## **Vector Spaces**

Linear algebra is the theory of *vector spaces* (or *linear spaces* - these are synonyms). It was developed as an abstract theory of vectors as we know them from physics, and also as a tool for systematic studies of systems of linear equations.

**Definition 1.1.** Suppose  $(\mathbf{F}, +, \times)$  is a field with 0 and 1 the identity elements for + and  $\times$  respectively, and V is a nonempty set. Suppose there is an operation  $\oplus$  defined on V, and for each  $p \in \mathbf{F}$  and for each  $v \in V$  there exists a unique element  $p \cdot v \in V$  called the multiple (or product) of v by p. If

- (i)  $(V, \oplus)$  is an abelian group,
- (ii)  $(\forall p \in \mathbf{F})(\forall u, v \in V) p \cdot (u \oplus v) = (p \cdot u) \oplus (p \cdot v)$

(iii)  $(\forall p,q \in \mathbf{F})(\forall v \in V) (q+p) \cdot v = (q \cdot v) \oplus (p \cdot v)$ 

- (iv)  $(\forall p,q \in \mathbf{F})(\forall v \in V) (q \times p) \cdot v = q \cdot (p \cdot v)$
- (v)  $(\forall v \in V) \ 1 \cdot v = v$

then the set V with the operations  $\oplus$  and  $\cdot$  is called a *vector space* (or a *linear space*) over **F**.

## **Remarks**

- A vector space is not an algebra in the usual sense, as  $\cdot$  is not an operation on V.
- Elements of V are called "vectors" and those of F "scalars". In this approach, the frequently asked question "but what actually IS a vector?" is meaningless. Absolutely everything, from a set to an elephant, can be a "vector" if somebody defines the right operations on the right sets. One rather "serves" as a vector than "is" a vector. In other words being a vector is more an occupation than an attribute. Presented with an object and asked "is this a vector?" the right answer is "in what vector space?"
- We usually ignore both symbols  $\cdot$  and  $\times$  and we write (qp)v instead of (q×p)·v. The meaning of the expression is usually clear from the context.
- Traditionally, we use the + symbol for both additions of vectors and of scalars. Under these conventions part (iii) looks like (q+p)v=qv+pv.
- We will use the 0 (zero) symbol for the identity element of scalar addition and Θ (theta) for the zero vector, i.e. the identity element of vector addition.

**Example 1.1.** Let  $(\mathbf{F},+,\cdot)$  be a field. We denote by  $\mathbf{F}^n$  the set  $\mathbf{F} \times \mathbf{F} \times \ldots \times \mathbf{F}$  of all n-tuples of elements of  $\mathbf{F}$ . For every two n-tuples  $(a_1,a_2, \ldots, a_n)$  and  $(b_1,b_2, \ldots, b_n)$  and every scalar  $p \in \mathbf{F}$  we define  $(a_1,a_2, \ldots, a_n)+(b_1,b_2, \ldots, b_n) = (a_1+b_1,a_2+b_2, \ldots, a_n+b_n)$  and  $p(a_1,a_2, \ldots, a_n)=(pa_1,pa_2, \ldots, pa_n)$ . These operations are known as *componentwise* addition and scalar multiplication.  $\mathbf{F}^n$  is a vector space over  $\mathbf{F}$  with respect to these two operations. In particular, when n=1 we get that every field is a vector space over itself.

**Example 1.2.**  $2^{X}$  is a vector space over the field  $\mathbb{Z}_{2}$  with respect to the symmetric difference as vector addition and  $0 \cdot A = \emptyset$  and  $1 \cdot A = A$  as the scalar multiplication. Here vectors are sets and scalars are the numbers 0 and 1.

**Example 1.3.**  $\mathbf{R}^{\mathbf{R}}$  is a vector space over  $\mathbf{R}$  with respect to ordinary addition of functions and multiplications of a function by a constant. Here vectors are functions and scalars are real numbers.

**Example 1.4.** For any set X and any field  $\mathbf{F}$ ,  $\mathbf{F}^{X}$  is a vector space with respect to ordinary function addition and scaling. See example ??? in chapter 2.

**Example 1.5. R** over **Q** with ordinary addition and multiplication. Here vectors are real numbers and scalars are rational numbers.

**Example 1.6. Q** over **R** is NOT a vector space. Here the would-be vectors are rational numbers and scalars are real numbers and the product of a rational number by a real number may an irrational number.

**Example 1.7.**  $V = \{(a_1, a_2, a_3, a_4) \in \mathbb{R}^4: a_1+a_2+2a_3-a_4=0 \text{ and } 3a_1-2a_2+a_3+a_4=0\}$  is a vector space over  $\mathbb{R}$  with respect to componentwise addition and multiplication. Let  $\mathbf{x}=(x_1, x_2, x_3, x_4)$  and  $\mathbf{y}=(y_1, y_2, y_3, y_4)$  be vectors from V. We must show that both  $\mathbf{x}+\mathbf{y}$  and  $\mathbf{p}\mathbf{y}$  belong to V.  $\mathbf{x}+\mathbf{y}=(x_1, x_2, x_3, x_4) + (y_1, y_2, y_3, y_4) = (x_1+y_1, x_2+y_2, x_3+y_3, x_4+y_4)$ . This vector belongs to v if and only if  $(x_1+y_1) + (x_2+y_2) + 2(x_3+y_3) - (x_4+y_4) = 0$  and  $3(x_1+y_1) - 2(x_2+y_2) + (x_3+y_3) + (x_4+y_4) = 0$ . The first equality is equivalent to  $(x_1+x_2+2x_3-x_4) + (y_1+y_2+2y_3-y_4)=0$  which is true because  $\mathbf{x}, \mathbf{y} \in V$ . The second one is equivalent to  $(3x_1+-2x_2+x_3+x_4)+(3y_1-2y_2+y_3+y_4)=0$  which is true for the same reason. A similar argument shows that  $\mathbf{p}\mathbf{v} \in V$ . The zero vector (0,0,0,0) and the inverse to any vector  $\mathbf{x}$  from V belong to V for obvious reasons. The remaining conditions hold because V is a subset of  $\mathbb{R}^4$ .

**Definition 1.2.** Suppose V is a vector space over **F**. Let W be a subset of V. If W is a vector space over **F** (with respect to the same vector addition and scalar multiplication then W is called a *subspace* of V.

Proposition 1.1. Let V be a vector space over F. Then

(a) 
$$(\forall v \in V) \ 0v = \Theta$$

(b) 
$$(\forall p \in \mathbf{F}) p \Theta = \Theta$$

(c)  $(\forall p \in \mathbf{F})(\forall v \in V) (-p)v = p(-v) = -(pv)$ 

**<u>Proof.</u>** (a) 0v=(0+0)v=0v+0v. Hence, by the cancellation law, we have  $0v=\Theta$ .

(b)  $p\Theta = p(\Theta + \Theta) = p\Theta + p\Theta$  and we use the cancellation law again.

(c) To prove that (-p)v=-pv it is enough to notice that  $(-p)v+pv=(-p+p)v=0v=\Theta$  by (a). To prove that p(-v)=-(pv) we write  $p(-v)+pv=p(-v+v)=p\Theta=\Theta$  by(b).

Theorem 1.1 A nonempty subset W of V is a subspace of V iff

- (a)  $(\forall u, w \in W) u + w \in W$
- (b)  $(\forall p \in \mathbf{F})(\forall u \in W) pu \in W$

**<u>Proof.</u>** " $\Rightarrow$ " is obvious because a subspace must be closed under both operations.

"⇐" It is enough to show that W is subgroup of V. The conditions (ii) – (v) are obviously inherited by every subset of V. Condition (a) guarantees that W is an algebra, associativity and commutativity of the vector addition are inherited from V. Since W is nonempty we can chose any w∈ W and putting p=0 in (b) we obtain  $0w \in W$ , which, thanks to Proposition 1.1 means  $\Theta \in W$ . Moreover, for every w∈ W, its inverse,  $-w \in W$  because, by Proposition 1.1, -w = (-1)w, and by the condition (b) of the present theorem with p=-1, (-1)w∈ W.□

**Example 1.8.** The set of all differentiable functions is a subspace in  $\mathbf{R}^{\mathbf{R}}$ .

**Example 1.9.** The set of all finite subsets of **N** is a subspace in the space  $2^{N}$  defined in Example 1.2. Obviously, the symmetric difference of two finite sets is finite,  $0A=\emptyset$  is finite and 1A=A is finite whenever A is finite.

**Example 1.10.**  $W = \{(x,y) \in \mathbb{R}^2 : x \ge 0\}$ . W is not a subspace of  $\mathbb{R}^2$ . It is closed under vector addition but is not closed under scalar multiplication,  $(-1)(1,1) = (-1,-1) \notin W$ .

**Example 1.11.** W={ $(x,y,z) \in \mathbb{R}^3$ :  $x^2-y^2=0$ }. W is not a subspace of  $\mathbb{R}^3$ . It is closed under scalar multiplication, but not under vector addition, for example  $(1,1,1), (-1,1,1) \in \mathbb{W}$  and  $(1,1,1) + (-1,1,1)=(0,2,2) \notin \mathbb{W}$ .

Recall that the relation of inclusion is a partial order on any family of sets, in particular on any family of subspaces of a vector space.

**Definition 1.3.** Let  $S \subseteq V$  (a subset, not necessarily a subspace). Then by *span(S)* we will denote the smallest subspace of V containing S. We call span(S) the *subspace spanned by S*. **Proposition 1.2.** Let V(S) denote the set of all subspaces of V containing S. Then

$$span(S) = \bigcap V(S)$$

**<u>Proof.</u>** It is enough to show that  $\bigcap V(S)$  is a subspace of V and that is an easy consequence of Theorem 1.1

**Definition 1.4.** Given vectors  $v_1, v_2, \ldots, v_n$  and scalars  $a_1, a_2, \ldots, a_n$ , the sum  $a_1v_1+a_2v_2+\ldots+a_nv_n$  is called the *linear combination* of vectors  $v_1, v_2, \ldots, v_n$  with coefficients  $a_1, a_2, \ldots, a_n$ .

**Example 1.12.** The vector (3,0,4) from  $\mathbf{R}^3$  is a linear combination of (1,2,2),(2,1,3) and (0,3,1) with coefficients 1,1 and -1. Indeed 1(1,2,2)+1(2,1,3)-1(0,3,1)=(1+2-0,2+1-3,2+3-1)=(3,0,4). Notice that we may also write (3,0,4)=-1(1,2,2)+2(2,1,3)+0(0,3,1) or (3,0,4)=-3(1,2,2)+3(2,1,3) +1(0,3,1). The lesson from this example is that a vector may be represented in various ways as a linear combination of given vectors.

**Definition 1.5.** For every set  $S \subseteq V$ , *lin*(*S*) will denote the set of all linear combinations of vectors from S.

**Example 1.13.**  $\mathbf{R}_n[x] = \lim\{1, x, x^2, ..., x^n\}$ **Example 1.14.**  $\mathbf{Q}(\sqrt{2}) = \lim(\{1, \sqrt{2}\})$ . Here we treat  $\mathbf{Q}(\sqrt{2})$  as a vector space over the field of rational numbers  $\mathbf{Q}$ .

**Proposition 1.3.** For every set  $S \subseteq V$ , lin(S) is a subspace of V.

**Proof.** Suppose w,u $\in$  span(S) and  $p\in \mathbf{F}$ . Then, for some scalars  $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_k$  and for some vectors  $v_1, v_2, \ldots, v_n, x_1, x_2, \ldots, x_k$ , all coming from V, we have  $u = a_1v_1 + a_2v_2 + \ldots + a_nv_n$  and  $w = b_1x_1 + b_2x_2 + \ldots + b_kx_k$ . Now  $u + w = a_1v_1 + a_2v_2 + \ldots + a_nv_n + b_1x_1 + b_2x_2 + \ldots + b_kx_k$  and  $pu = pa_1v_1 + pa_2v_2 + \ldots + pa_nv_n$ , which means that u + w and pu belong to lin(S), hence, by Theorem 1.1 lin(S) is a subspace of V. $\Box$ 

**Theorem 1.2** For every nonempty subset S of a space V lin(S)=span(S).

**Proof**. We must show that  $lin(S) \subseteq span(S)$  and  $span(S) \subseteq lin(S)$ . In order to show the first inclusion it is enough to notice that lin(S) is a subset of every subspace of V that contains S. This is obvious thanks to Theorem 1.1. To prove the second inclusion we point out that lin(S) contains S – which is obvious and, that lin(S) is a subspace of V by Proposition 1.3. Since span(S) is contained in every subspace of V that contains S, span(S) is contained in lin(S).

One of the central concepts of linear algebra is that of linear independence of sets of vectors.

**Definition 1.6.** A finite set of vectors  $\{v_1, v_2, ..., v_n\}$  from a vector space V is said to be *linearly independent* iff

 $(\forall a_1, a_2, \dots, a_n \in \mathbf{F}) [a_1v_1 + a_2v_2 + \dots + a_nv_n = \Theta \Longrightarrow a_1 = a_2 = \dots = a_n = 0]$ 

A set of vectors that is not linearly independent is called (surprise! surprise!) linearly dependent.

Notice that the linear independence is a property of sets of vectors, rather than vectors. Whenever we say *vectors* x,y,z *are linearly independent* what we really mean is *The set* {x,y,z} *is linearly independent*.

**Example 1.15.** The set  $\{(1,0,0,\ldots,0), (0,1,0,\ldots,0), \ldots, (0,0,\ldots,0,1)\}$  is a linearly independent subset of  $\mathbf{F}^{n}$ .

**Example 1.16.**  $\{1, x, x^2, ..., x^n\}$  is a linearly independent subset of  $\mathbf{R}_n[x]$ .

A very important property of any linearly independent set of vectors S is that every vector from span(S) is uniquely represented as a linear combination of vectors from S. More precisely we have

**Theorem 1.3** A set { $v_1, v_2, ..., v_n$ } of vectors from V is linearly independent iff  $(\forall v \in V)(\forall a_1, a_2, ..., a_n, b_1, b_2, ..., b_n \in \mathbf{F})$  [ $v=a_1v_1+a_2v_2+ ... +a_nv_n = b_1v_1+b_2v_2+ ... +b_nv_n \Rightarrow (\forall i) a_i=b_i$ ] **Proof.** ( $\Rightarrow$ ) Suppose  $v=a_1v_1+a_2v_2+ ... +a_nv_n$  and  $v=b_1v_1+b_2v_2+ ... +b_nv_n$ . Subtracting these equations side from side, we obtain  $\Theta = (a_1-b_1)v_1+(a_2-b_2)v_2+ ... +(a_n-b_n)v_n$ . Since the set { $v_1, v_2, ..., v_n$ } is linearly independent we have  $a_i-b_i=0$ , which means  $a_i=b_i$ , for each i=1,2,...,n. ( $\Leftarrow$ ) Let us take  $v=\Theta$ . One way of writing  $\Theta$  as a linear combination of vectors  $v_1, v_2, ..., v_n$ , is to put  $b_1=b_2=...=b_n=0$ , i.e.  $\Theta=0v_1+0v_2+...+0v_n$ , so our condition implies that whenever  $\Theta=a_1v_1+a_2v_2+...+a_nv_n$  we have  $a_i=0$  for each i=1,2,...,n.

What Theorem 1.3 really says is that if a vector can be represented as a linear combination of linearly independent vectors then the coefficients are unique and vice versa.

**Theorem 1.4** The set  $S = \{v_1, v_2, ..., v_n\}$  is linearly independent iff none of the vectors from S is a linear combination of the others.

**Proof.** ( $\Rightarrow$ ) Suppose one of the vectors is a linear combination of the others. Without loss of generality we may assume that  $v_n$  is the one, i.e.  $v_n = a_1v_1 + a_2v_2 + \dots + a_{n-1}v_{n-1}$ . Then we may write  $\Theta = a_1v_1 + a_2v_2 + \dots + a_{n-1}v_{n-1} + (-1)v_n$ . Since  $(-1) \neq 0$  the set  $\{v_1, v_2, \dots, v_n\}$  is linearly dependent. ( $\Leftarrow$ ) Suppose now that  $\{v_1, v_2, \dots, v_n\}$  is linearly dependent, i.e. there exist coefficients  $a_1, a_2, \dots$ ,  $a_n$ , not all of them zeroes, such that  $\Theta = a_1v_1 + a_2v_2 + \dots + a_nv_n$ . Again, without losing generality, we may assume that  $a_n \neq 0$  (we can always renumber the vectors so that the one with nonzero coefficient is the last). Since nonzero scalars are invertible, we have  $v_n = (-a_1a_n^{-1})v_1 + (-a_2a_n^{-1})v_2 + \dots + (-a_{n-1}a_n^{-1})v_{n-1}\Box$ 

**Theorem 1.5** If  $v \in \text{span}(v_1, v_2, \dots, v_n)$  then  $\text{span}(v_1, v_2, \dots, v_n) = \text{span}(v_1, v_2, \dots, v_n, v)$ **Proof.** Obviously  $\text{span}(v_1, v_2, \dots, v_n) \subseteq \text{span}(v_1, v_2, \dots, v_n, v)$ , even if v is not a linear combination of  $v_1, v_2, \dots, and v_n$ . Suppose now that  $w \in \text{span}(v_1, v_2, \dots, v_n, v)$ . This means that for some scalars  $p_1, p_2, \dots, p_n, p$  we have  $w = p_1v_1 + p_2v_2 + \dots + p_nv_n + pv$ . Since  $v \in \text{span}(v_1, v_2, \dots, v_n)$ , there exist scalars  $r_1, r_2, \dots, r_n$  such, that  $v = r_1v_1 + r_2v_2 + \dots + r_nv_n$ . Hence  $w = p_1v_1 + p_2v_2 + \dots + p_nv_n + p(r_1v_1 + r_2v_2 + \dots + r_nv_n) = (p_1 + pr_1)v_1 + (p_2 + pr_2)v_2 + \dots + (p_n + pr_n)v_n$  which means  $w \in \text{span}(v_1, v_2, \dots, v_n)$ .

The concepts of linear independence and spanning meet in the definition of a basis of a vector space.

Definition 1.7. The set S={v<sub>1</sub>,v<sub>2</sub>, ...,v<sub>n</sub>} is a basis of a vector space V iff (1) span(S)=V, and
(2) S is linearly independent.

**Example 1.17.** The set  $S = \{(1,0,0, \dots, 0), (0,1,0, \dots, 0), \dots, (0,0, \dots, 0,1)\}$  is a basis for  $\mathbf{F}^n$ . This basis is called *the standard* or *the canonical* basis for  $\mathbf{F}^n$ .

**Example 1.18.**  $S = \{1, x, x^2, ..., x^n\}$  is a basis of  $\mathbf{R}_n[x]$ . Obviously,  $span(\{1, x, x^2, ..., x^n\}) = \mathbf{R}_n[x]$  and in Example 1.16 we explained why S is linearly independent.

**<u>Theorem 1.6</u>** S={ $v_1, v_2, ..., v_n$ } is a basis of a vector space V iff for every  $v \in V$  there exist unique scalars  $p_1, p_2, ..., p_n$  such that  $v=p_1v_1+p_2v_2+...+p_nv_n$ .

**<u>Proof.</u>** ( $\Rightarrow$ ) Since V=span(S), all we have to show is that scalars p<sub>1</sub>,p<sub>2</sub>, ..., p<sub>n</sub> are unique, but this is guaranteed by Theorem 1.3.

( $\Leftarrow$ ) Obviously span(S)=V and the linear independence of S follows from Theorem 1.3

**Definition 1.8.** Let  $S = \{v_1, v_2, \dots, v_n\}$  be a basis of V and let  $v \in V$ . then the scalars  $p_1, p_2, \dots, p_n$  such that  $v = p_1v_1 + p_2v_2 + \dots + p_nv_n$  are called <u>coordinates of v with respect to S</u>. We will denote the sequence by  $[v]_S = [p_1, p_2, \dots, p_n]_S$ .

**Example 1.19.** The standard basis S for  $\mathbf{F}^n$  is special in that the coordinates of a vector  $(x_1, x_2, ..., x_n)$  are its "natural" coordinates  $x_1, x_2, ..., x_n$ , i.e.  $[(x_1, x_2, ..., x_n)]_S = [x_1, x_2, ..., x_n]_S$ . This is obvious as vectors  $v_1 = (1, 0, 0, ..., 0), v_2 = (0, 1, 0, ..., 0), ..., v_n = (0, 0, ..., 0, 1)$  constituting S are structured so that in every linear combination v of  $v_1, v_2, ..., v_n$  the i-th coordinate is equal to the i-th coefficient.

**Example 1.20.** The set  $S = \{(1,1,1), (1,1,0), (1,0,0)\}$  is a basis for  $\mathbb{R}^3$ . It is easy to check that  $[(1,0,0)]_S = [0,0,1], [(1,2,3)]_S = [3,-1,-1]$  and [(2,3,2)] = [2,1,-1].

According to our definition, a basis is a finite set of vectors. Not every vector space contains a finite basis (for example, there is no finite basis for  $\mathbf{R}$  over  $\mathbf{Q}$ ). In this course, though, we will only consider those vector spaces who do posses finite bases. Such vector spaces are called <u>finite dimensional</u>. This term is justified by the following definition.

**Definition 1.9.** Let S be a basis for a vector space V. Then |S| is called the <u>dimension</u> of V. We write dim(V)=|S|.

A vector space can obviously have many bases, so a natural question arises, *can a vector space have more then one dimension?* Fortunately, the answer is NO, as follows from the next theorem.

**Theorem 1.7** In a finite dimensional vector space V every two bases have the same number of vectors.

The proof of this theorem makes heavy use of the following lemma

Lemma 1.1. (Replacement lemma, known also as Steinitz lemma)

Suppose the set  $\{v_1, v_2, ..., v_n\}$  is linearly independent and span $\{w_1, w_2, ..., w_k\}=V$ . Then n≤k and some n out of the k vectors in  $\{w_1, w_2, ..., w_k\}$  can be replaced be  $v_1, v_2, ..., and v_n$ . To be more precise, there exist subscripts  $i_1, i_2, ..., i_{k-n}$  such that the set  $\{w_{i_1}, w_{i_2}, ..., w_{i_{k-n}}, v_1, v_2, ..., v_n\}$  is a spanning set for V.

**Proof** We will prove the lemma by induction on n. For n=1 we have a linearly independent set  $\{v_1\}$  and a spanning set  $\{w_1, w_2, \dots, w_k\}$ . Since  $\{v_1\}$  is linearly independent, we have  $v_1 \neq \Theta$ . This implies that  $V \neq \{\Theta\}$  and every spanning set for V must be nonempty. Hence  $k \ge n$  (n=1). The set  $\{w_1, w_2, \dots, w_k\}$  spans V. This means that every vector from V, in particular  $v_1$  is a linear combination of  $w_1, w_2, \dots$ , and  $w_k$ , so there exist scalars  $p_1, p_2, \dots, p_k$  such that  $v_1 = p_1 w_1 + p_2 w_2 + \dots + p_k w_k$ . At least one  $p_i$  is different from 0 – otherwise  $v_1$  would be equal to  $\Theta$ . Without loss of generality we may assume that  $p_k \neq 0$ . This easily implies that  $w_k$  is a linear combination of  $v_1, w_1, w_2, \dots$ , and  $w_{k-1}$ . Hence, by Theorem 1.5, span $(v_1, w_1, w_2, \dots, w_{k-1}) = \text{span}(v_1, w_1, w_2, \dots, w_k) = \text{span}(w_1, w_2, \dots, w_k) = V$  and we are done.

Now, suppose the set  $\{v_1, v_2, \dots, v_{n+1}\}$  is linearly independent,  $\{w_1, w_2, \dots, w_k\}$  spans V, and the lemma is true for every n-element linearly independent set, in particular for  $\{v_1, v_2, \dots, v_n\}$ . Hence  $k \ge n$  and some n vectors from  $\{w_1, w_2, \dots, w_k\}$  can be replaced by  $v_1, v_2, \dots$ , and  $v_n$ spans V. Having done this, we wind up with a single element linearly independent set  $\{v_{n+1}\}$  and a spanning set  $\{w_{i_1}, w_{i_2}, \dots, w_{i_{k-n}}, v_1, v_2, \dots, v_n\}$ . Obviously, there exist scalars such that  $v_{n+1} =$  $=a_1w_{i_1} + a_2w_{i_2} + \dots + a_{k-n}w_{i_{k-n}} + b_1v_1 + b_2v_2 + \dots + b_nv_n$ . Moreover, at least one of the  $a_i$ 's, say  $a_1$ , is different from 0 – otherwise  $v_{n+1}$  would be a linear combination of  $v_1, v_2, ..., v_n$ . Now, like in the first part of the proof,  $w_i$  can be replace in the spanning set by  $v_{n+1}$ .

**Proof** of Theorem 1.7. Let  $B = \{v_1, v_2, \dots, v_n\}$  and  $D = \{w_1, w_2, \dots, w_k\}$  be two bases of V. B, being a basis, is linearly independent and D, being also a basis, is a spanning set for V. Hence, by the replacement lemma, we have  $n \le k$ . Now, if we reverse the roles of B and D, we will get  $k \le n$ .

Notice that the replacement lemma says that the size of any linearly independent subset of V is less than, or equal to the size of any spanning set for V.

The following theorem summarizes the most important properties of bases.

**Theorem 1.8** Suppose V is a vector space, dimV=n, n>0 and S $\subseteq$ V. Then

- (1) If |S|=n and S is linearly independent then S is a basis for V
- (2) If |S|=n and span(S)=V then S is a basis for V
- (3) If S is linearly independent then  $|S| \le n$
- (4) If span(S)=V then  $|S| \ge n$
- (5) If S is linearly independent then S is a subset of some basis of V
- (6) If S spans V then S contains a basis of V
- (7) S is a basis of V iff S is a maximal linearly independent subset of V
- (8) S is a basis of V iff S is a minimal spanning set for V

**<u>Proof</u>** (1) S is linearly independent, and a basis B of V is a spanning set for V. Since |S|=n=|B|, the replacement lemma says that all vectors from B can be replaced by all vectors from S and the result (which is S) is a spanning set for V.

(2) If S is linearly dependent then, by Theorem 1.5, S contains a proper subset S' such that span(S')=span(S)=V. But the replacement lemma says that no set containing fewer than n=|S|==dim(V) vectors spans V, hence S cannot be linearly dependent.

(3) and (4) follow immediately from the replacement lemma.

(5) Let |S|=k. By the replacement lemma k vectors from a basis B of V can be replaced by vectors from S in such a way that the resulting set is a spanning set of V. Since the resulting set consists of n vectors, by (2) it is a basis.

(6) There exist linearly independent subsets of S because there exist non-zero vectors in S (otherwise the space spanned by S would have dimension 0) and a single-element set with nonzero element is linearly independent. Let T be a linearly independent subset of S with the largest number of elements. Such a set T exists because every linearly independent subset of V

(hence also of S) has no more than n elements. Thus  $S \subseteq span(T)$ , hence  $span(S) \subseteq span(span(T))$ . Since span(S)=V and span(span(T))=span(T) we get that T is a linearly independent spanning set for V, i.e. a basis contained in S.

(7) If S is a basis for V then obviously S is linearly independent. Since span(S)=V, every vector v from V is a linear combination of vectors from S, so  $S \cup \{v\}$  is linearly dependent. Hence S is a maximal linearly independent subset of V. On the other hand, if S is a maximal linearly independent subset of V then, for every  $v \in V$ -S, adding v to S turns it into a linearly dependent subset of V, which means that  $v \in \text{span}(S)$ . Hence S is a basis for V.

(8) If S is a basis then, by (4) S is a minimal spanning set for V. On the other hand suppose S is a minimal spanning set. Then, by Theorem 1.5, S is linearly independent, i.e. S is a basis for V. $\Box$ 

<u>**Theorem 1.9**</u> If W is a subspace of a finite dimensional space V then dimW $\leq$ dimV and dimW=dimV if and only if W=V.

**Proof.** Any basis of W is a linearly independent subset of V, so, by(6) is a subset of a basis of V, hence dimW $\leq$ dimV. If dimW=dimV then, by Theorem 1.8(1), every basis for W is also a basis for V, so W=V as spaces spanned by the same set.