Chapter 2
Vector Spaces

The notion of vector space can be defined over any field $\mathbb{K}$. We shall mainly consider the case $\mathbb{K} = \mathbb{R}$ and briefly mention the case $\mathbb{K} = \mathbb{C}$. Starting from our exposition, it is straightforward to generalise to any field.

2.1 Definition and Basic Properties

The model of the construction is the collection of all vectors in the space applied at a point with the operations of sum and multiplication by a scalar, as described in the Chap. 1.

Definition 2.1.1 A non empty set V is called a *vector space over* $\mathbb{R}$ (or a *real vector space* or an $\mathbb{R}$-*vector space*) if there are defined two operations,

(a) an internal one: a sum of vectors $s : V \times V \to V$,

$$V \times V \ni (v, v') \mapsto s(v, v') = v + v',$$

(b) an exterior one: the product by a scalar $p : \mathbb{R} \times V \to V$

$$\mathbb{R} \times V \ni (k, v) \mapsto p(k, v) = kv,$$

and these operations are required to satisfy the following conditions:

(1) There exists an element $0_V \in V$, which is neutral for the sum, such that $(V, +, 0_V)$ is an abelian group.
For any $k, k' \in \mathbb{R}$ and $v, v' \in V$ one has
(2) $(k + k')v = kv + k'v$
(3) $k(v + v') = kv + kv'$

G. Landi and A. Zampini, *Linear Algebra and Analytic Geometry for Physical Sciences*, Undergraduate Lecture Notes in Physics,
https://doi.org/10.1007/978-3-319-78361-1_2

(4) $k(k'v) = (kk')v$
(5) $1v = v$, with $1 = 1_{\mathbb{R}}$.

The elements of a vector space are called *vectors*; the element 0_V is the *zero* or *null* vector. A vector space is also called a *linear space*.

Remark 2.1.2 Given the properties of a group (see A.2.9), the null vector 0_V and the opposite $-v$ to any vector v are (in any given vector space) unique. The sums can be indeed simplified, that is $v + w = v + u \Longrightarrow w = u$. Such a statement is easily proven by adding to both terms in $v + w = v + u$ the element $-v$ and using the associativity of the sum.

As already seen in Chap. 1, the collections $\mathcal{V}_O^2$ (vectors in a plane) and $\mathcal{V}_O^3$ (vectors in the space) applied at the point O are real vector spaces. The bijection $\mathcal{V}_O^3 \longleftrightarrow \mathbb{R}^3$ introduced in the Definition 1.2.5, together with the Remark 1.2.9, suggest the natural definitions of sum and product by a scalar for the set $\mathbb{R}^3$ of ordered triples of real numbers.

Proposition 2.1.3 *The collection $\mathbb{R}^3$ of triples of real numbers together with the operations defined by*

I. $(x_1, x_2, x_3) + (y_1, y_2, y_3) = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$, *for any* (x_1, x_2, x_3), $(y_1, y_2, y_3) \in \mathbb{R}^3$,
II. $a(x_1, x_2, x_3) = (ax_1, ax_2, ax_3)$, *for any* $a \in \mathbb{R}$, $(x_1, x_2, x_3) \in \mathbb{R}^3$,

is a real vector space.

Proof We verify that the conditions given in the Definition 2.1.1 are satisfied. We first notice that (a) and (b) are fullfilled, since $\mathbb{R}^3$ is closed with respect to the operations in I. and II. of sum and product by a scalar. The neutral element for the sum is $0_{\mathbb{R}^3} = (0, 0, 0)$, since one clearly has

$$(x_1, x_2, x_3) + (0, 0, 0) = (x_1, x_2, x_3).$$

The datum $(\mathbb{R}^3, +, 0_{\mathbb{R}^3})$ is an abelian group, since one has

- The sum $(\mathbb{R}^3, +)$ is associative, from the associativity of the sum in $\mathbb{R}$:

$$\begin{aligned}
(x_1, x_2, x_3) &+ ((y_1, y_2, y_3) + (z_1, z_2, z_3)) \\
&= (x_1, x_2, x_3) + (y_1 + z_1, y_2 + z_2, y_3 + z_3) \\
&= (x_1 + (y_1 + z_1), x_2 + (y_2 + z_2), x_3 + (y_3 + z_3)) \\
&= ((x_1 + y_1) + z_1, (x_2 + y_2) + z_2, (x_3 + y_3) + z_3) \\
&= (x_1 + y_1, x_2 + y_2, x_3 + y_3) + (z_1, z_2, z_3) \\
&= ((x_1, x_2, x_3) + (y_1, y_2, y_3)) + (z_1, z_2, z_3).
\end{aligned}$$

- From the identity

$$(x_1, x_2, x_3) + (-x_1, -x_2, -x_3) = (x_1 - x_1, x_2 - x_2, x_3 - x_3) = (0, 0, 0)$$

  one has $(-x_1, -x_2, -x_3)$ as the opposite in $\mathbb{R}^3$ of the element (x_1, x_2, x_3).
- The group $(\mathbb{R}^3, +)$ is commutative, since the sum in $\mathbb{R}$ is commutative:

$$\begin{aligned}(x_1, x_2, x_3) + (y_1, y_2, y_3) &= (x_1 + y_1, x_2 + y_2, x_3 + y_3) \\ &= (y_1 + x_1, y_2 + x_2, y_3 + x_3) \\ &= (y_1, y_2, y_3) + (x_1, x_2, x_3).\end{aligned}$$

We leave to the reader the task to show that the conditions (1), (2), (3), (4) in Definition 2.1.1 are satisfied: for any $\lambda, \lambda' \in \mathbb{R}$ and any $(x_1, x_2, x_3), (y_1, y_2, y_3) \in \mathbb{R}^3$ it holds that

1. $(\lambda + \lambda')(x_1, x_2, x_3) = \lambda(x_1, x_2, x_3) + \lambda'(x_1, x_2, x_3)$
2. $\lambda((x_1, x_2, x_3) + (y_1, y_2, y_3)) = \lambda(x_1, x_2, x_3) + \lambda(y_1, y_2, y_3)$
3. $\lambda(\lambda'(x_1, x_2, x_3)) = (\lambda\lambda')(x_1, x_2, x_3)$
4. $1(x_1, x_2, x_3) = (x_1, x_2, x_3)$. □

The previous proposition can be generalised in a natural way. If $n \in \mathbb{N}$ is a positive natural number, one defines the n-th cartesian product of $\mathbb{R}$, that is the collection of ordered n-tuples of real numbers

$$\mathbb{R}^n = \{X = (x_1, \ldots, x_n) \ : \ x_k \in \mathbb{R}\},$$

and the following operations, with $a \in \mathbb{R}$, $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{R}^n$:

In. $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$
IIn. $a(x_1, \ldots, x_n) = (ax_1, \ldots, ax_n)$.

The previous proposition can be directly generalised to the following.

Proposition 2.1.4 *With respect to the above operations, the set $\mathbb{R}^n$ is a vector space over $\mathbb{R}$.*

The elements in $\mathbb{R}^n$ are called *n-tuples* of real numbers. With the notation $X = (x_1, \ldots, x_n) \in \mathbb{R}^n$, the scalar x_k, with $k = 1, 2, \ldots, n$, is the *k-th component* of the vector X.

Example 2.1.5 As in the Definition A.3.3, consider the collection of all polynomials in the indeterminate x and coefficients in $\mathbb{R}$, that is

$$\mathbb{R}[x] = \left\{ f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \ : \ a_k \in \mathbb{R}, \ n \geq 0 \right\},$$

with the operations of sum and product by a scalar $\lambda \in \mathbb{R}$ defined, for any pair of elements in $\mathbb{R}[x]$, $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ and $g(x) = b_0 + b_1 x + b_2 x^2 + \cdots + b_m x^m$, component-wise by

Ip. $f(x) + g(x) = a_0 + b_0 + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots$
IIp. $\lambda f(x) = \lambda a_0 + \lambda a_1 x + \lambda a_2 x^2 + \cdots + \lambda a_n x^n$.

Endowed with the previous operations, the set $\mathbb{R}[x]$ is a real vector space; $\mathbb{R}[x]$ is indeed closed with respect to the operations above. The null polynomial, denoted by $0_{\mathbb{R}[x]}$ (that is the polynomial with all coefficients equal zero), is the neutral element for the sum. The opposite to the polynomial $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ is the polynomial $(-a_0 - a_1 x - a_2 x^2 - \cdots - a_n x^n) \in \mathbb{R}[x]$ that one denotes by $-f(x)$. We leave to the reader to prove that $(\mathbb{R}[x], +, 0_{\mathbb{R}[x]})$ is an abelian group and that all the additional conditions in Definition 2.1.1 are fulfilled.

Exercise 2.1.6 We know from the Proposition A.3.5 that $\mathbb{R}[x]_r$, the subset in $\mathbb{R}[x]$ of polynomials with degree not larger than a fixed $r \in \mathbb{N}$, is closed under addition of polynomials. Since the degree of the polynomial $\lambda f(x)$ coincides with the degree of $f(x)$ for any $\lambda \neq 0$, we see that also the product by a scalar, as defined in IIp. above, is defined consistently on $\mathbb{R}[x]_r$. It is easy to verify that also $\mathbb{R}[x]_r$ is a real vector space.

Remark 2.1.7 The proof that $\mathbb{R}^n$, $\mathbb{R}[x]$ and $\mathbb{R}[x]_r$ are vector space over $\mathbb{R}$ relies on the properties of $\mathbb{R}$ as a field (in fact a ring, since the multiplicative inverse in $\mathbb{R}$ does not play any role).

Exercise 2.1.8 The set $\mathbb{C}^n$, that is the collection of ordered n-tuples of complex numbers, can be given the structure of a vector space over $\mathbb{C}$. Indeed, both the operations In. and IIn. considered in the Proposition 2.1.3 when intended for complex numbers make perfectly sense:

Ic. $(z_1, \ldots, z_n) + (w_1, \ldots, w_n) = (z_1 + w_1, \ldots, z_n + w_n)$
IIc. $c(z_1, \ldots, z_n) = (cz_1, \ldots, cz_n)$

with $c \in \mathbb{C}$, and $(z_1, \ldots, z_n), (w_1, \ldots, w_n) \in \mathbb{C}^n$.

The reader is left to show that $\mathbb{C}^n$ is a vector space over $\mathbb{C}$.

The space $\mathbb{C}^n$ can also be given a structure of vector space over $\mathbb{R}$, by noticing that the product of a complex number by a real number is a complex number. This means that $\mathbb{C}^n$ is closed with respect to the operations of (component-wise) product by a real scalar. The condition IIc. above makes sense when $c \in \mathbb{R}$.

We next analyse some elementary properties of general vector spaces.

Proposition 2.1.9 *Let V be a vector space over $\mathbb{R}$. For any $k \in \mathbb{R}$ and any $v \in V$ it holds that:*

(i) $0_{\mathbb{R}} v = 0_V$,
(ii) $k 0_V = 0_V$,
(iii) if $kv = 0_V$ *then it is either* $k = 0_{\mathbb{R}}$ *or* $v = 0_V$,
(iv) $(-k)v = -(kv) = k(-v)$.

Proof (i) From $0_{\mathbb{R}} v = (0_{\mathbb{R}} + 0_{\mathbb{R}})v = 0_{\mathbb{R}} v + 0_{\mathbb{R}} v$, since the sums can be simplified, one has that $0_{\mathbb{R}} v = 0_V$.
(ii) Analogously: $k0_V = k(0_V + 0_V) = k0_V + k0_V$ which yields $k0_V = 0_V$.

(iii) Let $k \neq 0$, so $k^{-1} \in \mathbb{R}$ exists. Then, $v = 1v = k^{-1}kv = k^{-1}0_V = 0_V$, with the last equality coming from (ii).
(iv) Since the product is distributive over the sum, from (i) it follows that $kv + (-k)v = (k + (-k))v = 0_{\mathbb{R}}v = 0_V$ that is the first equality. For the second, one writes analogously $kv + k(-v) = k(v - v) = k0_V = 0_V$ □

Relations (i), (ii), (iii) above are more succinctly expressed by the equivalence:

$$kv = 0_V \iff k = 0_{\mathbb{R}} \quad \text{or} \quad v = 0_V.$$

2.2 Vector Subspaces

Among the subsets of a real vector space, of particular relevance are those which inherit from V a vector space structure.

Definition 2.2.1 Let V be a vector space over $\mathbb{R}$ with respect to the sum s and the product p as given in the Definition 2.1.1. Let $W \subseteq V$ be a subset of V. One says that W is a *vector subspace* of V if the restrictions of s and p to W equip W with the structure of a vector space over $\mathbb{R}$.

In order to establish whether a subset $W \subseteq V$ of a vector space is a vector subspace, the following can be seen as *criteria*.

Proposition 2.2.2 *Let W be a non empty subset of the real vector space V. The following conditions are equivalent.*

(i) W is a vector subspace of V,
(ii) W is closed with respect to the sum and the product by a scalar, that is

(a) $w + w' \in W$, for any $w, w' \in W$,
(b) $kw \in W$, for any $k \in \mathbb{R}$ and $w \in W$,

(iii) $kw + k'w' \in W$, for any $k, k' \in \mathbb{R}$ and any $w, w' \in W$.

Proof The implications (i) $\Longrightarrow$ ii) and (ii) $\Longrightarrow$ (iii) are obvious from the definition.

(iii) $\Longrightarrow$ (ii): By taking $k = k' = 1$ one obtains (a), while to show point (b) one takes $k' = 0_{\mathbb{R}}$.

(ii) $\Longrightarrow$ (i): Notice that, by hypothesis, W is closed with respect to the sum and product by a scalar. Associativity and commutativity hold in W since they hold in V.

One only needs to prove that W has a neutral element 0_W and that, for such a neutral element, any vector in W has an opposite in W. If $0_V \in W$, then 0_V is the zero element in W: for any $w \in W$ one has $0_V + w = w + 0_V = w$ since $w \in V$; from ii, (b) one has $0_{\mathbb{R}}w \in W$ for any $w \in W$; from the Proposition 2.1.9 one has $0_{\mathbb{R}}w = 0_V$; collecting these relations, one concludes that $0_V \in W$. If $w \in W$, again from the Proposition 2.1.9 one gets that $-w = (-1)w \in W$. □

Exercise 2.2.3 Both $W = \{0_V\} \subset V$ and $W = V \subseteq V$ are *trivial* vector subspaces of V.

Exercise 2.2.4 We have already seen that $\mathbb{R}[x]_r \subseteq \mathbb{R}[x]$ are vector spaces with respect to the same operations, so we may conclude that $\mathbb{R}[x]_r$ is a vector subspace of $\mathbb{R}[x]$.

Exercise 2.2.5 Let $v \in V$ a non zero vector in a vector space, and let

$$\mathcal{L}(v) = \{av \;:\; a \in \mathbb{R}\} \subset V$$

be the collection of all multiples of v by a real scalar. Given the elements $w = av$ and $w' = a'v$ in $\mathcal{L}(v)$, from the equality

$$\alpha w + \alpha' w' = (\alpha a + \alpha' a')v \in \mathcal{L}(v)$$

for any $\alpha, \alpha' \in \mathbb{R}$, we see that, from the Proposition 2.2.2, $\mathcal{L}(v)$ is a vector subspace of V, and we call it the (*vector*) *line generated by* v.

Exercise 2.2.6 Consider the following subsets $W \subset \mathbb{R}^2$:

1. $W_1 = \{(x, y) \in \mathbb{R}^2 \;:\; x - 3y = 0\}$,
2. $W_2 = \{(x, y) \in \mathbb{R}^2 \;:\; x + y = 1\}$,
3. $W_3 = \{(x, y) \in \mathbb{R}^2 \;:\; x \in \mathbb{N}\}$,
4. $W_4 = \{(x, y) \in \mathbb{R}^2 \;:\; x^2 - y = 0\}$.

From the previous exercise, one sees that W_1 is a vector subspace since $W_1 = \mathcal{L}((3, 1))$. On the other hand, W_2, W_3, W_4 are not vector subspaces of $\mathbb{R}^2$. The zero vector $(0, 0) \notin W_2$; while W_3 and W_4 are not closed with respect to the product by a scalar, since, for example, $(1, 0) \in W_3$ but $\frac{1}{2}(1, 0) = (\frac{1}{2}, 0) \notin W_3$. Analogously, $(1, 1) \in W_4$ but $2(1, 1) = (2, 2) \notin W_4$.

The next step consists in showing how, given two or more vector subspaces of a real vector space V, one can define new vector subspaces of V via suitable operations.

Proposition 2.2.7 *The intersection $W_1 \cap W_2$ of any two vector subspaces W_1 and W_2 of a real vector space V is a vector subspace of V.*

Proof Consider $a, b \in \mathbb{R}$ and $v, w \in W_1 \cap W_2$. From the Propostion 2.2.2 it follows that $av + bw \in W_1$ since W_1 is a vector subspace, and also that $av + bw \in W_2$ for the same reason. As a consequence, one has $av + bw \in W_1 \cap W_2$. □

Remark 2.2.8 In general, the union of two vector subspaces of V is *not* a vector subspace of V. As an example, the Fig. 2.1 shows that, if $\mathcal{L}(v)$ and $\mathcal{L}(w)$ are generated by different $v, w \in \mathbb{R}^2$, then $\mathcal{L}(v) \cup \mathcal{L}(w)$ is not closed under the sum, since it does not contain the sum $v + w$, for instance.

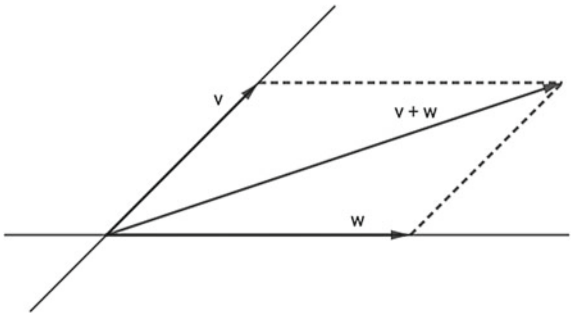

Fig. 2.1 The vector line $\mathcal{L}(v + w)$ with respect to the vector lines $\mathcal{L}(v)$ and $\mathcal{L}(w)$

Proposition 2.2.9 *Let W_1 and W_2 be vector subspaces of the real vector space V and let $W_1 + W_2$ denote*

$$W_1 + W_2 = \{v \in V \mid v = w_1 + w_2;\ w_1 \in W_1,\ w_2 \in W_2\} \subset V.$$

Then $W_1 + W_2$ is the smallest vector subspace of V which contains the union $W_1 \cup W_2$.

Proof Let $a, a' \in \mathbb{R}$ and $v, v' \in W_1 + W_2$; this means that there exist $w_1, w_1' \in W_1$ and $w_2, w_2' \in W_2$, so that $v = w_1 + w_2$ and $v' = w_1' + w_2'$. Since both W_1 and W_2 are vector subspaces of V, from the identity

$$av + a'v' = aw_1 + aw_2 + a'w_1' + a'w_2' = (aw_1 + a'w_1') + (aw_2 + a'w_2'),$$

one has $aw_1 + a'w_1' \in W_1$ and $aw_2 + a'w_2' \in W_2$. It follows that $W_1 + W_2$ is a vector subspace of V.

It holds that $W_1 + W_2 \supseteq W_1 \cup W_2$: if $w_1 \in W_1$, it is indeed $w_1 = w_1 + 0_V$ in $W_1 + W_2$; one similarly shows that $W_2 \subset W_1 + W_2$.

Finally, let Z be a vector subspace of V containing $W_1 \cup W_2$; then for any $w_1 \in W_1$ and $w_2 \in W_2$ it must be $w_1 + w_2 \in Z$. This implies $Z \supseteq W_1 + W_2$, and then $W_1 + W_2$ is the smallest of such vector subspaces Z. □

Definition 2.2.10 If W_1 and W_2 are vector subspaces of the real vector space V the vector subspace $W_1 + W_2$ of V is called the *sum* of W_1 e W_2.

The previous proposition and definition are easily generalised, in particular:

Definition 2.2.11 If $W_1, \ldots, W_n$ are vector subspaces of the real subspace V, the vector subspace

$$W_1 + \cdots + W_n = \{v \in V \mid v = w_1 + \cdots + w_n;\ w_i \in W_i,\ i = 1, \ldots, n\}$$

of V is the *sum* of $W_1, \ldots, W_n$.

Definition 2.2.12 Let W_1 and W_2 be vector subspaces of the real vector space V. The sum $W_1 + W_2$ is called *direct* if $W_1 \cap W_2 = \{0_V\}$. A direct sum is denoted $W_1 \oplus W_2$.

Proposition 2.2.13 *Let W_1, W_2 be vector subspaces of the real vector space V. Their sum $W = W_1 + W_2$ is direct if and only if any element $v \in W_1 + W_2$ has a* unique *decomposition as $v = w_1 + w_2$ with $w_i \in W_i$, $i = 1, 2$.*

Proof We first suppose that the sum $W_1 + W_2$ is direct, that is $W_1 \cap W_2 = \{0_V\}$. If there exists an element $v \in W_1 + W_2$ with $v = w_1 + w_2 = w_1' + w_2'$, and $w_i, w_i' \in W_i$, then $w_1 - w_1' = w_2' - w_2$ and such an element would belong to both W_1 and W_2. This would then be zero, since $W_1 \cap W_2 = \{0_V\}$, and then $w_1 = w_1'$ and $w_2 = w_2'$.

Suppose now that any element $v \in W_1 + W_2$ has a unique decomposition $v = w_1 + w_2$ with $w_i \in W_i$, $i = 1, 2$. Let $v \in W_1 \cap W_2$; then $v \in W_1$ and $v \in W_2$ which gives $0_V = v - v \in W_1 + W_2$, so the zero vector has a unique decomposition. But clearly also $0_V = 0_V + 0_V$ and being the decomposition for 0_V unique, this gives $v = 0_V$. □

These proposition gives a natural way to generalise the notion of direct sum to an arbitrary number of vector subspaces of a given vector space.

Definition 2.2.14 Let $W_1, \ldots, W_n$ be vector subspaces of the real vector space V. The sum $W_1 + \cdots + W_n$ is called *direct* if any of its element has a unique decomposition as $v = w_1 + \cdots + w_n$ with $w_i \in W_i$, $i = 1, \ldots, n$. The direct sum vector subspace is denoted $W_1 \oplus \cdots \oplus W_n$.

2.3 Linear Combinations

We have seen in Chap. 1 that, given a cartesian coordinate system $\Sigma = (O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ for the space $\mathcal{S}$, any vector $\mathbf{v} \in \mathcal{V}_O^3$ can be written as $\mathbf{v} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$. One says that $\mathbf{v}$ is a *linear combination* of $\mathbf{i}, \mathbf{j}, \mathbf{k}$. From the Definition 1.2.5 we also know that, given Σ, the components (a, b, c) are uniquely determined by $\mathbf{v}$. For this one says that $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are *linearly independent*. In this section we introduce these notions for an arbitrary vector space.

Definition 2.3.1 Let $v_1, \ldots, v_n$ be arbitrary elements of a real vector space V. A vector $v \in V$ is a *linear combination* of $v_1, \ldots, v_n$ if there exist n scalars $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, such that

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n.$$

The collection of all linear combinations of the vectors $v_1, \ldots, v_n$ is denoted by $\mathcal{L}(v_1, \ldots, v_n)$. If $I \subseteq V$ is an arbitrary subset of V, by $\mathcal{L}(I)$ one denotes the collection of all possible linear combinations of vectors in I, that is

$$\mathcal{L}(I) = \{\lambda_1 v_1 + \cdots + \lambda_n v_n \mid \lambda_i \in \mathbb{R},\ v_i \in I,\ n \geq 0\}.$$

The set $\mathcal{L}(I)$ is also called the *linear span* of I.

Proposition 2.3.2 *The space $\mathcal{L}(v_1, \dots, v_n)$ is a vector subspace of V, called the space* generated *by $v_1, \dots, v_n$ or the* linear span *of the vectors $v_1, \dots, v_n$.*

Proof After Proposition 2.2.2, it is enough to show that $\mathcal{L}(v_1, \dots, v_n)$ is closed for the sum and the product by a scalar. Let $v, w \in \mathcal{L}(v_1, \dots, v_n)$; it is then $v = \lambda_1 v_1 + \dots + \lambda_n v_n$ and $w = \mu_1 v_1 + \dots + \mu_n v_n$, for scalars $\lambda_1, \dots, \lambda_n$ and $\mu_1, \dots, \mu_n$. Recalling point (2) in the Definition 2.1.1, one has

$$v + w = (\lambda_1 + \mu_1)v_1 + \dots + (\lambda_n + \mu_n)v_n \in \mathcal{L}(v_1, \dots, v_n).$$

Next, let $\alpha \in \mathbb{R}$. Again from the Definition 2.1.1 (point 4)), one has $\alpha v = (\alpha\lambda_1)v_1 + \dots + (\alpha\lambda_n)v_n$, which gives $\alpha v \in \mathcal{L}(v_1, \dots, v_n)$. □

Exercise 2.3.3 The following are two examples for the notion just introduced.

(1) Clearly one has $\mathcal{V}_O^2 = \mathcal{L}(\mathbf{i}, \mathbf{j})$ and $\mathcal{V}_O^3 = \mathcal{L}(\mathbf{i}, \mathbf{j}, \mathbf{k})$.
(2) Let $v = (1, 0, -1)$ and $w = (2, 0, 0)$ be two vectors in $\mathbb{R}^3$; it is easy to see that $\mathcal{L}(v, w)$ is a proper subset of $\mathbb{R}^3$. For example, the vector $\mathbf{u} = (0, 1, 0) \notin \mathcal{L}(v, w)$. If $\mathbf{u}$ were in $\mathcal{L}(v, w)$, there should be $\alpha, \beta \in \mathbb{R}$ such that
$$(0, 1, 0) = \alpha(1, 0, -1) + \beta(2, 0, 0) = (\alpha + 2\beta, 0, -\alpha).$$
No choice of $\alpha, \beta \in \mathbb{R}$ can satisfy this vector identity, since the second component equality would give $1 = 0$, independently of α, β.

It is interesting to explore which subsets $I \subseteq V$ yield $\mathcal{L}(I) = V$. Clearly, one has $V = \mathcal{L}(V)$. The example (1) above shows that there are proper subsets $I \subset V$ whose linear span coincides with V itself. We already know that $\mathcal{V}_O^2 = \mathcal{L}(\mathbf{i}, \mathbf{j})$ and that $\mathcal{V}_O^3 = \mathcal{L}(\mathbf{i}, \mathbf{j}, \mathbf{k})$: both $\mathcal{V}_O^3$ and $\mathcal{V}_O^2$ are generated by a *finite* number of (their) vectors. This is not always the case, as the following exercise shows.

Exercise 2.3.4 The real vector space $\mathbb{R}[x]$ is *not* generated by a finite number of vectors. Indeed, let $f_1(x), \dots, f_n(x) \in \mathbb{R}[x]$ be arbitrary polynomials. Any $p(x) \in \mathcal{L}(f_1, \dots, f_n)$ is written as

$$p(x) = \lambda_1 f_1(x) + \dots + \lambda_n f_n(x)$$

with suitable $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. If one writes $d_i = \deg(f_i)$ and $d = \max\{d_1, \dots, d_n\}$, from Remark A.3.5 one has that

$$\deg(p(x)) = \deg(\lambda_1 f_1(x) + \dots + \lambda_n f_n(x)) \leq \max\{d_1, \dots, d_n\} = d.$$

This means that any polynomial of degree $d + 1$ or higher is *not* contained in $\mathcal{L}(f_1, \dots, f_n)$. This is the case for any *finite* n, giving a *finite* d; we conclude that, if n is finite, any $\mathcal{L}(I)$ with $I = (f_1(x), \dots, f_n(x))$ is a proper subset of $\mathbb{R}[x]$ which can then *not* be generated by a *finite* number of polynomials.

On the other hand, $\mathbb{R}[x]$ is indeed the linear span of the infinite set

$$\{1, x, x^2, \ldots, x^i, \ldots\}.$$

Definition 2.3.5 A vector space V over $\mathbb{R}$ is said to be *finitely generated* if there exists a finite number of elements $v_1, \ldots, v_n$ in V which are such that $V = \mathcal{L}(v_1, \ldots, v_n)$. In such a case, the set $\{v_1, \ldots, v_n\}$ is called a *system of generators* for V.

Proposition 2.3.6 *Let $I \subseteq V$ and $v \in V$. It holds that*

$$\mathcal{L}(\{v\} \cup I) = \mathcal{L}(I) \quad \Longleftrightarrow \quad v \in \mathcal{L}(I).$$

Proof "$\Rightarrow$" Let us assume that $\mathcal{L}(\{v\} \cup I) = \mathcal{L}(I)$. Since $v \in \mathcal{L}(\{v\} \cup I)$, then $v \in \mathcal{L}(I)$.

"$\Leftarrow$" We shall prove the claim under the hypothesis that we have a finite system $\{v_1, \ldots, v_n\}$. The inclusion $\mathcal{L}(I) \subseteq \mathcal{L}(\{v\} \cup I)$ is obvious. To prove the inclusion $\mathcal{L}(\{v\} \cup I) \subseteq \mathcal{L}(I)$, consider an arbitrary element $w \in \mathcal{L}(\{v\} \cup I)$, so that $w = \alpha v + \mu_1 v_1 + \cdots + \mu_n v_n$. By the hypothesis, $v \in \mathcal{L}(I)$ so it is $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$. We can then write

$$w = \alpha(\lambda_1 v_1 + \cdots + \lambda_n v_n) + \mu_1 v_1 + \cdots + \mu_n v_n.$$

From the properties of the sum of vectors in V, one concludes that $w \in \mathcal{L}(v_1, \ldots, v_n) = \mathcal{L}(I)$. □

Remark 2.3.7 From the previous proposition one has also the identity

$$\mathcal{L}(v_1, \ldots, v_n, 0_V) = \mathcal{L}(v_1, \ldots, v_n)$$

for any $v_1, \ldots, v_n \in V$.

If I is a system of generators for V, the next question to address is whether I contains a *minimal* set of generators for V, that is whether there exists a set $J \subset I$ (with $J \neq I$) such that $\mathcal{L}(J) = \mathcal{L}(I) = V$. The answer to this question leads to the notion of *linear independence* for a set of vectors.

Definition 2.3.8 Given a collection $I = \{v_1, \ldots, v_n\}$ of vectors in a real vector space V, the elements of I are called *linearly independent* on $\mathbb{R}$, and the system I is said to be *free*, if the following implication holds,

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0_V \quad \Longrightarrow \quad \lambda_1 = \cdots = \lambda_n = 0_{\mathbb{R}}.$$

That is, if the *only* linear combination of elements of I giving the zero vector is the one whose coefficients are all zero.

Analogously, an infinite system $I \subseteq V$ is said to be *free* if any of its finite subsets is free.

The vectors $v_1, \dots, v_n \in V$ are said to be *linearly dependent* if they are not linearly independent, that is if there are scalars $(\lambda_1, \dots, \lambda_n) \neq (0, \dots, 0)$ such that $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0_V$.

Exercise 2.3.9 It is clear that $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are linearly independent in $\mathcal{V}_O^3$, while the vectors $v_1 = \mathbf{i} + \mathbf{j}$, $v_2 = \mathbf{j} - \mathbf{k}$ and $v_3 = 2\mathbf{i} - \mathbf{j} + 3\mathbf{k}$ are linearly dependent, since one computes that $2v_1 - 3v_2 - v_3 = 0$.

Proposition 2.3.10 *Let V be a real vector space and $I = \{v_1, \dots, v_n\}$ be a collection of vectors in V. The following properties hold true:*

(i) if $0_V \in I$, then I is not free,
(ii) I is not free if and only if one of the elements v_i is a linear combination of the other elements $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$,
(iii) if I is not free, then any $J \supseteq I$ is not free,
(iv) if I is free, then any J such that $J \subseteq I$ is free; that is any subsystem of a free system is free.

Proof i) Without loss of generality we suppose that $v_1 = 0_V$. Then, one has

$$1_{\mathbb{R}} v_1 + 0_{\mathbb{R}} v_2 + \cdots + 0_{\mathbb{R}} v_n = 0_V,$$

which amounts to say that the zero vector can be written as a linear combination of elements in I with a non zero coefficients.

(ii) Suppose I is not free. Then, there exists scalars $(\lambda_1, \dots, \lambda_n) \neq (0, \dots, 0)$ giving the combination $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0_V$. Without loss of generality take $\lambda_1 \neq 0$; so λ_1 is invertible and we can write

$$v_1 = \lambda_1^{-1}(-\lambda_2 v_2 - \cdots - \lambda_n v_n) \in \mathcal{L}(v_2, \dots, v_n).$$

In order to prove the converse, we start by assuming that a vector v_i is a linear combination

$$v_i = \lambda_1 v_1 + \cdots + \lambda_{i-1} v_{i-1} + \lambda_{i+1} v_{i+1} + \cdots + \lambda_n v_n.$$

This identity can be written in the form

$$\lambda_1 v_1 + \cdots + \lambda_{i-1} v_{i-1} - v_i + \lambda_{i+1} v_{i+1} + \cdots + \lambda_n v_n = 0_V.$$

The zero vector is then written as a linear combination with coefficients not all identically zero, since the coefficient of v_i is -1. This amounts to say that the system I is not free.

We leave the reader to show the obvious points (iii) and (iv). □

2.4 Bases of a Vector Space

Given a real vector space V, in this section we determine its smallest possible systems of generators, together with their cardinalities.

Proposition 2.4.1 *Let V be a real vector space, with $v_1, \dots, v_n \in V$. The following facts are equivalent:*

(i) the elements $v_1, \dots, v_n$ are linearly independent,
(ii) $v_1 \neq 0_V$ and, for any $i \geq 2$, the vector v_i is not a linear combination of $v_1, \dots, v_{i-1}$.

Proof The implication (i) $\Longrightarrow$ (ii) directly follows from the Proposition 2.3.10.

To show the implication (ii) $\Longrightarrow$ (i) we start by considering a combination $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0_V$. Under the hypothesis, v_n is not a linear combination of $v_1, \dots, v_{n-1}$, so it must be $\lambda_n = 0$: were it not, one could write $v_n = \lambda_n^{-1}(-\lambda_1 v_1 - \cdots - \lambda_{n-1} v_{n-1})$. We are then left with $\lambda_1 v_1 + \cdots + \lambda_{n-1} v_{n-1} = 0_V$, and an analogous reasoning leads to $\lambda_{n-1} = 0$. After $n - 1$ similar steps, one has $\lambda_1 v_1 = 0$; since $v_1 \neq 0$ by hypothesis, it must be (see 2.1.5) that $\lambda_1 = 0$. □

Theorem 2.4.2 *Any finite system of generators for a vector space V contains a free system of generators for V.*

Proof Let $I = \{v_1, \dots, v_s\}$ be a system of generators for a real vector space V. Recalling the Remark 2.3.7, we can take $v_i \neq 0$ for any $i = 1, \dots, s$. We define iteratively a system of subsets of I, as follows:

- take $I_1 = I = \{v_1, \dots, v_s\}$,
- if $v_2 \in \mathcal{L}(v_1)$, take $I_2 = I_1 \setminus \{v_2\}$; if $v_2 \neq \mathcal{L}(v_1)$, take $I_2 = I_1$,
- if $v_3 \in \mathcal{L}(v_1, v_2)$, take $I_3 = I_2 \setminus \{v_3\}$; if $v_3 \neq \mathcal{L}(v_1, v_2)$, take $I_3 = I_2$,
- Iterate the steps above.

The whole procedure consists in examining any element in the starting $I_1 = I$, and deleting it if it is a linear combination of the previous ones. After s steps, one ends up with a chain $I_1 \supseteq \cdots \supseteq I_s \supseteq I$.

Notice that, for any $j = 2, \dots, s$, it is $\mathcal{L}(I_j) = \mathcal{L}(I_{j-1})$. It is indeed either $I_j = I_{j-1}$ (which makes the claim obvious) or $I_{j-1} = I_j \cup \{v_j\}$, with $v_j \in \mathcal{L}(v_1, \dots, v_{j-1}) \subseteq \mathcal{L}(I_{j-1})$; from Proposition 2.3.6, it follows that $\mathcal{L}(I_j) = \mathcal{L}(I_{j-1})$.

One has then $\mathcal{L}(I) = \mathcal{L}(I_1) = \cdots = \mathcal{L}(I_s)$, and I_s is a system of generators of V. Since no element in I_s is a linear combination of the previous ones, the Proposition 2.4.1 shows that I_s is free. □

Definition 2.4.3 Let V be a real vector space. An ordered system of vectors $I = (v_1, \dots, v_n)$ in V is called a *basis* of V if I is a free system of generators for V, that is $V = \mathcal{L}(v_1, \dots, v_n)$ and $v_1, \dots, v_n$ are linearly independent.

Corollary 2.4.4 *Any finite system of generators for a vector space contains (at least) a basis. This means also that any finitely generated vector space has a basis.*

Proof It follows directly from the Theorem 2.4.2. □

Exercise 2.4.5 Consider the vector space $\mathbb{R}^3$ and the system of vectors $I = \{v_1, \ldots, v_5\}$ with

$$v_1 = (1, 1, -1), \quad v_2 = (-2, -2, 2), \quad v_3 = (2, 0, 1), \quad v_4 = (1, -1, 2), \quad v_5 = (0, 1, 1).$$

Following Theorem 2.4.2, we determine a basis for $\mathcal{L}(v_1, v_2, v_3, v_4, v_5)$.

- At the first step $I_1 = I$.
- Since $v_2 = -2v_1$, so that $v_2 \in \mathcal{L}(v_1)$, delete v_2 and take $I_2 = I_1 \setminus \{v_2\}$.
- One has $v_3 \notin \mathcal{L}(v_1)$, so keep v_3 and take $I_3 = I_2$.
- One has $v_4 \in \mathcal{L}(v_1, v_3)$ if and only if there exist $\alpha, \beta \in \mathbb{R}$ such that $v_4 = \alpha v_1 + \beta v_3$, that is $(1, -1, 2) = (\alpha + 2\beta, \alpha, -\alpha + \beta)$. By equating components, one has $\alpha = -1$, $\beta = 1$. This shows that $v_4 = -v_1 + v_3 \in \mathcal{L}(v_1, v_3)$; therefore delete v_4 and take $I_4 = I_3 \setminus \{v_4\}$.
- Similarly one shows that $v_5 \notin \mathcal{L}(v_1, v_3)$. A basis for $\mathcal{L}(I)$ is then $I_5 = I_4 = (v_1, v_3, v_5)$.

The next theorem characterises free systems.

Theorem 2.4.6 *A system $I = \{v_1, \ldots, v_n\}$ of vectors in V is free if and only if any element in $\mathcal{L}(v_1, \ldots, v_n)$ can be written in a* unique *way as a linear combination of the elements $v_1, \ldots, v_n$.*

Proof We assume that I is free and that $\mathcal{L}(v_1, \ldots, v_n)$ contains a vector, say v, which has two linear decompositions with respect to the vectors v_i:

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n = \mu_1 v_1 + \cdots + \mu_n v_n.$$

This identity would give $(\lambda_1 - \mu_1)v_1 + \cdots + (\lambda_n - \mu_n)v_n = 0_V$; since the elements v_i are linearly independent it would read

$$\lambda_1 - \mu_1 = 0, \quad \cdots \quad , \lambda_n - \mu_n = 0,$$

that is $\lambda_i = \mu_i$ for any $i = 1, \ldots, n$. This says that the two linear expressions above coincide and v is written in a unique way.

We assume next that any element in $\mathcal{L}(v_1, \ldots, v_n)$ as a unique linear decomposition with respect to the vectors v_i. This means that the zero vector $0_V \in \mathcal{L}(v_1, \ldots, v_n)$ has the unique decomposition $0_V = 0_{\mathbb{R}} v_1 + \cdots + 0_{\mathbb{R}} v_n$. Let us consider the expression $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0_V$; since the linear decomposition of 0_V is unique, it is $\lambda_i = 0$ for any $i = 1, \ldots, n$. This says that the vectors $v_1, \ldots, v_n$ are linearly independent. □

Corollary 2.4.7 *Let $v_1, \ldots, v_n$ be elements of a real vector space V. The system $I = (v_1, \ldots, v_n)$ is a basis for V if and only if any element $v \in V$ can be written in a unique way as $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$.*

Definition 2.4.8 Let $I = (v_1, \ldots, v_n)$ be a basis for the real vector space V. Any $v \in V$ is then written as a linear combination $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$ is a unique way. The scalars $\lambda_1, \ldots, \lambda_n$ (which are uniquely determined by Corollary 2.4.7) are called the *components* of v on the basis I. We denote this by

$$v = (\lambda_1, \ldots, \lambda_n)_I.$$

Remark 2.4.9 Notice that we have taken a free system in a vector space V and a system of generators for V not to be ordered sets while on the other hand, the Definition 2.4.3 refers to a basis as an ordered set. This choice is motivated by the fact that it is more useful to consider the components of a vector on a given basis as an ordered array of scalars. For example, if $I = (v_1, v_2)$ is a basis for V, so it is $J = (v_2, v_1)$. But one considers I equivalent to J as systems of generators for V, not as bases.

Exercise 2.4.10 With $\mathcal{E} = (\mathbf{i}, \mathbf{j})$ and $\mathcal{E}' = (\mathbf{j}, \mathbf{i})$ two bases for $\mathcal{V}_O^2$, the vector $\mathbf{v} = 2\mathbf{i} + 3\mathbf{j}$ has the following components

$$\mathbf{v} = (2, 3)_{\mathcal{E}} = (3, 2)_{\mathcal{E}'}$$

when expressed with respect to them.

Remark 2.4.11 Consider the real vector space $\mathbb{R}^n$ and the vectors

$$\begin{aligned} e_1 &= (1, 0, \ldots, 0), \\ e_2 &= (0, 1, \ldots, 0), \\ &\vdots \\ e_n &= (0, 0, \ldots, 1). \end{aligned}$$

Since any element $v = (x_1, \ldots, x_n)$ can be uniquely written as

$$(x_1, \ldots, x_n) = x_1 e_1 + \cdots + x_n e_n,$$

the system $\mathcal{E} = (e_1, \ldots, e_n)$ is a basis for $\mathbb{R}^n$.

Definition 2.4.12 The system $\mathcal{E} = (e_1, \ldots, e_n)$ above is called the *canonical* basis for $\mathbb{R}^n$.

The canonical basis for $\mathbb{R}^2$ is $\mathcal{E} = (e_1, e_2)$, with $e_1 = (1, 0)$ and $e_2 = (0, 1)$; the canonical basis for $\mathbb{R}^3$ is $\mathcal{E} = (e_1, e_2, e_3)$, with $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ and $e_3 = (0, 0, 1)$.

Remark 2.4.13 We have meaningfully introduced the notion of a canonical basis for $\mathbb{R}^n$. Our analysis so far should nonetheless make it clear that for an arbitrary vector space V over $\mathbb{R}$ there is no canonical choice of a basis. The exercises that follow indeed show that some vector spaces have bases which appear more natural than others, in a sense.

Exercise 2.4.14 We refer to the Exercise 2.1.8 and consider $\mathbb{C}$ as a vector space over $\mathbb{R}$. As such it is generated by the two elements 1 and i since any complex number can be written as $z = a + \mathrm{i}b$, with $a, b \in \mathbb{R}$. Since the elements $1, \mathrm{i}$ are linearly independent over $\mathbb{R}$ they are a basis over $\mathbb{R}$ for $\mathbb{C}$.

As already seen in the Exercise 2.1.8, $\mathbb{C}^n$ is a vector space *both* over $\mathbb{C}$ *and* over $\mathbb{R}$. As a $\mathbb{C}$-vector space, $\mathbb{C}^n$ has canonical basis $\mathcal{E} = (e_1, \ldots, e_n)$, where the elements e_i are given as in the Remark 2.4.11. For example, the canonical basis for $\mathbb{C}^2$ is $\mathcal{E} = (e_1, e_2)$, with $e_1 = (1, 0)$, $e_2 = (0, 1)$.

As a real vector space, $\mathbb{C}^n$ has no canonical basis. It is natural to consider for it the following basis $\mathcal{B} = (b_1, c_1 \ldots, b_n, c_n)$, made of the $2n$ following elements,

$$\begin{aligned}
b_1 &= (1, 0, \ldots, 0), \quad c_1 = (\mathrm{i}, 0, \ldots, 0), \\
b_2 &= (0, 1, \ldots, 0), \quad c_2 = (0, \mathrm{i}, \ldots, 0), \\
&\vdots \\
b_n &= (0, 0, \ldots, 1), \quad c_n = (0, 0, \ldots, \mathrm{i}).
\end{aligned}$$

For $\mathbb{C}^2$ such a basis is $\mathcal{B} = (b_1, c_1, b_2, c_2)$, with $b_1 = (1, 0)$, $c_1 = (\mathrm{i}, 0)$, and $b_2 = (0, 1)$, $c_2 = (0, \mathrm{i})$.

Exercise 2.4.15 The real vector space $\mathbb{R}[x]_r$ has a natural basis given by all the monomials $(1, x, x^2, \ldots, x^r)$ with degree less than r, since any element $p(x) \in \mathbb{R}[x]_r$ can be written in a unique way as

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots a_r x^r,$$

with $a_i \in \mathbb{R}$.

Remark 2.4.16 We have seen in Chap. 1 that, by introducing a cartesian coordinate system in $\mathcal{V}_O^3$ and with the notion of components for the vectors, the vector space operations in $\mathcal{V}_O^3$ can be written in terms of operations among components. This fact is generalised in the following way.

Let $I = (v_1, \ldots, v_n)$ be a basis for V. Let $v, w \in V$, with $v = (\lambda_1, \ldots, \lambda_n)_I$ and $w = (\mu_1, \ldots, \mu_n)_I$ the corresponding components with respect to I. We compute the components, with respect to I, of the vectors $v + w$. We have

$$\begin{aligned}
v + w &= (\lambda_1 v_1 + \cdots + \lambda_n v_n) + (\mu_1 v_1 + \cdots + \mu_n v_n) \\
&= (\lambda_1 + \mu_1) v_1 + \cdots + (\lambda_n + \mu_n) v_n,
\end{aligned}$$

so we can write

$$v + w = (\lambda_1 + \mu_1, \ldots, \lambda_n + \mu_n)_I.$$

Next, with $a \in \mathbb{R}$ we also have

$$av = a(\lambda_1 v_1 + \cdots + \lambda_n v_n) = (a\lambda_1)v_1 + \cdots + (a\lambda_n)v_n,$$

so we can write

$$av = (a\lambda_1, \ldots, a\lambda_n)_I.$$

If $z = av + bw$ with, $z = (\xi_1, \ldots, \xi_n)_I$, it is immediate to see that

$$(\xi_1, \ldots, \xi_n)_I = (a\lambda_1 + b\mu_1, \ldots, a\lambda_n + b\mu_n)_I$$

or equivalently

$$\xi_i = a\lambda_i + b\mu_i, \qquad \text{for any} \quad i = 1, \ldots, n.$$

Proposition 2.4.17 *Let V be a vector space over $\mathbb{R}$, and $I = (v_1, \ldots, v_n)$ a basis for V. Consider a system*

$$w_1 = (\lambda_{11}, \ldots, \lambda_{1n})_I, \quad w_2 = (\lambda_{21}, \ldots, \lambda_{2n})_I, \quad \ldots \quad , \quad w_s = (\lambda_{s1}, \ldots, \lambda_{sn})_I$$

of vectors in V, and denote $z = (\xi_1, \ldots, \xi_n)_I$. One has that

$$z = a_1 w_1 + \cdots + a_s w_s \qquad \Longleftrightarrow \qquad \xi_i = a_1\lambda_{1i} + \cdots + a_s\lambda_{si} \;\; \textit{for any} \;\; i = 1, \ldots, n.$$

The i-th component of the linear combination z of the vectors w_k, is given by the same linear combination of the i-th components of the vectors w_k.

Proof It comes as a direct generalisation of the previous remark. □

Corollary 2.4.18 *With the same notations as before, one has that*

(a) the vectors $w_1, \ldots, w_s$ are linearly independent in V if and only if the corresponding n-tuples of components $(\lambda_{11}, \ldots, \lambda_{1n}), \ldots, (\lambda_{s1}, \ldots, \lambda_{sn})$ are linearly independent in $\mathbb{R}^n$,
(b) the vectors $w_1, \ldots, w_s$ form a system of generators for V if and only if the corresponding n-tuples of components $(\lambda_{11}, \ldots, \lambda_{1n}), \ldots, (\lambda_{s1}, \ldots, \lambda_{sn})$ generate $\mathbb{R}^n$.

A free system can be completed to a basis for a given vector space.

Theorem 2.4.19 *Let V be a finitely generated real vector space. Any free finite system is contained in a basis for V.*

Proof Let $I = \{v_1, \ldots, v_s\}$ be a free system for the real vector space V. By the Corollary 2.4.4, V has a basis, that we denote $\mathcal{B} = (e_1, \ldots, e_n)$. The set $I \cup \mathcal{B} = \{v_1, \ldots, v_s, e_1, \ldots, e_n\}$ obviously generates V. By applying the procedure given in the Theorem 2.4.2, the first s vectors are not deleted, since they are linearly independent by hypothesis; the subsystem $\mathcal{B}'$ one ends up with at the end of the procedure will then be a basis for V that contains I. □

2.5 The Dimension of a Vector Space

The following (somewhat intuitive) result is given without proof.

Theorem 2.5.1 *Let V be a vector space over $\mathbb{R}$ with a basis made of n elements. Then,*

(i) any free system I in V contains at most *n elements,*
(ii) any system of generators for V has at least *n elements,*
(iii) any basis for V has n elements.

This theorem makes sure that the following definition is consistent.

Definition 2.5.2 If there exists a positive integer $n > 0$, such that the real vector space V has a basis with n elements, we say that V has *dimension* n, and write $\dim V = n$. If V is not finitely generated we set $\dim V = \infty$. If $V = \{0_V\}$ we set $\dim V = 0$.

Exercise 2.5.3 Following what we have extensively described above, it is clear that $\dim \mathcal{V}^2_O = 2$ and $\dim \mathcal{V}^3_O = 3$. Also $\dim \mathbb{R}^n = n$, with $\dim \mathbb{R} = 1$, and we have that $\dim \mathbb{R}[x] = \infty$ while $\dim \mathbb{R}[x]_r = r + 1$. Referring to the Exercise 2.4.14, one has that $\dim_{\mathbb{C}} \mathbb{C}^n = n$ while $\dim_{\mathbb{R}} \mathbb{C}^n = 2n$.

We omit the proof of the following results.

Proposition 2.5.4 *Let V be a n-dimensional vector space, and W a vector subspace of V. Then,* $\dim(W) \leq n$*, while* $\dim(W) = n$ *if and only if $W = V$.*

Corollary 2.5.5 *Let V be a n-dimensional vector space, and $v_1, \dots, v_n \in V$. The following facts are equivalent:*

(i) the system $(v_1, \dots, v_n)$ is a basis for V,
(ii) the system $\{v_1, \dots, v_n\}$ is free,
(iii) the system $\{v_1, \dots, v_n\}$ generates V.

Theorem 2.5.6 (Grassmann) *Let V a finite dimensional vector space, with U and W two vector subspaces of V. It holds that*

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W).$$

Proof Denote $r = \dim(U)$, $s = \dim(W)$ and $p = \dim(U \cap W)$. We need to show that $U + W$ has a basis with $r + s - p$ elements.

Let $(v_1, \dots, v_p)$ be a basis for $U \cap W$. By the Theorem 2.4.19 such a free system can be completed to a basis $(v_1, \dots, v_p, u_1, \dots, u_{r-p})$ for U and to a basis $(v_1, \dots, v_p, w_1, \dots, w_{s-p})$ for W.

We then show that $I = (v_1, \dots, v_p, u_1, \dots, u_{r-p}, w_1, \dots, w_{s-p})$ is a basis for the vector space $U + W$. Since any vector in $U + W$ has the form $u + w$, with $u \in U$ and $w \in W$, and since u is a linear combination of $v_1, \dots, v_p, u_1, \dots, u_{r-p}$, while w

is a linear combination of $v_1, \ldots, v_p, w_1, \ldots, w_{s-p}$, the system I generates $U + W$. Next, consider the combination

$$\alpha_1 v_1 + \cdots + \alpha_p v_p + \beta_1 u_1 + \cdots + \beta_{r-p} u_{r-p} + \gamma_1 w_1 + \cdots + \gamma_{s-p} w_{s-p} = 0_V.$$

Denoting for brevity $v = \sum_{i=1}^{p} \alpha_i v_i, u = \sum_{j=1}^{r-p} \beta_j u_j$ and $w = \sum_{k=1}^{s-p} \gamma_k w_k$, we write this equality as

$$v + u + w = 0_V,$$

with $v \in U \cap W,\ u \in U,\ w \in W$. Since $v, u \in U$, then $w = -v - u \in U$; so $w \in U \cap W$. This implies

$$w = \gamma_1 w_1 + \cdots + \gamma_{s-p} w_{s-p} = \lambda_1 v_1 + \cdots + \lambda_p v_p$$

for suitable scalars λ_i: in fact we know that $\{v_1, \ldots, v_p, w_1, \ldots, w_{s-p}\}$ is a free system, so any γ_k must be zero. We need then to prove that, from

$$\alpha_1 v_1 + \cdots + \alpha_p v_p + \beta_1 u_1 + \cdots + \beta_{r-p} u_{r-p} = 0_V$$

it follows that all the coefficients α_i and β_j are zero. This is true, since $(v_1, \ldots, v_p, u_1, \ldots, u_{r-p})$ is a basis for U. Thus I is a free system. □

Corollary 2.5.7 *Let W_1 and W_2 be vector subspaces of V. If $W_1 \oplus W_2$ can be defined, then*

$$\dim(W_1 \oplus W_2) = \dim(W_1) + \dim(W_2).$$

Also, if $\mathcal{B}_1 = (w'_1, \ldots, w'_s)$ and $\mathcal{B}_2 = (w''_1, \ldots, w''_r)$ are basis for W_1 and W_2 respectively, a basis for $W_1 \oplus W_2$ is given by $\mathcal{B} = (w'_1, \ldots, w'_s, w''_1, \ldots, w''_r)$.

Proof By the Grassmann theorem, one has

$$\dim(W_1 + W_2) + \dim(W_1 \cap W_2) = \dim(W_1) + \dim(W_2)$$

and from the Definition 2.2.12 we also have $\dim(W_1 \cap W_2) = 0$, which gives the first claim.

With the basis $\mathcal{B}_1$ and $\mathcal{B}_2$ one considers $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ which obviously generates $W_1 \oplus W_2$. The second claim directly follows from the Corollary 2.5.5. □

The following proposition is a direct generalization.

Proposition 2.5.8 *Let $W_1, \ldots, W_n$ be subspaces of a real vector space V and let the direct sum $W_1 \oplus \cdots \oplus W_n$ be defined. One has that*

$$\dim(W_1 \oplus \cdots \oplus W_n) = \dim(W_1) + \cdots + \dim(W_n).$$