The span \( s(A) \) of a subset \( A \) in vector space \( V \) is a subspace of \( V \). This span may be the whole vector space \( V \) (we say then \( A \) spans \( V \)). In this paragraph, we study subsets \( A \) of \( V \) which span \( V \) and which are most economical in the sense that any proper subset of \( A \) spans a proper subspace of \( V \).

We begin with a definition that will be important for everything in the sequel.

**42.1 Definition:** Let \( V \) be a vector space over a field \( K \). A finite number of vectors \( v_1, v_2, \ldots, v_n \) in \( V \) are called linearly dependent over \( K \) if there are scalars \( \alpha_1, \alpha_2, \ldots, \alpha_n \) in \( K \), not all of them being zero, such that

\[
\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0
\]

(here 0 is the zero vector). If \( v_1, v_2, \ldots, v_n \) are not linearly dependent over \( K \), then \( v_1, v_2, \ldots, v_n \) are said to be linearly independent over \( K \).

A finite subset \( A \) of \( V \) is called linearly dependent (resp. linearly independent) over \( K \) if the finitely many vectors in \( A \) are linearly dependent (resp. linearly independent) over \( K \).

An infinite subset \( A \) of \( V \) is called linearly dependent over \( K \) if there is a finite subset of \( A \) which is linearly dependent over \( K \). An infinite subset \( A \) of \( V \) is called linearly independent over \( K \) if \( A \) is not linearly dependent over \( K \), i.e., \( A \) is called linearly independent over \( K \) if every finite subset of \( A \) is linearly independent over \( K \).

In place of the phrase "linearly (in)dependent over \( K \)" , we shall also use the expression "\( K \)-linearly (in)dependent". When the field of scalars is clear from the context, we drop the phrase "over \( K \)" or the prefix "\( K \)-".

According to our definition, the vectors \( v_1, v_2, \ldots, v_n \) of a vector space over \( K \) are linearly independent over \( K \) provided
\[ \alpha_1, \alpha_2, \ldots, \alpha_n \in K, \; \alpha_1 V_1 + \alpha_2 V_2 + \cdots + \alpha_n V_n = 0 \implies \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0. \]
That is to say, \( V_1, V_2, \ldots, V_n \) are \( K \)-linearly independent if the vector zero can be written as a linear combination of \( V_1, V_2, \ldots, V_n \) only in the trivial way where the scalars are zero.

42.2 Examples: (a) Let \( V \) be a vector space over a field \( K \) and let \( v \) be a nonzero vector in \( V \). Then \( \alpha v = 0 \) implies \( \alpha = 0 \) (Lemma 39.4(10)). Hence \( v \) (and \( \{ v \} \)) is linearly independent over \( K \). On the other hand, \( \{0\} \) is linearly dependent over \( K \) because \( 1.0 = 0 \) and \( 1 \neq 0 \).

(b) Consider the vector space \( \mathbb{Q}^3 \) over \( \mathbb{Q} \). The vectors \( u = (1,0,0), \; v = (0,1,0), \; w = (0,0,1) \) of \( \mathbb{Q}^3 \) are linearly independent over \( \mathbb{Q} \), for if \( \alpha, \beta, \gamma \in K \) and \( \alpha u + \beta v + \gamma w = 0 \), then
\[
\alpha(1,0,0) + \beta(0,1,0) + \gamma(0,0,1) = (0,0,0)
\]
\[
(\alpha,0,0) + (0,\beta,0) + (0,0,\gamma) = (0,0,0)
\]
\[
\alpha = \beta = \gamma = 0.
\]

(c) More generally, the vectors \( u_1 = (1,0,\ldots,0), \; u_2 = (0,1,\ldots,0), \; \ldots, u_n = (0,0,\ldots,1) \) in the vector space \( K^n \) over a field \( K \) are linearly independent over \( K \): if \( \alpha_1, \alpha_2, \ldots, \alpha_n \in K \) and \( \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n = 0 \), then
\[
\alpha_1(1,0,\ldots,0) + \alpha_2(0,1,\ldots,0) + \cdots + \alpha_n(0,0,\ldots,1) = (0,0,\ldots,0)
\]
\[
(\alpha_1,0,\ldots,0) + (0,\alpha_2,\ldots,0) + \cdots + (0,0,\ldots,\alpha_n) = (0,0,\ldots,0)
\]
\[
(\alpha_1,\alpha_2,\ldots,\alpha_n) = (0,0,\ldots,0)
\]
\[
\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.
\]
The reader is probably acquainted with the vectors \( u_1, u_2, u_3 \) in the vector space \( \mathbb{R}^3 \) over \( \mathbb{R} \) under the names \( \vec{i}, \vec{j}, \vec{k} \).

(d) The vectors (1,0) and (-1,0) in the \( \mathbb{R} \)-vector space \( \mathbb{R}^2 \) are linearly dependent over \( \mathbb{R} \) because \( 1 \neq 0 \) in \( \mathbb{R} \) and \( 1(1,0) + 1(-1,0) = (0,0) = \text{zero vector in } \mathbb{R}^2 \).

(e) Let \( V \) be a vector space over a field \( K \) and let \( V_1, V_2, \ldots, V_n \) be vectors in \( V \) which are linearly independent over \( K \). Then any nonempty subset of \( \{ V_1, V_2, \ldots, V_n \} \) is linearly independent over \( K \). In fact, if, say, \( V_1, V_2, \ldots, V_m \) are linearly dependent over \( K (m \leq n) \), then there are scalars \( \alpha_1, \alpha_2, \ldots, \alpha_m \) in \( K \), not all equal to zero, such that
\[
\alpha_1 V_1 + \alpha_2 V_2 + \cdots + \alpha_m V_m = 0;
\]
hence, when we put (in case $m < n$) $\alpha_{m+1} = \cdots = \alpha_n = 0$, we obtain

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m + \alpha_{m+1} v_{m+1} + \cdots + \alpha_n v_n = 0,$$

where not all of $\alpha_1, \alpha_2, \ldots, \alpha_m, \alpha_{m+1}, \ldots, \alpha_n$ are equal to zero, contradicting the assumption that $v_1, v_2, \ldots, v_n$ are linearly independent over $K$. Thus any nonempty subset of a linearly independent finite set of vectors is linearly independent. But this statement is true also for infinite linearly independent sets. Indeed, let $A$ be an infinite linearly independent subset of $V$ and let $B$ be a nonempty subset of $A$. If $B$ is finite, then $B$ is linearly independent by definition. If $B$ is infinite, then any finite subset of $B$, being a finite subset of $A$, is linearly independent over $K$ and hence $B$ itself is linearly independent over $K$. Thus we have shown that every nonempty subset of a linearly independent set of vectors is linearly independent. Equivalently, any set of vectors containing a linearly dependent subset is linearly dependent.

(f) Let $V$ be a vector space over a field $K$ and let $A$ be a subset of $V$ containing $0 \in V$. Then $A$ is linearly dependent over $K$ by Example 42.2(a) and Example 42.2(e). Alternatively, just choose a finite number of vectors $v_1, v_2, \ldots, v_n$ from $A$ including $0$, say $v_1 = 0$ and observe that

$$1v_1 + 0v_2 + \cdots + 0v_n = 0,$$

so that $v_1, v_2, \ldots, v_n$ are linearly dependent over $K$ and consequently $A$, too, is linearly dependent over $K$.

(g) Let $V$ be the vector space $\mathbb{C}^2$ over $\mathbb{C}$. The vectors $(1,0), (-i,0)$ in $V$ are linearly dependent over $\mathbb{C}$, because

$$i(1,0) + 1(-i,0) = (0,0) = \text{zero vector in } V.$$

However, when $V$ is regarded as an $\mathbb{R}$-vector space, these two vectors are not linearly dependent: if $\alpha, \beta \in \mathbb{R}$ and $\alpha(1,0) + \beta(-i,0) = (0,0)$, then $(\alpha - \beta i, 0) = (0,0)$, hence the complex number $\alpha - \beta i$ is equal to $0$, so $\alpha = \beta = 0$. Thus $(1,0), (-i,0)$ are linearly dependent over $\mathbb{C}$, but linearly independent over $\mathbb{R}$. This example shows that the field of scalars must be specified (unless it is clear from the context) whenever one discusses linear (in)dependence of vectors.

(h) Let $V$ be a vector space over a field $K$ and let $v_1, v_2, \ldots$ be infinitely many vectors in $V$. The linear dependence of $v_1, v_2, \ldots$ does not mean that there are scalars $\alpha_1, \alpha_2, \ldots$, not all equal to zero, such that

$$\sum_{k=1}^{n} \alpha_k v_k = 0.$$
This equation is meaningless, for its left hand side is not defined. What is defined (Definition 8.4) is a sum \( \sum_{k=1}^{n} \alpha_k v_k \) of a finite number \( n \) of vectors \( v_1, v_2, \ldots, v_n \) in \( V \). The definition of \( \sum_{k=1}^{n} \alpha_k v_k \) would involve some limiting process, and this is not possible in an arbitrary vector space.

(i) Consider the vector space \( C^1([0,1]) \) over \( \mathbb{R} \) (Example 40.6(j)). The functions \( f: [0,1] \to \mathbb{R} \) and \( g: [0,1] \to \mathbb{R} \), where \( f(x) = e^x \) and \( g(x) = e^{2x} \) for all \( x \) in \( [0,1] \), are vectors in \( C^1([0,1]) \). We claim that \( f \) and \( g \) are linearly independent over \( \mathbb{R} \). To prove this, let us assume \( \alpha, \beta \in \mathbb{R} \) and \( \alpha f + \beta g = \) zero vector in \( C^1([0,1]) \). The zero vector in \( C^1([0,1]) \) is the function \( z: [0,1] \to \mathbb{R} \) such that \( z(x) = 0 \) for all \( x \) in \( [0,1] \). Hence

\[
(\alpha f + \beta g)(x) = 0 \quad \text{for all } x \in [0,1],
\]

\[
\alpha f(x) + \beta g(x) = 0 \quad \text{for all } x \in [0,1],
\]

\[
\alpha e^x + \beta e^{2x} = 0 \quad \text{for all } x \in [0,1].
\]

Differentiating, we obtain

\[
\infty e^x + 2\beta e^{2x} = 0 \quad \text{for all } x \in [0,1].
\]

We have thus \( \beta e^{2x} = -\alpha e^x = 2\beta e^{2x} \) for all \( x \in [0,1] \), hence \( \beta = 0 \), so \( \alpha = 0 \). Therefore \( f \) and \( g \) are linearly independent over \( \mathbb{R} \).

(j) Let \( V \) be a vector space over a field \( K \) and let \( v_1, v_2 \) be vectors in \( V \) which are linearly dependent over \( K \). Then there are scalars \( \alpha, \beta \in K \), not both zero, such that \( \alpha v_1 + \beta v_2 = 0 \). If, say, \( \alpha \neq 0 \), then \( \alpha \) has an inverse \( \alpha^{-1} \) in \( K \) and we obtain \( v_1 + (\alpha^{-1}\beta) v_2 = \alpha^{-1}(\alpha v_1 + \beta v_2) = \alpha^{-1}0 = 0 \), so \( v_1 = \gamma v_2 \)

if we put \( \gamma = -\alpha^{-1}\beta \). So \( v_1 \) is a scalar multiple of \( v_2 \). Conversely, if \( v_1 \) and \( v_2 \) are vectors in \( V \) and if one of them is a scalar multiple of the other, for instance if \( v_1 = \gamma v_2 \) with some \( \gamma \in K \), then \( 1 v_1 + (-\gamma) v_2 = 0 \) and \( v_1, v_2 \) are linearly dependent over \( K \). Thus the linear dependence of two vectors means that one of them is a scalar multiple of the other.

We generalize the last example.

**42.3 Lemma:** Let \( V \) be a vector space over a field \( K \) and let \( v_1, v_2, \ldots, v_n \) be \( n \) vectors in \( V \), where \( n \geq 2 \). These vectors are linearly dependent over \( K \) if and only if one of them is a \( K \)-linear combination of the other vectors.

503
Proof: We first assume that \( v_1, v_2, \ldots, v_n \) are linearly dependent over \( K \). Then there are scalars \( \alpha_1, \alpha_2, \ldots, \alpha_n \), not all of them zero, such that
\[
\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0.
\]
To fix the ideas, let us suppose \( \alpha_1 \neq 0 \). Then \( \alpha_1 \) has an inverse \( \alpha_1^{-1} \) in \( K \) and we obtain
\[
\alpha_1^{-1}(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) = \alpha_1^{-1}0 = 0,
\]
\[
v_1 + \alpha_1^{-1}\alpha_2 v_2 + \cdots + \alpha_1^{-1}\alpha_n v_n = 0,
\]
\[
v_1 = \gamma_2 v_2 + \cdots + \gamma_n v_n
\]
where we put \( \gamma_j = \alpha_1^{-1}\alpha_j \in K \) (\( j = 2, \ldots, n \)). So \( v_1 \) is a \( K \)-linear combination of the vectors \( v_2, \ldots, v_n \).

Conversely, let us suppose that one of the vectors, for example \( v_1 \), is a linear combination of the rest, so that there are scalars \( \alpha_2, \ldots, \alpha_n \in K \) such that
\[
v_1 = \alpha_2 v_2 + \cdots + \alpha_n v_n.
\]
Then we get
\[
\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0
\]
when we write \( \alpha_1 = -1 \). Since \( \alpha_1 = -1 \neq 0 \), we see that \( v_1, v_2, \ldots, v_n \) are linearly dependent over \( K \).

A vector space over a field \( K \) can be spanned by many subsets of \( V \). Among the subsets of \( V \) which span \( V \), we want to find the ones with the least number of elements. The next two theorems, which are converses of each other, tell us that linearly dependent subsets are not useful for this purpose.

42.4 Theorem: Let \( V \) be a vector space over a field \( K \) and let \( A \) be a nonempty subset of \( V \). If \( A \) is linearly dependent over \( K \), then there is a proper subset \( B \) of \( A \) such that \( s_k(A) = s_k(B) \):

Proof: Suppose \( A \) is \( K \)-linearly dependent. If \( A \) is infinite, then, by definition, there is a finite linearly dependent subset \( A_0 \) of \( A \). If \( A \) is finite, let us put \( A_0 = A \). Hence, in both cases, \( A_0 \) is a finite linearly dependent subset of \( A \). Let \( A_0 = \{ v_0, v_1, v_2, \ldots, v_n \} \).

We first dispose of the trivial case \( |A_0| = 1, V = A_0 \). In this case we have \( n = 0 \) and \( A_0 = \{ v_0 \} \), so \( v_0 = 0 \) by Example 42.2(a), so \( A_0 = \{ 0 \} \). Thus
\[
\{ 0 \} = A_0 \subseteq A \subseteq s_k(A) \subseteq V = A_0 = \{ 0 \}
\]
and \( s_k(A) \) is equal to the \( K \)-span of the proper subset \( B = \emptyset \) of \( A \).
Suppose now $|A_0| \geq 1$ or $|A_0| = 1$ but $V \neq A_0$. Then we may and do join nonzero vectors $v_1, v_2, \ldots, v_n$ to $A_0$ without disturbing the linear dependence and finiteness of $A_0$. One of the vectors $v_0, v_1, v_2, \ldots, v_n$, which we may assume to be $v_0$ without loss of generality, is a $K$-linear combination of the others (Lemma 42.3). So there are scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that

$$v_0 = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n.$$  

We will show that $v_0$ is redundant. We put $B = A \backslash \{v_0\}$. Then $B$ is a proper subset of $A$ and $s_K(B) \subseteq s_K(A)$. We prove $s_K(A) \subseteq s_K(B)$.

Let $v \in s_K(A)$. Then there are vectors $w_1, w_2, \ldots, w_m$ in $A$ and scalars $\beta_1, \beta_2, \ldots, \beta_m$ in $K$ such that

$$v = \beta_1 w_1 + \beta_2 w_2 + \cdots + \beta_m w_m.$$  

Here we may suppose that $w_1, w_2, \ldots, w_m$ are pairwise distinct (if $w_1 = w_2$, we write $(\beta_1 + \beta_2) w_1$ instead of $\beta_1 w_1 + \beta_2 w_2$, etc.).

If none of the vectors $w_1, w_2, \ldots, w_m$ is equal to $v_0$, then $v$ is a $K$-linear combination of the vectors $w_1, w_2, \ldots, w_m$ in $B$, so $v \in s_K(B)$.

If one of the vectors $w_1, w_2, \ldots, w_m$ is equal to $v_0$, for instance if $w_1 = v_0$, then we have

$$v = \beta_1 w_1 + \beta_2 w_2 + \cdots + \beta_m w_m$$  

$$= \beta_1 (\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) + \beta_2 w_2 + \cdots + \beta_m w_m$$  

$$= \beta_1 \alpha_1 v_1 + \beta_1 \alpha_2 v_2 + \cdots + \beta_1 \alpha_n v_n + \beta_2 w_2 + \cdots + \beta_m w_m,$$

so $v$ is a $K$-linear combination of the vectors $v_1, v_2, \ldots, v_n, w_2, \ldots, w_m$ in $B = A \backslash \{v_0\}$ (some $v_i$ might equal a $w_j$, but this does not matter), so $v \in s_K(B)$.

In both cases, $v \in s_K(B)$. Thus $s_K(A) \subseteq s_K(B)$ and $s_K(A) = s_K(B)$, as was to be proved. \hfill \Box

42.5 Theorem: Let $V$ be a vector space over a field $K$ and let $A$ be a nonempty subset of $V$. If there is a proper subset $B$ of $A$ such that $s_K(B) = s_K(A)$, then $A$ is linearly dependent over $K$.

Proof: We first dispose of the trivial case $B = \emptyset$. If $B = \emptyset$, then $\emptyset \neq A \subseteq s_K(A) = s_K(B) = s_K(\emptyset) = \{0\}$; gives $A = \{0\}$ and $A$ is $K$-linearly dependent by Example 42.2(a).

Suppose now $B \neq \emptyset$. Since $B \subset A$, there is a vector $v$ in $A \backslash B$. From
\( v \in A \subseteq s_k(A) = s_k(B) \), we conclude that there are vectors \( w_1, w_2, \ldots, w_m \) in \( B \) and scalars \( \beta_1, \beta_2, \ldots, \beta_m \) in \( K \) with
\[
\nu = \beta_1 w_1 + \beta_2 w_2 + \cdots + \beta_m w_m.
\]
So the vector \( \nu \) in \( A \) is a \( K \)-linear combination of the vectors \( w_1, w_2, \ldots, w_m \) and the subset \( \{ \nu w_1, w_2, \ldots, w_m \} \) of \( A \) is \( K \)-linearly dependent by Lemma 42.3. From Example 42.2(e), it follows that \( A \) is \( K \)-linearly dependent.
\( \Box \)

The last two theorems lead us to consider linearly independent subsets of \( V \) spanning \( V \). Whether an arbitrary vector space does have such a subset will be discussed later. We give a name to the subsets in question.

42.6 Definition: Let \( V \) be a vector space over a field \( K \). A nonempty subset \( B \) of \( V \) is called a basis of \( V \) over \( K \), or a \( K \)-basis of \( V \), if \( B \) is linearly independent over \( K \) and spans \( V \) over \( K \) (i.e., \( s_k(B) = V \)). By convention, the empty set \( \emptyset \) will be called a \( K \)-basis of the vector space \( \{0\} \).

42.7 Examples: (a) Consider the vector space \( K^n \) over a field \( K \). The vectors \( u_1 = (1, 0, \ldots, 0) \), \( u_2 = (0, 1, \ldots, 0) \), \ldots, \( u_n = (0, 0, \ldots, 1) \) are linearly independent over \( K \) (Example 42.2(c)). Moreover, \( \{ u_1, u_2, \ldots, u_n \} \) spans \( K^n \) over \( K \) because any vector \( (\alpha_1, \alpha_2, \ldots, \alpha_n) \) in \( K^n \) is a \( K \)-linear combination
\[
\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n
\]
of the vectors \( u_1, u_2, \ldots, u_n \). Hence \( \{ u_1, u_2, \ldots, u_n \} \) is a basis of \( K^n \) over \( K \).

(b) Let \( V = \{ h \in C^2([0,1]): h''(x) - 3h'(x) + 2h(x) = 0 \text{ for all } x \in [0,1] \} \). Then \( V \) is an \( \mathbb{R} \)-subspace of \( C^2([0,1]) \), as can be verified directly and also follows from Example 40.6(k). From the theory of ordinary differential equations, it is known that every function in \( V \) (that is, every solution of \( y'' - 3y' + 2 = 0 \)) can be written in the form \( c_1 f + c_2 g \), where \( c_1, c_2 \in \mathbb{R} \) and \( f(x) = e^x, g(x) = e^{2x} \) for all \( x \in [0,1] \). Thus \( \{ f, g \} \) spans \( V \) over \( \mathbb{R} \). Also, \( \{ f, g \} \) is linearly independent over \( \mathbb{R} \) by Example 42.2(i). Hence \( \{ f, g \} \) is an \( \mathbb{R} \)-basis of \( V \).
42.8 Theorem: Let $V$ be a vector space over a field $K$ and let $B = \{v_1, v_2, \ldots, v_n\}$ be a nonempty subset of $V$. Then $B$ is a $K$-basis of $V$ if and only if every element of $V$ can be written in the form

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n, \quad (\alpha_1, \alpha_2, \ldots, \alpha_n \in K)$$

in a unique way (i.e., with unique scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$).

Proof: Assume first that $B$ is a $K$-basis of $V$. Then $V = s_K(B)$, and every element of $V$ can be written as

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$$

with suitable scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$. We are to show the uniqueness of this representation. In other words, we must prove that, if

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n$$

(1)

then $\alpha_1 = \beta_1$, $\alpha_2 = \beta_2$, $\ldots$, $\alpha_n = \beta_n$. This is easy: if (1) holds, then

$$(\alpha_1 - \beta_1) v_1 + (\alpha_2 - \beta_2) v_2 + \cdots + (\alpha_n - \beta_n) v_n = 0$$

and we obtain, since $B = \{v_1, v_2, \ldots, v_n\}$ is $K$-linearly independent, that

$$\alpha_1 - \beta_1 = \alpha_2 - \beta_2 = \cdots = \alpha_n - \beta_n = 0.$$ This proves uniqueness.

Conversely, let us suppose that every vector in $V$ can be written in the form

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$$

with unique scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$. Then $V = s_K(v_1, v_2, \ldots, v_n) = s_K(B)$. Moreover, $B$ is linearly independent over $K$, for if $\alpha_1, \alpha_2, \ldots, \alpha_n \in K$ are scalars such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0,$$

then

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0 v_1 + 0 v_2 + \cdots + 0 v_n$$

and the uniqueness of the scalars in the representation of $0 \in V$ as a $K$-linear combination of $v_1, v_2, \ldots, v_n$ implies that $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. Thus $B$ is $K$-linearly independent and consequently $B$ is a $K$-basis of $V$.  

We prove next that any finitely spanned vector space has a basis.

42.9 Theorem: Let $V$ be a vector space over a field $K$ and assume $T$ is a finite subset of $V$ spanning $V$, so that $s_K(T) = V$. Then $V$ has a finite $K$-basis. In fact, a suitable subset of $T$ is a $K$-basis of $V$.  

507
Proof: If \( V \) happens to be the vector space \( \{0\} \), then \( V \) has a \( K \)-basis, namely the empty set \( \emptyset \) (Definition 42.6) and \( \emptyset \subseteq T \). Having disposed of this degenerate case, let us assume \( V \not\equiv 0 \). Now \( s_k(T) = V \). Since \( V \not\equiv 0 \), we have \( T \not\subseteq \emptyset \). If \( T \) is linearly independent over \( K \), then \( T \) is a \( K \)-basis of \( V \). Otherwise, there is a proper subset \( T_1 \) of \( T \) with \( s_k(T_1) = s_k(T) = V \) (Theorem 42.4). Here \( T_1 \not\subseteq \emptyset \), because \( s_k(T_1) = V \not\equiv \{0\} \). If \( T_1 \) is linearly independent over \( K \), then \( T_1 \) is a \( K \)-basis of \( V \). Otherwise, there is a proper subset \( T_2 \) of \( T_1 \) with \( s_k(T_2) = s_k(T_1) = V \). Here \( T_2 \not\subseteq \emptyset \), because \( s_k(T_2) = V \not\equiv \{0\} \). If \( T_2 \) is linearly independent over \( K \), then \( T_2 \) is a \( K \)-basis of \( V \). Otherwise, there is a proper subset \( T_3 \) of \( T_2 \) with \( s_k(T_3) = s_k(T_2) = V \). Here \( T_3 \not\subseteq \emptyset \), because \( s_k(T_3) = V \not\equiv \{0\} \). We continue in this way. Each time, we get a nonempty subset \( T_{i+1} \) of \( T_i \) such that \( s_k(T_{i+1}) = V \) and \( T_{i+1} \) has less elements than \( T_i \). Since \( T \) is a finite set, this process cannot go on indefinitely. Sooner or later, we will meet a \( K \)-linearly independent subset \( T_m \) of \( T \) with \( s_k(T_m) = V \). This \( T_m \) is therefore a \( K \)-basis of \( V \), and of course \( T_m \) is finite.

Having convinced ourselves of the existence of bases in some vector spaces, we turn our attention to the number of vectors in a finite basis. We show that the number of linearly independent vectors in a subspace cannot exceed the number of vectors spanning the subspace. This theorem, due to E. Steinitz (1871-1928), is the source of many deep results concerning the dimension of a vector space. The idea is to replace some vectors in the spanning set by the vectors in the linearly independent set without changing the span.

42.10 Theorem (Steinitz' replacement theorem): Let \( V \) be a vector space over a field \( K \) and \( w_1, w_2, \ldots, w_m \) be finitely many vectors in \( V \). Let \( v_1, v_2, \ldots, v_n \) be \( n \) linearly independent vectors in the \( K \)-span
\[
s_k(w_1, w_2, \ldots, w_m)
\]
of \( w_1, w_2, \ldots, w_m \). Then \( n \leq m \). Moreover, there are \( n \) vectors among \( w_1, w_2, \ldots, w_m \), which we may assume to be \( w_1, w_2, \ldots, w_n \), such that

\[
s_k(v_1, v_2, \ldots, v_n, w_{n+1}, \ldots, w_m) = s_k(w_1, w_2, \ldots, w_m).
\]

Proof: For \( 1 \leq h \leq n \), let \( A_h \) be the assertion

"there are \( h \) vectors among \( w_1, w_2, \ldots, w_m \), say \( w_1, \ldots, w_h \), such that

\[
508
\[ s_K(v_1, \ldots, v_h, w_{h+1}, \ldots, w_m) = s_K(w_1, \ldots, w_h, w_{h+1}, \ldots, w_m) \]

We show that
1. If \( A_1 \) is true,
2. If \( 2 \leq h \leq n \) and \( A_{h-1} \) is true, then \( A_h \) is true.

This will establish \( A_1, A_2, \ldots, A_{n-1}, A_n \). The second claim \( A_n \) in the theorem will be proved in this way.

(1) \( A_1 \) is true. We have \( v_1 \in s_K(w_1, w_2, \ldots, w_m) \), so
\[ v_1 = \beta_1 w_1 + \beta_2 w_2 + \cdots + \beta_m w_m \]
with some scalars \( \beta_1, \beta_2, \ldots, \beta_m \in K \). Since \( v_1, v_2, \ldots, v_n \) are linearly independent over \( K \), \( v_1 \neq 0 \) (Example 42.2(f)), hence not all of \( \beta_1, \beta_2, \ldots, \beta_m \) are equal to 0 \( \in K \). So one of them is distinct from 0. Renaming \( w_1, w_2, \ldots, w_m \) if necessary, we may suppose \( \beta_1 \neq 0 \). Then \( \beta_1 \) has an inverse \( \beta_1^{-1} \) in \( K \) and we get
\[ w_1 = \beta_1^{-1}(v_1 - \beta_2 w_2 - \cdots - \beta_m w_m) \]
\[ (w_1, w_2, \ldots, w_m) \subseteq s_K(v_1, w_2, \ldots, w_m). \]

Since \( v_1 \in s_K(w_1, w_2, \ldots, w_m) \), we also have \( (w_1, w_2, \ldots, w_m) \subseteq s_K(v_1, w_2, \ldots, w_m) \). Using (i) and (ii) and applying Lemma 40.11 (with \( A = \{ w_1, w_2, \ldots, w_m \} \) and \( B = \{ v_1, w_2, \ldots, w_m \} \), we obtain
\[ s_K(v_1, w_2, \ldots, w_m) = s_K(w_1, w_2, \ldots, w_m). \]

This proves \( A_1 \).

(2) Suppose \( 2 \leq h \leq n \) and \( A_{h-1} \) is true. Then \( A_h \) is true. The truth of \( A_{h-1} \) means
\[ s_K(v_1, \ldots, v_{h-1}, w_h, \ldots, w_m) = s_K(w_1, w_2, \ldots, w_m) \]
provided the \( w_i \)'s are indexed suitably. We have
\[ v_h \in s_K(w_1, \ldots, w_{h-1}, w_h, \ldots, w_m) \]
\[ v_h \in s_K(v_1, \ldots, v_{h-1}, w_h, \ldots, w_m), \]
so
\[ v_h = \alpha_1 v_1 + \cdots + \alpha_{h-1} v_{h-1} + \alpha_h w_h + \cdots + \alpha_m w_m \]
for some appropriate \( \alpha_1, \ldots, \alpha_{h-1}, \alpha_h, \ldots, \alpha_m \in K \). Here not all of \( \alpha_h, \ldots, \alpha_m \) are equal to 0 \( \in K \), for then \( v_h \) would be a \( K \)-linear combination
\[ \alpha_1 v_1 + \cdots + \alpha_{h-1} v_{h-1} \]
of the vectors \( v_1, \ldots, v_{h-1} \) and the vectors \( v_1, \ldots, v_h \) would not be linearly independent over \( K \) (Lemma 42.3), so \( v_1, v_2, \ldots, v_n \) would not be linearly independent over \( K \) (Example 42.2(e)), contrary to the hypothesis. So one of \( \alpha_h, \ldots, \alpha_m \) is distinct from 0. Renaming \( w_h, \ldots, w_m \) if necessary, we may suppose \( \alpha_h \neq 0 \). Then \( \alpha_h \) has an inverse \( \alpha_h^{-1} \) in \( K \) and we get
\[ \alpha_h w_h = \alpha_1 v_1 + \cdots + \alpha_{h-1} v_{h-1} - v_h + \alpha_{h+1} w_{h+1} + \cdots + \alpha_m w_m. \]
\[ F_h = -\sigma_h^{-1}(\sigma_1 V_1 + \cdots + \sigma_{h-1} V_{h-1} - V_h + \sigma_{h+1} W_{h+1} + \cdots + \sigma_m W_m), \]

\[ W_h \in s_K(V_1, \ldots, V_{h-1}, V_h, W_{h+1}, \ldots, W_m). \]

Now each one of the vectors \( W_1, \ldots, W_{h-1} \), being an element of the span

\[ s_K(W_1, W_2, \ldots, W_m) = s_K(V_1, \ldots, V_{h-1}, W_h, \ldots, W_m), \]

can be written in the form

\[ \gamma_1 V_1 + \cdots + \gamma_{h-1} V_{h-1} + \gamma_h W_h + \gamma_{h+1} W_{h+1} + \cdots + \gamma_m W_m \]

with scalars \( \gamma_1, \ldots, \gamma_{h-1}, \gamma_h, \gamma_{h+1}, \ldots, \gamma_m \in K \). Thus each one of \( W_1, \ldots, W_{h-1} \) can be written as

\[ \gamma_1 V_1 + \cdots + \gamma_{h-1} V_{h-1} + \gamma_h (-\sigma_h^{-1}(\sigma_1 V_1 + \cdots + \sigma_{h-1} V_{h-1} - V_h + \sigma_{h+1} W_{h+1} + \cdots + \sigma_m W_m)) + \gamma_{h+1} W_{h+1} + \cdots + \gamma_m W_m, \]

and so \( \{ W_1, \ldots, W_{h-1} \} \subseteq s_K(V_1, \ldots, V_{h-1}, V_h, W_{h+1}, \ldots, W_m) \).

Therefore \( \{ W_1, \ldots, W_{h-1}, W_h, W_{h+1}, \ldots, W_m \} \subseteq s_K(V_1, \ldots, V_{h-1}, V_h, W_{h+1}, \ldots, W_m) \). (i’)

Since

\[ V_1, \ldots, V_{h-1}, V_h \in s_K(W_1, \ldots, W_{h-1}, W_h, W_{h+1}, \ldots, W_m), \]

we also have

\[ \{ V_1, \ldots, V_{h-1}, V_h, W_{h+1}, \ldots, W_m \} \subseteq s_K(W_1, \ldots, W_{h-1}, W_h, W_{h+1}, \ldots, W_m) \). (ii’)

Using (i’) and (ii’) and applying Lemma 40.11 with

\[ A = \{ W_1, \ldots, W_{h-1}, W_h, W_{h+1}, \ldots, W_m \}, B = \{ V_1, \ldots, V_{h-1}, V_h, W_{h+1}, \ldots, W_m \}, \]

we obtain

\[ s_K(V_1, \ldots, V_{h-1}, V_h, W_{h+1}, \ldots, W_m) = s_K(W_1, W_2, \ldots, W_m). \]

Thus \( A_n \) is true.

As remarked earlier, this establishes the truth of \( A_1 A_2 \cdots A_{n-1} A_n \). In particular, \( A_n \) is true, and the second statement in the enunciation is proved. Now it remains to establish \( n \leq m \).

If we had \( m < n \), then \( A_m \) would be true and we would get

\[ s_K(V_1, V_2, \ldots, V_m) = s_K(W_1, W_2, \ldots, W_m). \]

Then

\[ V_n \in s_K(W_1, W_2, \ldots, W_m) \]

would give

\[ V_n \in s_K(V_1, V_2, \ldots, V_m), \]

contrary to the hypothesis that \( V_1, V_2, \ldots, V_m, \ldots, V_n \) are linearly independent over \( K \). So \( m < n \) is impossible and necessarily \( n \leq m \). This completes the proof. □
42.11 Theorem: Let $V$ be a vector space over a field $K$ and assume that $V$ has a finite $K$-basis. Then any two $K$-bases of $V$ have the same number of elements.

Proof: There is a finite $K$-basis of $V$ by hypothesis, say $B$. Assume that $B$ has exactly $n$ vectors ($n \geq 0$). We prove that any $K$-basis $B_1$ of $V$ has also $n$ vectors in it.

If $n = 0$, then $B = \emptyset$ and $V = \{0\}$. Thus $\emptyset$ and $\{0\}$ are the only subsets of $V$ and $B = \emptyset$ is the only $K$-basis of $V$. Then any $K$-basis of $V$ has exactly 0 elements.

Suppose now $n \geq 1$ and let $B_1$ be any $K$-basis of $V$. First we show that $B_1$ cannot be infinite. Otherwise, $B_1$ would be an infinite $K$-linearly independent subset of $V$. Every finite subset of $B_1$ would be $K$-linearly independent by definition. Let $v_1, v_2, \ldots, v_n, v_{n+1}$ be $n + 1$ $K$-linearly independent vectors in $B_1$. These $n + 1$ vectors lie in the $K$-span $s_k(B)$ of $B$ and $B$ has $n$ elements. Steinitz’ replacement theorem gives $n + 1 \leq n$, which is absurd. Thus $B_1$ cannot be infinite.

We put $|B_1| = n_1$. Here $n_1 \neq 0$, because $n_1 = 0$ would imply $B_1 = \emptyset$ and $\emptyset \neq B \subseteq s_k(B) = V = s_k(B_1) = s_k(\emptyset) = \{0\}$, so $B = \{0\}$ and $B$ would be linearly independent over $K$, contrary to the hypothesis that $B$ is a $K$-basis of $V$. So $n_1 \in \mathbb{N}$.

$B_1$ is a $K$-linearly independent subset of $V$ in $s_k(B)$, therefore $n_1 \leq n$ by Steinitz’ replacement theorem. Likewise, $B$ is a $K$-linearly independent subset of $V$ in $s_k(B_1)$, so $n \leq n_1$. Therefore $n = n_1$, as was to be proved. \(\square\)

42.12 Definition: Let $V$ be a vector space over a field $K$. If $V$ has a finite $K$-basis, the number of elements in any $K$-basis of $V$, which is the same for all $K$-bases of $V$ by Theorem 42.11, is called the dimension of $V$ over $K$, or the $K$-dimension of $V$. It is denoted as $\dim_k V$ or as $\dim V$. If $V$ has no finite $K$-basis, then the $K$-dimension of $V$ is defined to be infinity, and we write in this case $\dim_k V = \infty$. 

511
Thus $\dim_k k^n = n$ (Example 42.7(a)) and $\dim_\mathbb{R} V = 2$, where $V$ is the $\mathbb{R}$-vector space of Example 42.7(b).

We frequently say that $V$ is $n$-dimensional when $\dim V = n$. A vector space is said to be finite dimensional if $\dim V$ is a nonnegative integer and infinite dimensional if $\dim V = \infty$. Notice that the dimension of the vector space $\{0\}$ is zero.

**42.13 Lemma:** Let $V$ be a vector space over a field $K$, let $\dim_k V = n \in \mathbb{N}$ and let $\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_n$ be $n$ vectors in $V$.

1. If $\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_n$ are linearly independent over $K$, then $s_k(\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_n) = V$.
2. If $s_k(\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_n) = V$, then $\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_n$ are linearly independent over $K$.

**Proof:**

1. We are given $\dim_k V = n \in \mathbb{N}$. Let $\{w_1, w_2, \ldots, w_n\}$ be a basis of $V$ over $K$. If $\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_n$ are $K$-linearly independent vectors in $V = s_k(w_1, w_2, \ldots, w_n)$, then we obtain $s_k(\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_n) = s_k(w_1, w_2, \ldots, w_n)$ by Steinitz' replacement theorem.

2. Suppose $s_k(\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_n) = V$. If $\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_n$ are not linearly independent over $K$, then there is a proper subset $T \subset \{\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_n\}$ of $\{\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_n\}$ with $s_k(T) = s_k(\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_n) = V$ (Theorem 42.4), and there is a $K$-basis $B$ of $V$ such that $B \subseteq T$ (Theorem 42.9). Using Theorem 42.11, we obtain the contradiction

$$n = \dim_k V = |B| \leq |T| < n.$$

Hence $\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_n$ have to be linearly independent over $K$. \hfill \Box

Any finite set spanning a vector space can be stripped off to a basis of that vector space (Theorem 42.9). Similarly, any linearly independent subset of a vector space can be extended to a basis, as we show now.

**42.14 Theorem:** Let $V$ be an $m$-dimensional vector space over a field $K$, with $m \geq 1$. Let $n \geq 1$ and let $\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_n$ be $n$ linearly independent vectors in $V$. Then there is a $K$-basis $B$ of $V$ such that $\{\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_n\} \subseteq B$. 512
**Proof:** Let \( \{ w_1, w_2, \ldots, w_m \} \) be a \( K \)-basis of \( V \). Then \( v_1, v_2, \ldots, v_n \) are linearly independent vectors in \( V = s_k(w_1, w_2, \ldots, w_m) \), and Steinitz' replacement theorem gives

\[
s_k(v_1, v_2, \ldots, v_n, w_{n+1}, \ldots, w_m) = V \quad \text{and} \quad n \leq m
\]
on indexing \( u \)'s suitably. Then the \( m = \dim_k V \) vectors

\[
v_1, v_2, \ldots, v_n, w_{n+1}, \ldots, w_m
\]
are linearly independent over \( K \) by Lemma 42.13(2). Hence

\[
B = \{ v_1, v_2, \ldots, v_n, w_{n+1}, \ldots, w_m \}
\]
is a \( K \)-basis of \( V \) containing the vectors \( v_1, v_2, \ldots, v_n \).

\[
\square
\]

42.15 Lemma: Let \( V \) be a finite dimensional vector space over a field \( K \) and let \( W \) be a subspace of \( V \).
(1) \( W \) is finite dimensional; in fact \( \dim_k W \leq \dim_k V \).
(2) \( \dim_k W = \dim_k V \) if and only if \( W = V \).

**Proof:** Let \( n = \dim_k V \).

(1) The assertion is trivial when \( W = \{0\} \), so let us assume \( W \neq \{0\} \). Then there is a nonzero vector \( v \) in \( V \), and \( \{ v \} \) is a \( K \)-linearly independent subset with one element. On the other hand, any \( n + 1 \) vectors in \( V \) (in fact in \( V \)) are linearly dependent over \( K \) by Steinitz' replacement theorem. Therefore there exists a natural number \( m \) such that

(a) \quad \( 1 \leq m \leq n \),

(b) there are \( m \) linearly independent vectors in \( W \),

(c) any \( m + 1 \) vectors in \( W \) are linearly dependent.

This \( m \) is clearly unique in view of (b) and (c). The natural number \( m \) having been defined in this way, let \( w_1, w_2, \ldots, w_m \) be \( m \) \( K \)-linearly independent vectors in \( W \). We claim that \( \{ w_1, w_2, \ldots, w_m \} \) is a \( K \)-basis of \( W \).

To show this, we must prove only that these vectors span \( W \) over \( K \). Let \( \alpha \) be an arbitrary vector in \( W \). Then \( w, w_1, w_2, \ldots, w_m \) are linearly dependent over \( K \) by (c). Hence

\[
\beta w + \beta_1 w_1 + \beta_2 w_2 + \cdots + \beta_m w_m = 0
\]

with some scalars \( \beta, \beta_1, \beta_2, \ldots, \beta_m \) in \( K \). Here \( \beta \neq 0 \), for otherwise the equation above would imply that \( w_1, w_2, \ldots, w_m \) are \( K \)-linearly dependent. Hence \( \beta \) has an inverse \( \beta^{-1} \) in \( K \) and we get

\[
\alpha = (-\beta^{-1} \beta_1) w_1 + (-\beta^{-1} \beta_2) w_2 + \cdots + (-\beta^{-1} \beta_m) w_m,
\]

\[
\alpha \in s_k(w_1, w_2, \ldots, w_m).
\]
This gives \( W \subseteq s_k(\omega_1, \omega_2, \ldots, \omega_m) \). But \( \omega_1, \omega_2, \ldots, \omega_m \) belong to \( W \), so \( s_k(\omega_1, \omega_2, \ldots, \omega_m) \subseteq W \) (Lemma 40.5). The vectors \( \omega_1, \omega_2, \ldots, \omega_m \) therefore span \( W \over K \), so \( \{ \omega_1, \omega_2, \ldots, \omega_m \} \) is a \( K \)-basis of \( W \). Thus \( W \) is finite dimensional and in fact \( \dim_k W = m \leq n = \dim_k V \).

(2) If \( W = V \), then of course \( \dim_k W = \dim_k V \). Suppose conversely \( \dim_k W = \dim_k V = n \) and let \( A \) be a \( K \)-basis of \( W \). Then there is a \( K \)-basis \( B \) of \( V \) with \( A \subseteq B \): this follows from Theorem 42.14 when \( A \neq \emptyset \) and is obvious when \( A = \emptyset \). Then

\[
n = \dim_k W = |A| \leq |B| = \dim_k V = n
\]

implies that \( A = B \). Thus \( W = s_k(A) = s_k(B) = V \). \( \square \)

42.16 Lemma: Let \( V, U \) be vector spaces over a field \( K \). Suppose \( V \) is finite dimensional and let \( \varphi: V \to U \) be a vector space homomorphism. Let \( v_1, v_2, \ldots, v_n \) be vectors in \( V \).

(1) If \( \varphi \) is one-to-one and \( \{v_1, v_2, \ldots, v_n\} \) is linearly independent over \( K \), then \( \{v_1\varphi, v_2\varphi, \ldots, v_n\varphi\} \) is linearly independent over \( K \).

(2) If \( \varphi \) is onto \( U \) and \( \{v_1, v_2, \ldots, v_n\} \) spans \( V \) over \( K \), then \( \{v_1\varphi, v_2\varphi, \ldots, v_n\varphi\} \) spans \( U \) over \( K \).

(3) If \( \varphi \) is a vector space isomorphism and \( \{v_1, v_2, \ldots, v_n\} \) is a \( K \)-basis of \( V \), then \( \{v_1\varphi, v_2\varphi, \ldots, v_n\varphi\} \) is a \( K \)-basis of \( U \). In particular, \( \dim_k U = \dim_k V \).

Proof: (1) Suppose \( \alpha_1, \alpha_2, \ldots, \alpha_n \) are scalars such that

\[
\alpha_1(v_1\varphi) + \alpha_2(v_2\varphi) + \cdots + \alpha_n(v_n\varphi) = 0.
\]

Then

\[
(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n)\varphi = 0,
\]

so

\[
\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \in \ker \varphi,
\]

and

\[
\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0
\]

since \( \ker \varphi = 0 \) as \( \varphi \) is one-to-one. Since \( v_1, v_2, \ldots, v_n \) are \( K \)-linearly independent, we get \( \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0 \). Hence \( v_1\varphi, v_2\varphi, \ldots, v_n\varphi \) are linearly independent over \( K \).

(2) We must show that any element of \( U \) can be written as a \( K \)-linear combination of the vectors \( v_1\varphi, v_2\varphi, \ldots, v_n\varphi \). Let \( u \in U \). Then there is a \( v \in V \) with \( \varphi = u \) and

\[
u = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n, \quad v_1 v_1 + \cdots + v_n v_n \text{ are suitable scalars in } K.
\]

This yields

\[
\varphi = (\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n)\varphi
\]

\[
= \alpha_1 (v_1\varphi) + \alpha_2 (v_2\varphi) + \cdots + \alpha_n (v_n\varphi) \in s_k(v_1\varphi, v_2\varphi, \ldots, v_n\varphi).
\]

514
as was to be proved.

(3) This follows immediately from (1) and (2). □

From Lemma 42.16, it follows that \( \dim_k U = n \) whenever \( U \cong K^n \). The converse of this statement is also true.

**42.17 Theorem:** Let \( V \) be a vector space over a field \( K \). Then \( \dim_k V = n \in \mathbb{N} \) if and only if \( V \cong K^n \) (as vector spaces).

**Proof:** If \( V \cong K^n \), then \( \dim_k V = \dim_k K^n = n \) by Lemma 42.16(3). Suppose conversely that \( \dim_k V = n \) and let \( \{ v_1, v_2, \ldots, v_n \} \) be a \( K \)-basis of \( V \). Every element \( v \) of \( V \) can be written in a unique way as \( \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \), where \( \alpha_1, \alpha_2, \ldots, \alpha_n \in K \). We consider the mapping

\[ \varphi: V \longrightarrow K^n \]

\[ \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \longrightarrow (\alpha_1, \alpha_2, \ldots, \alpha_n). \]

This \( \varphi \) is a \( K \)-linear transformation, since, for any \( \alpha, \beta \in K \) and \( v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \), \( w = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n \) in \( V \), we have

\[
(\alpha v + \beta w) \varphi = \left((\alpha \alpha_1 + \beta \beta_1) v_1 + (\alpha \alpha_2 + \beta \beta_2) v_2 + \cdots + (\alpha \alpha_n + \beta \beta_n) v_n\right) \varphi
\]

\[
= (\alpha \varphi + \beta \varphi).
\]

Furthermore, \( v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \in V \) belongs to \( \text{Ker} \ \varphi \) if and only if \( (\alpha_1, \alpha_2, \ldots, \alpha_n) = (0, 0, \ldots, 0) \), thus if and only if \( v = 0 v_1 + 0 v_2 + \cdots + 0 v_n = 0 \). So \( \text{Ker} \ \varphi = \{0\} \) and \( \varphi \) is one-to-one.

Since any \( n \)-tuple \( (\alpha_1, \alpha_2, \ldots, \alpha_n) \) in \( K^n \) is the image, under \( \varphi \), of the vector \( \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \) in \( V \), we see that \( \varphi \) is onto.

Hence \( \varphi \) is a vector space isomorphism and \( V \cong K^n \). □
42.18 Theorem: Let $V$ and $U$ be finite dimensional vector spaces over a field $K$. Then $V \cong U$ if and only if $\dim_k V = \dim_k U$.

Proof: The case when $\dim_k V = 0$ or $\dim_k U = 0$ is trivial. Let us suppose $\dim_k V \geq 1$ and $\dim_k U \geq 1$. If $V \cong U$, then $\dim_k V = \dim_k U$ by Lemma 42.16(3). If $\dim_k V = \dim_k U$, then $V \cong K^{\dim_k V} = K^{\dim_k U} \cong U$ by Theorem 42.17, hence $V \cong U$. \hfill $\square$

42.19 Theorem: Let $V$ be a vector space over a field $K$ and let $W$ be a subspace of $V$. If $V$ is finite dimensional, then $V/W$ is finite dimensional. In fact,

$$\dim_k V = \dim_k W + \dim_k V/W.$$ 

Proof: We eliminate the trivial cases. We know that $W$ is finite dimensional (Lemma 42.15(1)). If $\dim_k W = 0$, then $W = \{0\}$, so $V \cong V/\{0\} = V/W$, so $\dim_k V/W = \dim_k V$ and $\dim_k V = 0 + \dim_k V = \dim_k W + \dim_k V/W$. If $\dim_k W = \dim_k V$, then $W = V$ (Lemma 42.15(2)), so $V/W \cong \{0\}$ and $\dim_k V = \dim_k V + 0 = \dim_k W + \dim_k V/W$. Thus the theorem is proved in case $\dim_k W = 0$ or $\dim_k W = \dim_k V$ (in particular in case $\dim_k V = 0$).

Let us assume now $0 < \dim_k W < \dim_k V$. Let $\dim_k W = m$ and let 

$\{w_1, w_2, \ldots, w_m\}$ be a $K$-basis of $W$. There are vectors $u_1, u_2, \ldots, u_k$ in $V$ such that 

$\{w_1, w_2, \ldots, w_m, u_1, u_2, \ldots, u_k\}$ is a $K$-basis of $V$ (Theorem 42.14). Here $k \geq 1$ and $m + k = \dim_k V$. We claim that $\{u_1 + W, u_2 + W, \ldots, u_k + W\}$ is a $K$-basis of $V/W$. This will imply $k = \dim_k V/W$, hence $\dim_k V = m + k = \dim_k W + \dim_k V/W$.

To establish our claim, we note first that $u_1 + W, u_2 + W, \ldots, u_k + W$ are $K$-linearly independent vectors in $V/W$. Indeed, if $\alpha_1, \alpha_2, \ldots, \alpha_k$ are scalars such that

$$\alpha_1 (u_1 + W) + \alpha_2 (u_2 + W) + \cdots + \alpha_k (u_k + W) = 0 + W,$$

then

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k \in W = s_k(w_1, w_2, \ldots, w_m),$$

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k = \beta_1 w_1 + \beta_2 w_2 + \cdots + \beta_m w_m$$

where $\beta_1, \beta_2, \ldots, \beta_m$ are appropriate scalars in $K$. Then

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k = -\beta_1 w_1 - \beta_2 w_2 - \cdots - \beta_m w_m = 0$$

and linearly independence of $w_1, w_2, \ldots, w_m, u_1, u_2, \ldots, u_k$ implies that $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$. Thus $u_1 + W, u_2 + W, \ldots, u_k + W$ in $V/W$ are linearly independent over $K$. 

516
Secondly, these vectors span $V/W$. To see this, let us take an arbitrary vector $\nu + W$ in $V/W$, where $\nu \in V$. Then
\[
\nu = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k + \beta_1 w_1 + \beta_2 w_2 + \cdots + \beta_m w_m
\]
where $\alpha_1, \alpha_2, \ldots, \alpha_k, \beta_1, \beta_2, \ldots, \beta_m$ are scalars, and thus
\[
\nu + W = (\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k + \beta_1 w_1 + \beta_2 w_2 + \cdots + \beta_m w_m) + W
\]
\[
= \alpha_1 (u_1 + W) + \alpha_2 (u_2 + W) + \cdots + \alpha_k (u_k + W) + 
+ \beta_1 (w_1 + W) + \beta_2 (w_2 + W) + \cdots + \beta_m (w_m + W)
\]
\[
\in s_k(u_1 + W, u_2 + W, \ldots, u_k + W)
\subseteq V/W,
\]

so $V/W = s_k(u_1 + W, u_2 + W, \ldots, u_k + W)$. This proves that
\[
\{u_1 + W, u_2 + W, \ldots, u_k + W\}
\]
is a basis of $V/W$ over $K$. As we remarked above, this gives $\dim_K V = \dim_K W + \dim_K V/W$.

We deduce important corollaries from Theorem 42.19.

42.20 Theorem: Let $V$ be a vector space over a field $K$. Let $W, U$ be finite dimensional subspaces of $V$. Then $W + U$ is a finite dimensional subspace of $V$ and in fact
\[
\dim_K (W + U) = \dim_K W + \dim_K U - \dim_K (W \cap U).
\]

Proof: If $\{w_1, w_2, \ldots, w_m\}$ is a $K$-basis of $W$ and $\{u_1, u_2, \ldots, u_k\}$ is a $K$-basis of $U$, then $W + U = \{w + u \in V : w \in W, u \in U\}$ is clearly spanned by the finite set $\{w_1, w_2, \ldots, w_m, u_1, u_2, \ldots, u_k\}$, hence $W + U$ is finite dimensional (Theorem 42.9). From $W + U/U \cong W/W \cap U$ (Theorem 41.16), we obtain then
\[
\dim_K (W + U) - \dim_K U = \dim_K (W + U/U)
\]
\[
= \dim_K (W/W \cap U)
\]
\[
= \dim_K W - \dim_K (W \cap U).
\]

42.21 Theorem: Let $V$ be a vector space over a field $K$ and let $\varphi$ be a $K$-linear transformation from $V$. If $V$ is finite dimensional, then
\[
\dim_K \ker \varphi + \dim_K \operatorname{im} \varphi = \dim_K V.
\]
Proof: Theorem 41.13 tells us \( V / \text{Ker } \varphi \cong \text{Im } \varphi \) and Theorem 42.19 gives
\[
\dim_k V - \dim_k \text{Ker } \varphi = \dim_k \text{Im } \varphi.
\]

42.22 Theorem: Let \( V, U \) be vector spaces over a field \( K \) and let \( \varphi: V \to U \) be a \( K \)-linear mapping. Suppose that \( V \) and \( U \) have the same finite dimension. Then the following statements are equivalent.

1. \( \varphi \) is one-to-one.
2. \( \varphi \) is onto.
3. \( \varphi \) is a vector space isomorphism.

Proof: (1) \( \implies \) (2) If \( \varphi \) is one-to-one, then \( \text{Ker } \varphi = \{0\} \), so \( \dim_k \text{Ker } \varphi = 0 \) and \( \dim_k \text{Im } \varphi = \dim_k \text{Ker } \varphi + \dim_k \text{Im } \varphi = \dim_k V = \dim_k U \). Thus \( \text{Im } \varphi \) is a subspace of \( U \) with \( \dim_k \text{Im } \varphi = \dim_k U \), and Lemma 42.15(2) gives then \( \text{Im } \varphi = U \). Hence \( \varphi \) is onto.

(2) \( \implies \) (1) If \( \varphi \) is onto, then \( \text{Im } \varphi = U \), so \( \dim_k \text{Im } \varphi = \dim_k U \) and \( \dim_k \text{Ker } \varphi = \dim_k V - \dim_k \text{Im } \varphi = \dim_k U - \dim_k U = 0 \). Thus \( \text{Ker } \varphi = \{0\} \) and \( \varphi \) is one-to-one.

Hence any one of (1),(2) implies the other, and these together imply (3). Conversely, if \( \varphi \) is an isomorphism, then of course \( \varphi \) is one-to-one and onto. Thus (3) implies both (1) and (2).

We close this paragraph with a brief discussion of infinite dimensional vector spaces. Do infinite dimensional vector spaces have bases? From Theorem 42.9, we know that such a vector space cannot be spanned by a finite set. But if \( B \) is a spanning set, necessarily infinite, the argument of Theorem 42.9 does not work. To prove the existence of bases of infinite dimensional vector spaces, we have to resort to more sophisticated means.

It is in fact true that every vector space has a basis, and a proof is given in the appendix. The proof of this statement for infinite dimensional vector spaces requires a fundamental tool known as Zorn's lemma. This lemma can be used in a variety of situations to establish the existence of certain objects.
The existence of bases having been assured by Zorn's lemma, we might ask whether any two bases have the same cardinality. The answer turned out to be "yes" in the finite dimensional case (Theorem 42.11), and this was proved by using Steinitz' replacement theorem. The proof of Steinitz' replacement theorem does not extend to the infinite dimensional case. Nevertheless, theorems of set theory can be employed to show that two bases of a vector space have the same cardinal number. Thus renders it possible to define the dimension of a vector space as the cardinality of a basis. Hence it is possible to distinguish between various types of infinities. This is much finer than Definition 42.12, by which infinite dimensionality is merely a crude negation of finite dimensionality.

Theorem 42.14, which states that any linearly independent subset can be extended to a basis, is true in the infinite dimensional case, too. The proof makes use of Zorn's lemma.

Lemma 42.15(1) remains valid also in the infinite dimensional case, in the sense that a basis of a subspace has a cardinal number less than or equal to the cardinality of a basis of the whole space. Lemma 42.15(2), however, is not necessarily true for infinite dimensional vector spaces: a proper subspace may have the same dimension as the whole space (think of $\mathbb{R}$ and $\mathbb{C}$ as $\mathbb{Q}$-vector spaces).

Lemma 42.16 and its proof works in the infinite dimensional case.

Lemma 42.19 and its proof works in the infinite dimensional case, provided we refer to the generalization of Theorem 42.14 at the appropriate place.

Generally speaking, infinite dimensional vector spaces are wild objects. To render them more manageable, one equips them with some additional structure, perhaps with a topological or analytic one.

Exercises

1. Let $V$ be a vector space over a field $K$ and let $W$ be a subspace of $V$. Show that there is a subspace $U$ of $V$ such that $V = W + U$ and
\( W \cap U = \{0\} \). (\( U \) is called a direct complement of \( W \) in \( V \). We write then \( V = W \oplus U \) and call \( V \) the direct sum of \( W \) and \( U \).)

2. Is \( \{(1,1,1), (1,1,0), (1,0,0)\} \) an \( \mathbb{R} \)-basis of \( \mathbb{R}^3 \)?

3. Is \( \{(1,2,0), (0,0,1), (\bar{2},1,0)\} \) a \( \mathbb{Z}_3 \)-basis of \( \mathbb{Z}_3^2 \)?

4. Find an \( \mathbb{R} \)-basis of \( \{f \in C^2([0,1]): f''(x) - 7f'(x) + 12f(x) = 0 \text{ for all } x \in [0,1]\} \).

5. Find all \( \mathbb{R} \)-linear mappings from \( \mathbb{R}^4 \) onto \( \mathbb{R}^5 \).

6. Find all \( \mathbb{Z}_2 \)-bases of \( \mathbb{Z}_2^3 \) and \( \mathbb{Z}_3 \)-bases of \( \mathbb{Z}_3^2 \).

7. Show that the vectors \( (1,2,1), (0,2,0), (1,2,1) \) and also the vectors \( (1,1,0), (1,0,1), (1,1,1) \) in \( \mathbb{O}^3 \) are linearly independent over \( \mathbb{O} \).

8. Let \( f_k(x) = \sin kx \) for \( x \in [0,1] \) \( (k = 1,2,3, \ldots) \). Prove that the functions \( \{f_1,f_2,f_3,\ldots\} \) in \( C^\infty([0,1]) \) are linearly independent over \( \mathbb{R} \).