\text{Span}\{v_1,v_2,\dots\}$ and $\{v_1,v_2,\dots\}$ is linearly independent.

Examples:

1. Let $V=\mathbb{F}^n$, let e_i be the vector with a 1 in the i-th entry and 0 in all other

entries, so for $i=1,\dots,n$ we have $e_1=\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}, \dots, e_n=\begin{bmatrix}0\\\vdots\\0\\1\end{bmatrix}$. Then $e_1,\dots,e_n$ are linearly independent and $\mathbb{F}^n = \text{Span}\{e_1,\dots,e_n\}$, so $\{e_1,\dots,e_n\}$ is a basis of $\mathbb{F}^n$.

2. Let $V=\mathbb{F}[x]$, the set $\{1,x,x^2,\dots\}$ is a basis of $\mathbb{F}[x]$.

3. Let $V=\mathbb{F}_n[x]$ be the vector space of polynomials of degree at most n, the set $\{1,x,\dots,x^n\}$ is a basis of $\mathbb{F}_n[x]$.

4. Let $V=M_{n\times m}(\mathbb{F})$, let E_{ij} be the matrix with a 1 in the ij-th entry and 0 in all other entries. The set $\{E_{ij}\}_{\substack{1\le j\le m\\ 1\le i\le n}}$ is a basis of $M_{n\times m}(\mathbb{F})$.

<u>Theorem 11:</u> Let V be a vector space. Let β be a finite set that is also a basis of V. Let β' be another basis of V. Then β and β' have the same number of elements, namely $|\beta|=|\beta'|$.

The above result is a consequence of the Replacement Theorem and of Theorem 9. In fact, Theorem 11 also holds when β is infinite. This enables the following definition.

<u>Definition:</u> Let V be a vector space, let β be a basis of V. The <u>dimension</u> of V, denoted $\dim_{\mathbb{F}}(V)$, is the number of elements in β.

A vector space can be finite dimensional or infinite dimensional.

Examples:

1. $\dim_{\mathbb{R}}(\mathbb{R}^n) = n$.

2. $\dim_{\mathbb{R}}(\mathbb{C}^n) = 2n$, $\dim_{\mathbb{C}}(\mathbb{C}^n) = n$.

3. $\dim_{\mathbb{F}}(M_{n\times m}(\mathbb{F})) = n \cdot m$.

4. $\dim_{\mathbb{F}}(\mathbb{F}[x]) = |\mathbb{N}|$.

Theorem 12: Let V be a vector space. The set $\{v_1, \dots, v_n\}$ is a basis of V if and only if every $v \in V$ can be expressed as a linear combination $v = a_1 v_1 + \dots + a_n v_n$ with $a_1, \dots, a_n \in \mathbb{F}$ in an unique way. Namely if $a_1 v_1 + \dots + a_n v_n = b_1 v_1 + \dots + b_n v_n$ for $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{F}$ then $a_1 = b_1, \dots, a_n = b_n$.

We can now answer the question of the existence of basis: every vector space has a basis.

Theorem 13: Let V be a vector space, $V = \operatorname{Span}\{v_1, \dots, v_n\}$. Then there exists a subset $\beta \subseteq \{v_1, \dots, v_n\}$ that is a basis of V.

Proof: We prove this by starting with the empty set, and adding one vector at a time. Let n be the number of vectors spanning V, namely $V = \operatorname{Span}\{v_1, \dots, v_n\}$. If $n = 0$

then $V = \text{Span}\{\}$. Now $V = 0$ and $\beta = \{\}$ is a basis. If $n = 1$ then $V = \text{Span}\{v_1\}$.

Now if $v_1 = \vec{0}$ then $V = 0$ and $\beta = \{\}$ is a basis, and if $v_1 \neq \vec{0}$ then $V \neq 0$ and $\beta = \{v_1\}$ is a basis. If $n \neq 0, 1$ then $V = \text{Span}\{v_1, \dots, v_n\}$. Consider $\text{Span}\{v_1\} \subseteq V$, if $\text{Span}\{v_1\} = V$ then $\beta = \{v_1\}$ is a basis. If $\text{Span}\{v_1\} \neq V$ check whether v_2 is in $\text{Span}\{v_1\}$. If $v_2 \notin \text{Span}\{v_1\}$ then consider $\text{Span}\{v_1, v_2\} \subseteq V$ with $\beta = \{v_1, v_2\}$ as potential basis. If $v_2 \in \text{Span}\{v_1\}$ check whether v_3 is in $\text{Span}\{v_1\}$.

Repeating this process, we obtain at every step a subset $\{v_{i_1}, v_{i_2}, \dots, v_{i_k}\} \subseteq \{v_1, \dots, v_n\}$ that is linearly independent. Moreover since at every step we always have that the vectors ommited in the candidate basis are in the span of the candidate basis, $\text{Span}\{v_{i_1}, \dots, v_{i_k}\} \supseteq \text{Span}\{v_1, \dots, v_n\} = V$ and thus $V = \text{Span}\{v_{i_1}, \dots, v_{i_k}\}$ so $\beta = \{v_{i_1}, \dots, v_{i_k}\}$ is a basis of V. Since we are only checking a finite number of conditions, because $\{v_1, \dots, v_n\}$ is finite, this process terminates. $\square$.

This proves that a finitely generated vector space has a basis. It does not say that infinitely generated vector spaces have a basis; although that statement is true, it requires a different proof.

Remark: There are several important results related to Theorem 13.

1. Any finite spanning set can be reduced to a basis.

2. Any spanning set with more vectors than $\dim_{\mathbb{F}}(V)$ is not a basis.

3. Any set with fewer vectors than $\dim_{\mathbb{F}}(V)$ does not span V.

Example: Let $V = \mathbb{R}^3$, let $S = \{(2,-3,5),(8,-12,20),(1,0,-2),(0,2,-1),(7,2,0)\}$. Check if we can extract a basis of V from S, and do so.

Since $\mathbb{R}^3$ has dimension 3, a basis for $\mathbb{R}^3$ has exactly 3 vectors. To build a basis, we follow the proof of Theorem 13:

Step 1: Pick two vectors in S that are not linearly dependent.

Since $(2,-3,5) \neq a\cdot(1,0,-2)$ for all $a \in \mathbb{R}$, these are enough.

Step 2: Pick one vector in S that is not in $\text{Span}\{(2,-3,5),(1,0,-2)\}$. We can simplify this a bit further using row reduction:

$$\begin{bmatrix} 2 & -3 & 5 \\ 1 & 0 & -2 \end{bmatrix} \xrightarrow{R_1 - 2R_2} \begin{bmatrix} 0 & -3 & 9 \\ 1 & 0 & -2 \end{bmatrix} \xrightarrow{\frac{-1}{3}R_1} \begin{bmatrix} 0 & 1 & -3 \\ 1 & 0 & -2 \end{bmatrix}$$

and thus $\text{Span}\{(2,-3,5),(1,0,-2)\} = \text{Span}\{(0,1,-3),(1,0,-2)\}$ so we immediately see that:

$$(8,-12,20) \in \text{Span}\{(0,1,-3),(1,0,-2)\},$$

$(0,2,-1), (7,2,0) \notin \text{Span}\{(0,1,-3),(1,0,-2)\}$.

Thus $\beta = \{(2,-3,5),(1,0,-2),(0,2,-1)\}$ and $\beta' = \{(2,-3,5),(1,0,-2),(7,2,0)\}$ are basis of $\mathbb{R}^3$.

* Aside on cosets.

A fundamental construction in mathematics is the concept of equivalence, which gives relations between objects. One of the embodyments of equivalence relations are cosets.

Definition: Let V be a vector space, let $W \subseteq V$ be a vector subspace, let $v \in V$. The set

$v + W = \{v + w \mid w \in W\}$ is called a **coset**.

In general cosets are not vector subspaces of V, just subsets.

Definition: Let V be a vector space, let $W \subseteq V$ be a vector subspace. The set formed by the sets $v + W$ for $v \in V$ is called the **quotient space** of V modulo W, denoted V/W. Namely, $V/W = \{v + W \mid v \in V\}$.

Theorem 14: Let V be a vector space, let $W \subseteq V$ be a vector subspace. The set V/W is a vector space over $\mathbb{F}$ with the operations:

$$+ : V/W \times V/W \longrightarrow V/W \qquad (v_1 + W, v_2 + W) \mapsto (v_1 + v_2) + W)$$

and

$$\cdot : \mathbb{F} \times V/W \longrightarrow V/W \qquad (a, v + W) \longmapsto (a \cdot v) + W)$$

Example: Let $V=\mathbb{R}^3$, let $W=\operatorname{Span}\left\{\begin{bmatrix}0\\0\\0\end{bmatrix}\right\}=\left\{\begin{bmatrix}0\\0\\0\end{bmatrix}\,\middle|\, r\in\mathbb{R}\right\}$. We know that any vector v in $\mathbb{R}^3$ has the form $\begin{bmatrix}a\\b\\c\end{bmatrix}$ with $a,b,c\in\mathbb{R}$, so such a vector is in W if and only if $b=c=0$. Now:

$$v+W=\left\{\begin{bmatrix}a\\b\\c\end{bmatrix}+\begin{bmatrix}r\\0\\0\end{bmatrix}\,\middle|\, r\in\mathbb{R}\right\}=\left\{\begin{bmatrix}a+r\\b\\c\end{bmatrix}\,\middle|\, r\in\mathbb{R}\right\}.$$

In particular if $v\in W$ then it has the form $\begin{bmatrix}a\\0\\0\end{bmatrix}$ for some $a\in\mathbb{R}$, hence:

$$v+W=\left\{\begin{bmatrix}a\\0\\0\end{bmatrix}+\begin{bmatrix}r\\0\\0\end{bmatrix}\,\middle|\, r\in\mathbb{R}\right\}=\left\{\begin{bmatrix}a+r\\0\\0\end{bmatrix}\,\middle|\, r\in\mathbb{R}\right\}=\left\{\begin{bmatrix}s\\0\\0\end{bmatrix}\,\middle|\, s\in\mathbb{R}\right\}=W.$$

The space V/W is formed by the sets:

$$\begin{bmatrix}0\\b\\c\end{bmatrix}+W=\left\{\begin{bmatrix}r\\b\\c\end{bmatrix}\,\middle|\, r\in\mathbb{R}\right\}\qquad \text{for each choice of } b,c\in\mathbb{R}.$$

Now V/W is spanned by the two sets $\begin{bmatrix}0\\1\\0\end{bmatrix}+W$ and $\begin{bmatrix}0\\0\\1\end{bmatrix}+W$, and they are linearly independent since for all $s\in\mathbb{R}$:

$$s\cdot\left(\begin{bmatrix}0\\1\\0\end{bmatrix}+W\right)=\left(s\cdot\begin{bmatrix}0\\1\\0\end{bmatrix}\right)+W=\begin{bmatrix}0\\s\\0\end{bmatrix}+W \quad \text{is different to} \quad \begin{bmatrix}0\\0\\1\end{bmatrix}+W.$$

Hence $\left\{\begin{bmatrix}0\\1\\0\end{bmatrix}+W,\ \begin{bmatrix}0\\0\\1\end{bmatrix}+W\right\}$ is a basis of V/W.

Moreover, the set W is the zero vector in V/W.

End of the aside. *

Theorem 15: (Replacement Theorem). Let V be a vector space, $V = \text{Span}\{v_1, \dots, v_n\}$, let $\{u_1, \dots, u_m\} \subset V$ be linearly independent. Then:

1) The cardinality of $\{u_1, \dots, u_m\}$ is, at most, the cardinality of $\{v_1, \dots, v_n\}$.

Namely $m \leq n$.

2) There is a subset $\{v_{i_1}, \dots, v_{i_{n-m}}\} \subseteq \{v_1, \dots, v_n\}$ containing $n-m$ vectors

such that $V = \text{Span}\{u_1, \dots, u_m, v_{i_1}, \dots, v_{i_{n-m}}\}$.

Proof: We use induction on m. Suppose first that $m = 0$, now $0 \leq n$ and setting

$v_{i_j} = v_j$ for $j = 1, \dots, n$ then $V = \text{Span}\{v_1, \dots, v_n\} = \text{Span}\{v_{i_1}, \dots, v_{i_{n-0}}\}$, as

desired. Suppose second that the statement holds for $m-1$, namely given any

linearly independent set with $m-1$ elements $\{w_1, \dots, w_{m-1}\}$ then $m-1 \leq n$ and

there is a subset $\{v_{i_1}, \dots, v_{i_{n-(m-1)}}\} \subset \{v_1, \dots, v_n\}$ such that

$V = \text{Span}\{w_1, \dots, w_{m-1}, v_{i_1}, \dots, v_{i_{n-(m-1)}}\}$. We now prove that the statement

holds for m. Let $\{u_1, \dots, u_m\} \subset V$ be linearly independent, then $\{u_1, \dots, u_{m-1}\}$

is linearly independent, so by induction hypothesis we have $m-1 \leq n$ and

a subset $\{v_{i_1}, \dots, v_{i_{n-m+1}}\} \subset \{v_1, \dots, v_n\}$ with

$V = \text{Span}\{u_1, \dots, u_{m-1}, v_{i_1}, \dots, v_{i_{n-m+1}}\}$. Since $u_m \in V$, we can write:

$$u_m = a_1 u_1 + \cdots + a_{m-1} u_{m-1} + a_m v_{i_1} + \cdots + a_n v_{i_{n-m+1}}$$

where $a_1, \ldots, a_m \in \mathbb{F}$, at least one of them is not zero. If $n = m-1$ then this can be simplified to $u_m = a_1 u_1 + \cdots + a_{m-1} u_{m-1}$, where at least one of the coefficients is not zero. Thus $u_m \in \text{Span}\{u_1, \ldots, u_{m-1}\}$ so by Theorem 9 then $\{u_1, \ldots, u_{m-1}, u_m\}$ is linearly dependent, a contradiction. If $n \neq m-1$, since we know $m-1 \leq n$, then $m-1 < n$ so $m \leq n$, giving the first part of the result. Moreover if all the coefficients $a_m, \ldots, a_n$ are zero then $u_m = a_1 u_1 + \cdots + a_{m-1} u_{m-1}$, where at least one of the coefficients is not zero, giving a contradiction as before. Thus at least one of the coefficients $a_m, \ldots, a_n$ is not zero, say $a_j \neq 0$ for $m \leq j \leq n$. Thus we can rearrange $u_m = a_1 u_1 + \cdots + a_{m-1} u_{m-1} + a_m v_{i_1} + \cdots + a_n v_{i_{n-m+1}}$ as:

$$v_{i_j} = -\frac{a_1}{a_j} u_1 - \cdots - \frac{a_{m-1}}{a_j} u_{m-1} + \frac{1}{a_j} u_m$$
$$- \frac{a_m}{a_j} v_{i_1} - \cdots - \frac{a_{(m-1)+(j-1)}}{a_j} v_{i_{j-1}} - \frac{a_{(m-1)+(j+1)}}{a_j} v_{i_{j+1}} - \cdots - \frac{a_n}{a_j} v_{i_{n-m+1}}.$$

Hence $v_{i_j} \in \text{Span}\{u_1, \ldots, u_m, v_{i_1}, \ldots, v_{i_{j-1}}, v_{i_{j+1}}, \ldots, v_{i_{n-m+1}}\}$, meaning that $\{u_1, \ldots, u_m, v_{i_1}, \ldots, v_{i_{n-m+1}}\} \subset \text{Span}\{u_1, \ldots, u_m, v_{i_1}, \ldots, v_{i_{j-1}}, v_{i_{j+1}}, \ldots, v_{i_{n-m+1}}\}$, so

$V = \text{Span}\{u_1, \ldots, u_m, v_{i_1}, \ldots, v_{i_{n-m+1}}\} \subset \text{Span}\{u_1, \ldots, u_m, v_{i_1}, \ldots, v_{i_{j-1}}, v_{i_{j+1}}, \ldots, v_{i_{n-m+1}}\}$

and of course $\text{Span}\{u_1,\dots,u_m,v_{i_1},\dots,v_{i_{j-1}},v_{i_{j+1}},\dots,v_{i_{n-m+1}}\} \subseteq V$, whence

$V = \text{Span}\{u_1,\dots,u_m,v_{i_1},\dots,v_{i_{j-1}},v_{i_{j+1}},\dots,v_{i_{n-m+1}}\}$. Since the set

$\{v_{i_1},\dots,v_{i_{j-1}},v_{i_{j+1}},\dots,v_{i_{n-m+1}}\}$ has $(n-m+1)-1 = n-m$ elements and is a subset

of $\{v_1,\dots,v_n\}$, this yields the second part of the result. $\square$.

Examples:

1. Let $V = \mathbb{R}^3$, note that $V = \text{Span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$. The set $\left\{\begin{bmatrix}1\\1\\0\end{bmatrix}\right\}$ is linearly independent. Now Theorem 15 says that we can find a subset of $\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$ that complements $\left\{\begin{bmatrix}1\\1\\0\end{bmatrix}\right\}$, that is, two of the vectors $\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix}$ together with $\begin{bmatrix}1\\1\\0\end{bmatrix}$ span $\mathbb{R}^3$. There is more than one choice:

   (a) $\mathbb{R}^3 \neq \text{Span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}1\\1\\0\end{bmatrix}\right\}$.

   (b) $\mathbb{R}^3 = \text{Span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix},\begin{bmatrix}1\\1\\0\end{bmatrix}\right\}$.

   (c) $\mathbb{R}^3 = \text{Span}\left\{\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix},\begin{bmatrix}1\\1\\0\end{bmatrix}\right\}$.

2. Let $V = M_{2\times 2}(\mathbb{Z}_3)$ be the vector space of 2 by 2 matrices with entries in $\mathbb{Z}_3$. Now $V = \text{Span}\left\{\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\end{bmatrix}\right\}$. The set $\left\{\begin{bmatrix}1&2\\0&0\end{bmatrix},\begin{bmatrix}0&0\\2&0\end{bmatrix}\right\}$ is linearly independent, and there are two subsets of the

original generating set that complement it:

(a) $M_{2\times 2}(\mathbb{F}) = \text{Span}\left\{\begin{bmatrix}1&0\\0&0\end{bmatrix}, \begin{bmatrix}0&0\\0&1\end{bmatrix}, \begin{bmatrix}1&2\\0&0\end{bmatrix}, \begin{bmatrix}0&0\\2&0\end{bmatrix}\right\}$.

(b) $M_{2\times 2}(\mathbb{F}) = \text{Span}\left\{\begin{bmatrix}0&1\\0&0\end{bmatrix}, \begin{bmatrix}0&0\\0&1\end{bmatrix}, \begin{bmatrix}1&2\\0&0\end{bmatrix}, \begin{bmatrix}0&0\\2&0\end{bmatrix}\right\}$.

Corollary 16: (Theorem 11) Let V be a vector space. Let β be a finite set that is also a basis of V. Let β' be another basis of V. Then β and β' have the same number of elements, namely $|\beta| = |\beta'|$.

Proof: Since β is a basis then $V = \text{Span}(\beta)$, since β' is a basis then β' is linearly independent, so by Theorem 15 then $|\beta'| \leq |\beta|$. Exchanging the roles of β and β' gives $|\beta| \leq |\beta'|$. Thus $|\beta| = |\beta'|$. □.

Remark: Given a vector space V of dimension $n \in \mathbb{N}$, then:

1. A basis β of V has exactly n elements.

2. A subset of V spanning V has at least n elements.

3. A linearly independent subset of V has at most n elements.

In particular any spanning set of V with exactly n vectors is a basis of V.

Examples:

1. $\dim_{\mathbb{R}}(\mathbb{R}^n) = n$ with basis $\beta = \{e_i = (0,\dots,\overset{i}{1},\dots,0) \mid 1 \leq i \leq n\}$.

2.1. $\dim_{\mathbb{C}}(\mathbb{C}^n) = n$ with basis $\beta = \{ e_i = (0,\dots,\overset{i}{1},\dots,0) \mid 1 \le i \le n \}$.

2.2. $\dim_{\mathbb{R}}(\mathbb{C}^n) = 2n$ with basis

$$\beta = \{ e_i = (0,\dots,\overset{i}{1},\dots,0) \mid 1 \le i \le n \} \cup \{ f_j = (0,\dots,\overset{j}{i},\dots,0) \mid 1 \le j \le n \}.$$

3. $\dim_{\mathbb{F}}(M_{n\times m}(\mathbb{F})) = n \cdot m$ with basis $\beta = \{ E_{ij} \mid 1 \le i \le n,\ 1 \le j \le m \}$.

4. $\dim_{\mathbb{F}}(\mathbb{F}[x]) = |\mathbb{N}|$ with basis $\beta = \{ x^i \mid i \in \mathbb{N} \}$.

5. $\dim_{\mathbb{F}}(\mathbb{F}_n[x]) = n+1$ with basis $\beta = \{ x^i \mid 0 \le i \le n \}$.

A vector subspace W of a vector space V is a vector space on its own, so it will have a basis and a dimension. However, our options will be restricted by what is a basis in V, and by the dimension of V.

Theorem 17: Let V be a finite dimensional vector space, $W \subseteq V$ a vector subspace. Then:

(1) $\dim_{\mathbb{F}}(W) \le \dim_{\mathbb{F}}(V)$,

(2) $\dim_{\mathbb{F}}(W) = \dim_{\mathbb{F}}(V)$ if and only if $V = W$.

In particular, W is finite dimensional.

Proof: (1) Let β be a basis of W. Since β is linearly independent in W it is also linearly independent in V. Now V can be generated with $\dim_{\mathbb{F}}(V)$ vectors, so by Theorem 15 then $\dim_{\mathbb{F}}(W) = |\beta| \le \dim_{\mathbb{F}}(V)$.

(2) We want to prove that $\dim_{\mathbb{F}}(W) = \dim_{\mathbb{F}}(V)$ if and only if $V = W$.

($\Leftarrow$) Suppose $V = W$, then $\dim_{\mathbb{F}}(V) = \dim_{\mathbb{F}}(W)$.

($\Rightarrow$) Suppose $\dim_{\mathbb{F}}(W) = \dim_{\mathbb{F}}(V)$, let β be a basis of W, let γ be a basis of V. By Theorem 15 there is a subset S of γ containing $\dim_{\mathbb{F}}(V) - \dim_{\mathbb{F}}(W)$ vectors such that $V = \text{Span}(\beta \cup S)$. Since $\dim_{\mathbb{F}}(W) = \dim_{\mathbb{F}}(V)$ then S contains zero vectors, so $S = \emptyset$ and $\beta \cup S = \beta$. Hence $V = \text{Span}(\beta) = W$. $\square$.

Thus finite dimensional vector spaces can only have finite dimensional subspaces.

Inspecting the proof of Theorem 17, we find a corollary of the Replacement Theorem.

Corollary 18: Let V be a finite dimensional vector space, $W \subsetneq V$ a vector subspace with basis β. Then we can extend β to a basis of V.

Proof: Let γ be a basis of V. By Theorem 15 there is a subset S of γ containing $\dim_{\mathbb{F}}(V) - \dim_{\mathbb{F}}(W)$ vectors such that $V = \text{Span}(\beta \cup S)$. Note that $S \neq \emptyset$ since $W \neq V$. Now $\beta \cup S$ is a set with at most $\dim_{\mathbb{F}}(V)$ elements that generates V. By the Remark after Corollary 16 then $\beta \cup S$ has exactly $\dim_{\mathbb{F}}(V)$ elements, and thus is a basis of V. $\square$.

Remark: If W is a vector subspace of V with basis $\{w_1, \dots, w_k\}$, we can add vectors $v_{k+1}, \dots, v_n$ to it so that $\{w_1, \dots, w_k, v_{k+1}, \dots, v_n\}$ is a basis of V.

Now setting $W^c = \text{Span}\{v_{k+1}, \dots, v_n\}$ the complement subspace of W we find $V = W \oplus W^c$.

2. Linear transformations.

We have given the structure of a vector space (over a field) to some sets, we now investigate how different objects with this structure interact. Namely, we are interested in relations between vector spaces, and in how to obtain information about one if we know information about another. We will achieve this by using functions that preserve this vector space structure, the so called linear transformations.

Definition: Let V and W be vector spaces over the same field $\mathbb{F}$. A linear transformation is a function $T: V \to W$ satisfying for all $x, y \in V$ and $a \in \mathbb{F}$:

(1) $T(x+y) = T(x) + T(y)$

(2) $T(a \cdot x) = a \cdot T(x)$.

Remark: In particular, a linear transformation $T: V \to W$ preserves linear combinations, namely $T\left(\sum_{i=1}^{n} a_i \cdot v_i\right) = \sum_{i=1}^{n} a_i \cdot T(v_i)$ for all $v_1, \dots, v_n \in V$ and all

$a_1, \ldots, a_n \in \mathbb{F}$. A linear transformation also sends the zero in V to the zero in W: $T(\vec{0}) = \vec{0}$.

Theorem 19: Let V and W be vector spaces over the same field $\mathbb{F}$.

1) Let $T_1 : V \to W$ and $T_2 : V \to W$ be linear transformations, then:

$$T_1 + T_2 : V \longrightarrow W \qquad \text{is a linear transformation.}$$
$$x \mapsto T_1(x) + T_2(x)$$

2) Let $T : V \to W$ be a linear transformation and $c \in \mathbb{F}$, then:

$$c \cdot T : V \longrightarrow W \qquad \text{is a linear transformation.}$$
$$x \longmapsto c \cdot T(x)$$

3) The set $\mathcal{L}(V,W) = \{T : V \to W \mid T \text{ is a linear transformation}\}$ is a vector space over $\mathbb{F}$ with the operations above.

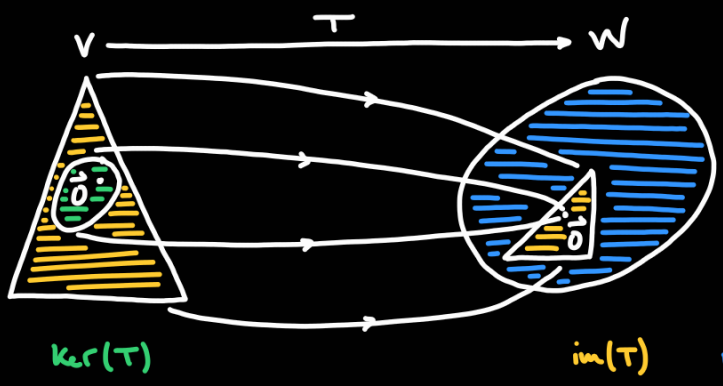
Proof: Straightforward, use the definition for 1) and 2), use Theorem 4 for 3). □.

Definition: Let V and W be vector spaces over the same field $\mathbb{F}$, let $T : V \to W$ be a linear transformation. The kernel or null space of T, denoted $\ker(T)$, is the set:

$$\ker(T) = \{x \in V \mid T(x) = \vec{0}\}.$$

The image or range of T, denoted $\text{im}(T)$, is the set:

$$\text{im}(T) = \{y \in W \mid \text{there exists } x \in V \text{ with } T(x) = y\} = \{T(x) \in W \mid x \in V\}.$$



Examples:

1. A function T between a set S and a field $\mathbb{F}$, namely $T \in \mathcal{F}(S, \mathbb{F})$, that is a polynomial of degree less than or equal to 1 is a linear transformation on every vector space structure on S.

2. The function $T: M_{n\times m}(\mathbb{F}) \longrightarrow M_{n\times m}(\mathbb{F})$ is a linear transformation since
$$A \longmapsto A^t$$
$T(A+B) = (A+B)^t = A^t + B^t = T(A) + T(B)$ and $T(c\cdot A) = (c\cdot A)^t = c\cdot A^t = c\cdot T(A)$.

3. The function $T: \mathbb{F}_n[x] \longrightarrow \mathbb{F}_{n-1}[x]$ is a linear transformation since
$$f \longmapsto \frac{d}{dx} f$$
$T(f+g) = \frac{d}{dx}(f+g) = \frac{d}{dx}f + \frac{d}{dx}g = T(f) + T(g)$ and

$T(c\cdot f) = \frac{d}{dx}(c\cdot f) = c\cdot \frac{d}{dx} f = c\cdot T(f)$.

4. The function $T: \mathcal{C}(\mathbb{R}, \mathbb{R}) \longrightarrow \mathbb{R}$ is a linear transformation.
$$f \longmapsto \int_a^b f\, dx$$

5. The functions from $\mathbb{R}$ to $\mathbb{R}$ given by $T(x) = x^2$, $T(x) = \frac{1}{x}$, $T(x) = e^x$, $T(x) = \sin(x)$, $T(x) = \cos(x)$, and combinations of these, are not linear transformations.

Example: Compute the kernel and image of the linear transformation $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^2$.

$(x,y,z) \mapsto (x-y, 2z)$

For the kernel, we want $\ker(T) = \{(x,y,z) \mid T(x,y,z) = (0,0)\}$ so:

$(x,y,z) \in \ker(T)$ if and only if $T(x,y,z) = (0,0)$, namely $(x-y, 2z) = (0,0)$

so $x-y = 0$ and $2z = 0$ so $x = y$ and $z = 0$.

Thus (x,y,z) should look like $(x,x,0)$, and $\ker(T) = \{(x,x,0) \mid x \in \mathbb{R}\}$.

For the image, we want $\text{im}(T) = \{T(x,y,z) \mid (x,y,z) \in \mathbb{R}^3\}$ so:

$(u,v) \in \text{im}(T)$ if and only if $(u,v) = T(x,y,z) = (x-y, 2z)$,

so $u = x-y$ and $v = 2z$.

Now since any real number u can be obtained as a difference of two real numbers $x-y$, and any real number v can be obtained by doubling another real number z, then u and v take all the possible real values. Hence $\text{im}(T) = \mathbb{R}^2$.

Theorem 20: Let V and W be vector spaces over the same field $\mathbb{F}$, let $T: V \longrightarrow W$ be a linear transformation. Then:

1) $\ker(T)$ is a subspace of V.

2) $\text{im}(T)$ is a subspace of W.

Proof: We use Theorem 4:

1) Since $T(\vec{0}) = \vec{0}$ then $\vec{0} \in \ker(T)$.

Suppose that $x, y \in \ker(T)$, namely $T(x) = \vec{0} = T(y)$. Then:

$T(x+y) = T(x) + T(y) = \vec{0} + \vec{0} = \vec{0}$ so $x+y \in \ker(T)$.

Suppose that $x \in \ker(T)$, so $T(x) = \vec{0}$, and $c \in \mathbb{F}$. Then:

$T(c \cdot x) = c \cdot T(x) = c \cdot \vec{0} = \vec{0}$ so $c \cdot x \in \ker(T)$.

2) Since $T(\vec{0}) = \vec{0}$ then $\vec{0} \in \text{im}(T)$.

Suppose that $u, v \in \text{im}(T)$, so there are $x, y \in V$ with $T(x) = u$ and $T(y) = v$.

Then $T(x+y) = T(x) + T(y) = u + v$ so $u+v \in \text{im}(T)$.

Suppose that $u \in \text{im}(T)$, so there is $x \in V$ with $T(x) = u$, and $c \in \mathbb{F}$. Then:

$T(c \cdot x) = c \cdot T(x) = c \cdot u$ so $c \cdot u \in \text{im}(T)$. $\square$.

Theorem 21: Let V and W be vector spaces over the same field $\mathbb{F}$, let $T: V \longrightarrow W$ be a linear transformation, let $\beta = \{v_1, \dots, v_n\}$ be a basis of V. Then

$T(\beta) = \{T(v_1), \dots, T(v_n)\}$ spans $\text{im}(T)$, namely $\text{im}(T) = \text{Span}\{T(v_1), \dots, T(v_n)\}$.

Proof: We prove the double inclusion of $\text{Span}\{T(v_1), \dots, T(v_n)\}$ in $\text{im}(T)$ and of $\text{im}(T)$ in $\text{Span}\{T(v_1), \dots, T(v_n)\}$.

2) Since $T(v_1), \dots, T(v_n) \in \text{im}(T)$ then $\text{Span}\{T(v_1), \dots, T(v_n)\} \subseteq \text{im}(T)$.

$\subseteq$) Let $u \in \text{im}(T)$, then there exists $v \in V$ with $T(v) = u$. Since β is a basis of V, we can write $v = \sum_{i=1}^{n} a_i \cdot v_i$ for some coefficients $a_1, \dots, a_n \in \mathbb{F}$, and thus:

$$u = T(v) = T\left(\sum_{i=1}^{n} a_i \cdot v_i\right) = \sum_{i=1}^{n} T(a_i \cdot v_i) = \sum_{i=1}^{n} a_i \cdot T(v_i).$$

Hence $u \in \text{Span}\{T(v_1), \dots, T(v_n)\}$. Since this holds for all $u \in \text{im}(T)$, we have that $\text{im}(T) \subseteq \text{Span}\{T(v_1), \dots, T(v_n)\}$. □.

<u>Theorem 22:</u> (Rank-Nullity Theorem) Let V and W be vector spaces over the same field $\mathbb{F}$, with V finite dimensional, and let $T: V \longrightarrow W$ be a linear transformation. Then:

$$\dim(V) = \dim(\ker(T)) + \dim(\text{im}(T)).$$

<u>Proof:</u> Since V is finite dimensional, set $n = \dim(V)$. Writing $k = \dim(\ker(T))$, since $\ker(T)$ is a subspace of V, then $k \leq n$. Given a basis $\{v_1, \dots, v_k\}$ of $\ker(T)$, by Corollary 18 we can extend it to a basis $\{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$ of V. As we remarked this allows us to write $V = \ker(T) \oplus \ker(T)^c$ where $\ker(T)^c$ has basis $\{v_{k+1}, \dots, v_n\}$.

We now claim that $\{T(v_{k+1}), \dots, T(v_n)\}$ is a basis of $\text{im}(T)$. To prove this, we have to prove that $\{T(v_{k+1}), \dots, T(v_n)\}$ spans $\text{im}(T)$ and is linearly independent.

To prove spanning, note that since $\{v_1, \dots, v_n\}$ is a basis of V then

$\{T(v_1), \ldots, T(v_n)\}$ spans $\text{im}(T)$ by Theorem 21. Now $v_1, \ldots, v_k \in \ker(T)$ so $T(v_1) = \cdots = T(v_k) = \vec{0}$, which do not contribute to the span. Thus $\{T(v_1), \ldots, T(v_n)\} \in \text{Span}\{T(v_{k+1}), \ldots, T(v_n)\}$ so $\text{im}(T) = \text{Span}\{T(v_1), \ldots, T(v_n)\} \subseteq \text{Span}\{T(v_{k+1}), \ldots, T(v_n)\}$. Since $\text{Span}\{T(v_{k+1}), \ldots, T(v_n)\} \subseteq \text{im}(T)$ then $\text{im}(T) = \text{Span}\{T(v_{k+1}), \ldots, T(v_n)\}$.

To prove linear independence, we proceed by contradiction. Suppose that the set is linear dependent, namely there are scalars $a_{k+1}, \ldots, a_n \in \mathbb{F}$, at least one non-zero, such that $a_{k+1} \cdot T(v_{k+1}) + \cdots + a_n T(v_n) = \vec{0}$. Then $T(a_{k+1} \cdot v_{k+1} + \cdots + a_n v_n) = \vec{0}$ so $a_{k+1} \cdot v_{k+1} + \cdots + a_n v_n \in \ker(T)$. Since $\{v_1, \ldots, v_k\}$ is a basis of $\ker(T)$ then $a_{k+1} \cdot v_{k+1} + \cdots + a_n v_n \in \text{Span}\{v_1, \ldots, v_k\}$, so there are scalars $a_1, \ldots, a_k \in \mathbb{F}$ such that $a_{k+1} \cdot v_{k+1} + \cdots + a_n v_n = a_1 \cdot v_1 + \cdots + a_k \cdot v_k$, which we can rewrite as $-a_1 \cdot v_1 - \cdots - a_k \cdot v_k + a_{k+1} + \cdots + a_n v_n = \vec{0}$. This is a linear combination of the basis $\{v_1, \ldots, v_n\}$ of V that is zero, but at least one of the $a_{k+1}, \ldots, a_n \in \mathbb{F}$ is not zero, contradicting the linear independency of the basis. Thus our assumption that $\{T(v_{k+1}), \ldots, T(v_n)\}$ is linear dependent is false, hence it is linearly

independent.

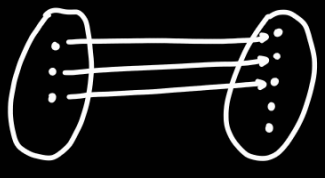
Hence $\{T(\vec{v}_{k+1}), \ldots, T(\vec{v}_n)\}$ is a basis of $\text{im}(T)$, meaning that $\dim(\text{im}(T)) = n-k$.

Now $\dim(\ker(T)) + \dim(\text{im}(T)) = k + (n-k) = n = \dim(V)$. □.

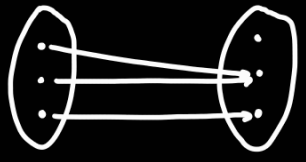
Remark: We call $\dim(\ker(T))$ the nullity of T and $\dim(\text{im}(T))$ the rank of T.

Definition: Let V and W be vector spaces over the same field $\mathbb{F}$, let $T: V \to W$ be a linear transformation. We say that T is injective or one-to-one if $T(x) = T(y)$ implies $x = y$. We say that T is surjective or onto if for every $y \in W$ there is $x \in V$ with $T(x) = y$.

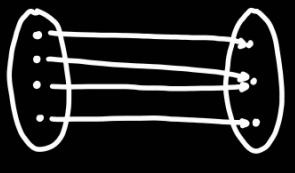
We often write $T: V \hookrightarrow W$ when T is injective, and $T: V \twoheadrightarrow W$ when T is surjective. Pictorially, we have:



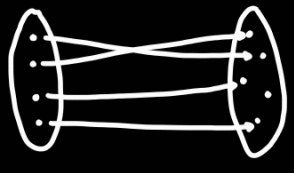
injective, but



not injective, and



surjective, but



not surjective.

Theorem 23: Let V and W be vector spaces over the same field $\mathbb{F}$, let $T: V \to W$ be a linear transformation.

(1) T is injective if and only if $\ker(T) = \{\vec{0}\}$.

(2) T is surjective if and only if $\text{im}(T) = W$.

Proof: (1) ($\Rightarrow$) Suppose T is injective. We prove $\ker(T) = \{\vec{0}\}$. Let $x \in \ker(T)$, so $T(x) = \vec{0}$.

Now $T(x) = \vec{0} = T(\vec{0})$ so by injectivity of T we have $x = \vec{0}$, the sole element in $\ker(T)$.

($\Leftarrow$) Suppose $\ker(T) = \{\vec{0}\}$. We prove T injective. Let $x, y \in V$ such that $T(x) = T(y)$, then $\vec{0} = T(x) - T(y) = T(x-y)$ so $x-y \in \ker(T)$. Hence $x-y = \vec{0}$ so $x = y$, and T is injective.

(2) ($\Rightarrow$) Suppose T is surjective. We prove $\text{im}(T) = W$. Since $\text{im}(T) \subseteq W$, it suffices to prove $W \subseteq \text{im}(T)$. Let $y \in W$, since T is surjective there is $x \in V$ with $T(x) = y$, and thus $y \in \text{im}(T)$.

($\Leftarrow$) Suppose $W = \text{im}(T)$. We prove T surjective. Let $y \in W$, since $W = \text{im}(T)$ then $y \in \text{im}(T)$, so there is $x \in V$ with $T(x) = y$. Thus T is surjective. $\square$.

The concepts of injectivity and surjectivity are different in general, but sometimes they coincide.

Theorem 24: Let V and W be vector spaces over the same field $\mathbb{F}$, let $T: V \longrightarrow W$ be a linear transformation, let $\dim(V) = \dim(W)$ be a finite natural number. The

following are equivalent:

(1) T is injective.

(2) T is surjective.

(3) The dimension of $\text{im}(T)$ equals the dimension of V.

Proof: We have to prove that (1) holds if and only if (2) holds if and only if (3) holds. It is enough to prove the equivalence between (1) and (3), and the equivalence between (2) and (3).

(1) $\Rightarrow$ (3) Suppose T is injective. Using Theorem 23 we know that $\ker(T) = \{\vec{0}\}$. Since $\dim(\ker(T)) = \dim(\{\vec{0}\}) = 0$ by Theorem 22 then $\dim(V) = \dim(\text{im}(T))$.

(3) $\Rightarrow$ (1) Suppose $\dim(V) = \dim(\text{im}(T))$. Since by Theorem 22 then $\dim(V) = \dim(\text{im}(T)) + \dim(\ker(T))$, we have $\dim(\ker(T)) = 0$ and thus $\ker(T) = \text{Span}\{\} = \{\vec{0}\}$. Now by Theorem 23 we have T injective.

(2) $\Rightarrow$ (3) Suppose T is surjective. Using Theorem 23 we know that $\text{im}(T) = W$, so $\dim(V) = \dim(W) = \dim(\text{im}(T))$.

(3) $\Rightarrow$ (2) Suppose $\dim(V) = \dim(\text{im}(T))$ of T [illegible] of

(3) $\Rightarrow$ (2) Suppose $\dim(V) = \dim(\text{im}(T))$. Now $\text{im}(T)$ is a vector subspace of W by Theorem 20, and since $\dim(W) = \dim(V) = \dim(\text{im}(T))$, by Theorem 17 then $W = \text{im}(T)$. Now by Theorem 23 then T is surjective. $\square$.

Given a linear transformation $T: V \to W$, often it is more convenient to compute the dimensions of $\ker(T)$ and $\text{im}(T)$ than to compute the dimensions of V and W. Using Theorem 22 and Theorem 24, we can often determine what we want from just knowing $\ker(T)$ and $\text{im}(T)$.

We will now see that a linear map $T: V \to W$ is completely determined by where it sends a basis of V. This means that to check a characteristic of a linear transformation, such as checking whether T is linear or compute its kernel and image, it suffices to work with a basis of V.

Theorem 25: Let V and W be vector spaces over the same field $\mathbb{F}$, let $\beta = \{v_1, \ldots, v_n\}$ be a basis of V, let $\gamma = \{w_1, \ldots, w_n\}$ be a basis of W. Then the function defined by $T: V \longrightarrow W$ is a linear transformation, and it is unique.
$$\sum_{i=1}^{n} a_i \cdot v_i \mapsto \sum_{i=1}^{n} a_i \cdot w_i$$

Proof: We first prove that $T: V \to W$ is linear. Let $x, y \in V$, since β is a basis of V we can write $x = \sum_{i=1}^{n} a_i \cdot v_i$ and $y = \sum_{i=1}^{n} b_i \cdot v_i$ for some scalars

$a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{F}$. Now:

$$T(x+y) = T\left(\sum_{i=1}^{n} a_i \cdot v_i + \sum_{i=1}^{n} b_i \cdot v_i\right) = T\left(\sum_{i=1}^{n} (a_i + b_i) \cdot v_i\right) = \sum_{i=1}^{n} (a_i + b_i) \cdot w_i =$$

$$= \sum_{i=1}^{n} a_i \cdot w_i + \sum_{i=1}^{n} b_i \cdot w_i = T(x) + T(y).$$

Similarly if $a \in \mathbb{F}$ then:

$$T(a \cdot x) = T\left(a \cdot \sum_{i=1}^{n} a_i \cdot v_i\right) = T\left(\sum_{i=1}^{n} (a \cdot a_i) \cdot v_i\right) = \sum_{i=1}^{n} (a \cdot a_i) \cdot w_i =$$

$$= a \cdot \sum_{i=1}^{n} a_i \cdot w_i = a \cdot T(x).$$

Hence $T: V \to W$ is linear. Suppose that there is another linear transformation $T^1: V \to W$ such that $T^1(v_i) = w_i$ for all $i = 1, \dots, n$. Since any $v \in V$ can be written as $v = \sum_{i=1}^{n} a_i \cdot v_i$ for some $a_1, \dots, a_n \in \mathbb{F}$, then:

$$T(v) = T\left(\sum_{i=1}^{n} a_i \cdot v_i\right) = \sum_{i=1}^{n} a_i \cdot T(v_i) = \sum_{i=1}^{n} a_i \cdot T^1(v_i) = T^1\left(\sum_{i=1}^{n} a_i \cdot v_i\right) = T^1(v).$$

Hence $T(v) = T^1(v)$ for all $v \in V$, so $T = T^1$ and T is unique. □.

Note that we only used that $T: V \to W$ is defined on a basis of V, the values $T(v_1), \dots, T(v_n)$ do not matter. Namely, a linear transformation $T: V \to W$ is uniquely determined by its image on a basis of V.

Example: Consider the function $T: \mathbb{R}_2[x] \longrightarrow \mathbb{R}_3[x]$. Is it linear? Is it

$$f(x) \longmapsto 2 \cdot f'(x) + \int_0^x 3 \cdot f(t)\, dt$$

injective? Is it surjective?

Note that differentiation and integration are linear transformations, and T is obtained by multiplying linear transformations by scalars, and then adding up the remaining linear transformations. Hence T is indeed a linear transformation.

We now compute the dimensions of the image and kernel of T, using that $\{1, x, x^2\}$ is a basis of $\mathbb{R}_2[x]$. We obtain:

$$T(1) = 3x, \quad T(x) = 2 + \frac{3}{2}x^2, \quad T(x^2) = 4x + x^3,$$

so $\text{im}(T) = \text{Span}\{3x, 2+\frac{3}{2}x^2, 4x+x^3\}$. Since $3x, 2+\frac{3}{2}x^2, 4x+x^3$ all have different degrees, they are linearly independent and thus $\dim(\text{im}(T)) = 3$.

Since $\mathbb{R}_3[x]$ has basis $\{1, x, x^2, x^3\}$, then $\dim(\mathbb{R}_3[x]) = 4$ and $\text{im}(T) \subsetneq \mathbb{R}_3[x]$, so T is not surjective. Moreover:

$$3 = \dim(\mathbb{R}_2[x]) = \dim(\text{im}(T)) + \dim(\ker(T)) = 3 + \dim(\ker(T))$$

and thus $\dim(\ker(T)) = 0$ so $\ker(T) = \{\vec{0}\}$ so T is injective.

We are now in a position to justify why "linear transformations are equivalent to matrices". First, we will introduce the notions of coordinate vector. Second, we will use this to define the matrix associated to a linear transformation. Third, we will explain how these constructions enable us to understand a finite dimensional vector

space V, say of dimension n, in terms of $\mathbb{F}^n$. In particular, a linear transformation from V to W will be identified with a linear transformation from $\mathbb{F}^n$ to $\mathbb{F}^m$, where m is the dimension of W, which in turn corresponds to an $m \times n$ matrix.

Definition: Let V be a finite dimensional vector space with basis $\beta = \{v_1, \dots, v_n\}$, let $v \in V$ be written as $v = \sum_{i=1}^{n} a_i \cdot v_i$ for some $a_1, \dots, a_n \in \mathbb{F}$. We say that the coordinate vector of v with respect to the basis β is:

$$[v]_\beta = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

Example: Let $V = \mathbb{R}_2[x]$, it has basis $\beta = \{1, x, x^2\}$ and $\gamma = \{1+x, 1-x, 3x^2\}$.

Let $f(x) = 3 - 2x + 4x^2$. We can also write $f(x) = \frac{1}{2}(1+x) + \frac{5}{2}(1-x) + \frac{4}{3} \cdot 3x^2$. Hence:

$$[f(x)]_\beta = \begin{bmatrix} 3 \\ -2 \\ 4 \end{bmatrix} \quad \text{and} \quad [f(x)]_\gamma = \begin{bmatrix} 1/2 \\ 5/2 \\ 4/3 \end{bmatrix}.$$

Definition: Let V and W be vector spaces over the same field $\mathbb{F}$, let $\beta = \{v_1, \dots, v_n\}$ be a basis of V, let $\gamma = \{w_1, \dots, w_m\}$ be a basis of W, let $T: V \to W$ be a linear transformation. Write $T(v_j) = \sum_{i=1}^{m} a_{ij} \cdot w_i$ with $a_{ij} \in \mathbb{F}$ for each $i = 1, \dots, m$ and $j = 1, \dots, n$. The matrix associated to T with respect to the basis β and γ is:

$$[T]_\beta^\gamma = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \Big[[T(v_1)]_\gamma \cdots [T(v_n)]_\gamma \Big].$$

We should think of $[v]_\beta$ and $[T]_\beta^\gamma$ as just notation for the linear combinations $v = \sum_{i=1}^{n} a_i \cdot v_i$ and $T(v_j) = \sum_{i=1}^{m} a_{ij} \cdot w_i$ for $j = 1, \ldots, n$. These vectors and matrices encode all the information about v and T. Since we understand matrices well, it will be useful to have access to a matrix containing all the information required to work with abstract v and T.

Theorem 26: Let V and W be vector spaces over the same field $\mathbb{F}$, let $\beta = \{v_1, \ldots, v_n\}$ be a basis of V, let $\gamma = \{w_1, \ldots, w_m\}$ be a basis of W. Let $T: V \longrightarrow W$ and $T': V \longrightarrow W$ be linear transformations, let $c \in \mathbb{F}$. Then:

1) $[T+T']_\beta^\gamma = [T]_\beta^\gamma + [T']_\beta^\gamma$ and

2) $[c \cdot T]_\beta^\gamma = c \cdot [T]_\beta^\gamma$.

Proof: We first establish the notation we will use. Suppose:

$$T(v_j) = \sum_{i=1}^{m} a_{ij} \cdot w_i, \quad T'(v_j) = \sum_{i=1}^{m} b_{ij} \cdot w_i, \quad (T+T')(v_j) = \sum_{i=1}^{m} c_{ij} \cdot w_i,$$

and $(c \cdot T)(v_j) = \sum_{i=1}^{m} d_{ij} \cdot w_i$, now:

$$(T+T')(v_j) = T(v_j) + T'(v_j) = \sum_{i=1}^{m} a_{ij} \cdot w_i + \sum_{i=1}^{m} b_{ij} \cdot w_i = \sum_{i=1}^{m} (a_{ij} + b_{ij}) \cdot w_i$$

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

and thus $c_{ij} = a_{ij} + b_{ij}$ and $d_{ij} = c\cdot a_{ij}$ by Theorem 12. Then:

1) $([T+T']_\beta^\gamma)_{ij} = c_{ij} = a_{ij} + b_{ij} = ([T]_\beta^\gamma)_{ij} + ([T']_\beta^\gamma)_{ij}$ for all $i=1,\dots,m$

and $j=1,\dots,n$. Thus $[T+T']_\beta^\gamma = [T]_\beta^\gamma + [T']_\beta^\gamma$.

2) $([c\cdot T]_\beta^\gamma)_{ij} = d_{ij} = c\cdot a_{ij} = c\cdot([T]_\beta^\gamma)_{ij}$ for all $i=1,\dots,m$ and $j=1,\dots,n$.

Thus $[c\cdot T]_\beta^\gamma = c\cdot[T]_\beta^\gamma$. □.

Theorem 27: Let V, W, X be vector spaces over the same field F, let $\alpha = \{v_1,\dots,v_n\}$, $\beta = \{w_1,\dots,w_m\}$, $\gamma = \{x_1,\dots,x_p\}$ be basis of V, W, X. Let $T: V \to W$ and $T': W \to X$ be linear transformations. Then $T'\circ T: V \to X$ is a linear transformation and $[T'\circ T]_\alpha^\gamma = [T']_\beta^\gamma \cdot [T]_\alpha^\beta$.

Proof: Let $c \in F$, $x, y \in V$, then:

$$(T'\circ T)(x+y) = T'(T(x+y)) = T'(T(x)+T(y)) = T'(T(x)) + T'(T(x)) =$$

$$(T'\circ T)(x) + (T'\circ T)(y),$$

$$(T'\circ T)(c\cdot x) = T'(T(c\cdot x)) = T'(c\cdot T(x)) = c\cdot T'(T(x)) = c\cdot(T'\circ T)(x).$$

Hence $T'\circ T$ is indeed linear. We now establish the notation we will use. Suppose:

$T(v_j) = \sum_{k=1}^{m} b_{kj}\cdot w_k$ and $T'(w_k) = \sum_{i=1}^{p} a_{ik}\cdot x_i$, now:

$$(T' \circ T)(v_j) = T'\left(T(v_j)\right) = T'\left(\sum_{k=1}^{m} b_{kj} \cdot w_k\right) = \sum_{k=1}^{m} b_{kj} \cdot T'(w_k) =$$

$$= \sum_{k=1}^{m} b_{kj} \cdot \left(\sum_{i=1}^{p} a_{ik} \cdot x_i\right) = \sum_{i=1}^{p} \left(\sum_{k=1}^{m} a_{ik} \cdot b_{kj}\right) \cdot x_i .$$

Thus $\left([T' \circ T]_\alpha^\gamma\right)_{ij} = \sum_{k=1}^{m} a_{ik} \cdot b_{kj} = \sum_{k=1}^{m} \left([T']_\beta^\gamma\right)_{ik} \cdot \left([T]_\alpha^\beta\right)_{kj} = \left([T']_\beta^\gamma \cdot [T]_\alpha^\beta\right)_{ij}$

so $[T' \circ T]_\alpha^\gamma = [T']_\beta^\gamma \cdot [T]_\alpha^\beta$. □.

Theorem 28: Let V and W be vector spaces over the same field $\mathbb{F}$, let $\beta = \{v_1, \ldots, v_n\}$ be a basis of V, let $\gamma = \{w_1, \ldots, w_m\}$ be a basis of W. Let $T: V \longrightarrow W$ be a linear transformation. Then $[T(v)]_\gamma = [T]_\beta^\gamma [v]_\beta$ for all $v \in V$.

Proof: We first establish the notation we will use. Suppose:

$$T(v_j) = \sum_{i=1}^{m} b_{ij} \cdot w_i \quad \text{and} \quad v = \sum_{j=1}^{n} a_j \cdot v_j , \quad \text{now:}$$

$$T(v) = T\left(\sum_{j=1}^{n} a_j \cdot v_j\right) = \sum_{j=1}^{n} a_j \cdot T(v_j) = \sum_{j=1}^{n} a_j \cdot \left(\sum_{i=1}^{m} b_{ij} \cdot w_i\right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} b_{ij} \cdot a_j\right) \cdot w_i .$$

Hence $\left([T(v)]_\gamma\right)_i = \sum_{j=1}^{n} b_{ij} \cdot a_j = \sum_{j=1}^{n} \left([T]_\beta^\gamma\right)_{ij} \cdot \left([v]_\beta\right)_j = \left([T]_\beta^\gamma \cdot [v]_\beta\right)_i$, and thus

$[T(v)]_\gamma = [T]_\beta^\gamma \cdot [v]_\beta$. □.

Example: Let $T: \mathbb{R}_2[x] \longrightarrow \mathbb{R}_3[x]$, $f(x) \longmapsto \int_0^x f(t)\,dt$, consider $\beta = \{1, x, x^2\}$ a basis of $\mathbb{R}_2[x]$ and $\gamma = \{1, x, x^2, x^3\}$ a basis of $\mathbb{R}_3[x]$. Now:

$$T(1) = x , \quad T(x) = \frac{x^2}{2} , \quad T(x^2) = \frac{x^3}{3} .$$

Hence:

$$[T]_\beta^\gamma = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{bmatrix}. \quad \text{Not that } [1]_\beta = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad [T(1)]_\gamma = [x]_\gamma = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \text{ and}$$

$$[T]_\beta^\gamma \cdot [1]_\beta = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = [T(1)]_\gamma.$$

Consider $\tilde{T}: \mathbb{R}_3[x] \longrightarrow \mathbb{R}_2[x]$, we have:

$$f(x) \longmapsto f'(x)$$

$$\tilde{T}(1) = 0, \quad \tilde{T}(x) = 1, \quad \tilde{T}(x^2) = 2x, \quad \tilde{T}(x^3) = 3x^2.$$

Hence:

$$[\tilde{T}]_\gamma^\beta = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Moreover:

$$\tilde{T} \circ T(1) = 1, \quad \tilde{T} \circ T(x) = x, \quad \tilde{T} \circ T(x^2) = x^2, \text{ so } [\tilde{T} \circ T]_\beta^\beta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Now:

$$[\tilde{T}]_\gamma^\beta [T]_\beta^\gamma = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [T' \circ T]_\beta^\beta.$$

Our approach to the statement "linear transformations are equivalent to matrices" requires us to make this statement formal for $\mathbb{F}^n$ and $\mathbb{F}^m$. The following proves that the function $M_{m \times n}(\mathbb{F}) \longrightarrow \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$, $A \longmapsto T_A$ is a linear transformation, injective, and surjective.

Definition: Let $A \in M_{m \times n}(\mathbb{F})$, we say that the linear transformation $T_A: \mathbb{F}^n \longrightarrow \mathbb{F}^m$, $x \longmapsto Ax$

is the left-multiplication by A.

Theorem 29: Let $A \in M_{m\times n}(\mathbb{F})$, then $T_A : \mathbb{F}^n \longrightarrow \mathbb{F}^m$ is a linear transformation and:

1) $[T_A]_{\sigma_n}^{\sigma_m} = A$, where σ_n and σ_m are the standard basis of $\mathbb{F}^n$ and $\mathbb{F}^m$.

2) Let $B \in M_{m\times n}(\mathbb{F})$, then $T_A = T_B$ if and only if $A = B$.

3) Let $B \in M_{m\times n}(\mathbb{F})$, then $T_{A+B} = T_A + T_B$.

4) Let $c \in \mathbb{F}$, then $T_{c\cdot A} = c \cdot T_A$.

5) Let $B \in M_{n\times p}(\mathbb{F})$, then $T_{A\cdot B} = T_A \circ T_B$.

6) Let $I_{dn} \in M_{n\times n}(\mathbb{F})$ be the identity matrix, then $T_{I_{dn}} = id_{\mathbb{F}^n}$, where

$id_{\mathbb{F}^n}$ is the identity function on $\mathbb{F}^n$ (namely $id_{\mathbb{F}^n}(x) = x$ for all $x \in \mathbb{F}^n$).

Proof: It follows from the definitions of matrix operations, but only basic facts are used, and a straightforward computation suffices. □.

Remark: Let V be a vector space, the identity function on V is the linear transformation that sends every element in V to itself: $id_V : V \longrightarrow V$.

$v \longmapsto v$