BASES

BRANKO ĆURGUS

Throughout this note $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$ and $\mathcal{V}$ is a vector space over $\mathbb{F}$; $\mathbb{N}$ denotes the set of positive integers. For a finite set $\mathcal{S}$ by $\#\mathcal{S}$, or sometimes $\#(\mathcal{S})$, we denote the number of elements in $\mathcal{S}$.

1. Linear independence

Definition 1.1. If $m \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_m \in \mathbb{F}$ and $v_1, \ldots, v_m \in \mathcal{V}$, then

$$\alpha_1 v_1 + \cdots + \alpha_m v_m$$

is called a *linear combination* of vectors in $\mathcal{V}$. A linear combination is *trivial* if $\alpha_1 = \cdots = \alpha_m = 0$; otherwise it is a *nontrivial* linear combination. ♢

Definition 1.2. Let $\mathcal{A}$ be a nonempty subset of $\mathcal{V}$. The *span of* $\mathcal{A}$ is the set of all linear combinations of vectors in $\mathcal{A}$. The span of $\mathcal{A}$ is denoted by

$$\operatorname{span}(\mathcal{A}).$$

The span of the empty set is the trivial vector space $\{0_{\mathcal{V}}\}$; that is,

$$\operatorname{span}(\emptyset) = \{0_{\mathcal{V}}\}.$$

If

$$\mathcal{V} = \operatorname{span}(\mathcal{A}),$$

then $\mathcal{A}$ is said to be a *spanning set* for $\mathcal{V}$. ♢

It is useful to write down the definition of a span in set-builder notation. Let $\mathcal{A}$ be a nonempty subset of $\mathcal{V}$. Then

$$\operatorname{span}(\mathcal{A}) = \left\{ v \in \mathcal{V} : \begin{array}{l} \exists\, m \in \mathbb{N} \\ \exists\, \alpha_1, \ldots, \alpha_m \in \mathbb{F} \\ \exists\, u_1, \ldots, u_m \in \mathcal{A} \\ \text{such that } v = \sum_{k=1}^{m} \alpha_k u_k \end{array} \right\}$$

Theorem 1.3. *Let* $\mathcal{A} \subseteq \mathcal{V}$. *Then* $\operatorname{span}(\mathcal{A})$ *is a subspace of* $\mathcal{V}$.

Proof. Write a proof as an exercise. □

Proposition 1.4. *If* $\mathcal{U}$ *is a subspace of* $\mathcal{V}$ *and* $\mathcal{A} \subseteq \mathcal{U}$, *then* $\operatorname{span}(\mathcal{A}) \subseteq \mathcal{U}$.

Proof. Write a proof as an exercise. □

Date: January 16, 2025 at 14:31.

Definition 1.5. Let $m \in \mathbb{N}$. Vectors $v_1, \ldots, v_m \in \mathcal{V}$ are said to be *linearly dependent* if there exists $\alpha_1, \ldots, v_m \in \mathbb{F}$ such that

$$\exists k \in \{1, \ldots, m\} \ \text{ such that } \ \alpha_k \neq 0 \quad \wedge \quad \alpha_1 v_1 + \cdots + \alpha_m v_m = 0_{\mathcal{V}}.$$

A subset $\mathcal{A}$ of $\mathcal{V}$ is *linearly dependent* if there exist $m \in \mathbb{N}$ and distinct vectors $v_1, \ldots, v_m \in \mathcal{A}$ that are linearly dependent. ♢

Remark 1.6. The definition of linear dependence is equivalent to the following statement: Let $\mathcal{A} \subseteq \mathcal{V}$. The set $\mathcal{A}$ is linearly dependent if there exist $m \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_m \in \mathbb{F} \setminus \{0\}$ and distinct $v_1, \ldots, v_m \in \mathcal{A}$ such that

$$\alpha_1 v_1 + \cdots + \alpha_m v_m = 0_{\mathcal{V}}.$$

♢

Definition 1.7. Let $m \in \mathbb{N}$. Vectors $v_1, \ldots, v_m \in \mathcal{V}$ are said to be *linearly independent* if for all $\alpha_1, \ldots, v_m \in \mathbb{F}$ the following implication holds:

$$\alpha_1 v_1 + \cdots + \alpha_m v_m = 0_{\mathcal{V}} \quad \Rightarrow \quad \forall k \in \{1, \ldots, m\} \ \ \alpha_k = 0.$$

An infinite subset $\mathcal{A}$ of $\mathcal{V}$ is *linearly independent* if for each $m \in \mathbb{N}$ arbitrary distinct vectors $v_1, \ldots, v_m \in \mathcal{A}$ are linearly independent. The empty set is by definition linearly independent. ♢

Notice that the last two definitions are formal logical negations of each other. Notice also that the last two definitions can briefly be stated as follows: A set $\mathcal{A} \subseteq \mathcal{V}$ is linearly dependent if there exists a nontrivial linear combination of distinct vectors in $\mathcal{A}$ whose value is $0_{\mathcal{V}}$. A set $\mathcal{A} \subseteq \mathcal{V}$ is linearly independent if the only linear combination of distinct vectors in $\mathcal{A}$ whose value is $0_{\mathcal{V}}$ is the trivial linear combination.

The following proposition is an immediate consequence of the definitions.

Proposition 1.8. *Let $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{V}$. If $\mathcal{A}$ is linearly dependent, then $\mathcal{B}$ is linearly dependent. Equivalently, if $\mathcal{B}$ is linearly independent, then $\mathcal{A}$ is linearly independent.*

Proof. Write a proof as an exercise. □

Proposition 1.9. *Let $\mathcal{A}$ be a linearly independent subset of $\mathcal{V}$. Let $u \in \mathcal{V}$ be such that $u \notin \mathcal{A}$. Then $\mathcal{A} \cup \{u\}$ is linearly dependent if and only if $u \in \operatorname{span}(\mathcal{A})$. Equivalently, $\mathcal{A} \cup \{u\}$ is linearly independent if and only if $u \notin \operatorname{span}(\mathcal{A})$.*

Proof. Assume that $u \in \operatorname{span}(\mathcal{A})$. Then there exist $m \in \mathbb{N}, \alpha_1, \ldots, \alpha_m \in \mathbb{F}$ and distinct $v_1, \ldots, v_m \in \mathcal{A}$ such that $u = \sum_{j=1}^{m} \alpha_j v_j$. Then

$$1 \cdot u - \alpha_1 v_1 - \cdots - \alpha_m v_m = 0_{\mathcal{V}}.$$

Since $1 \neq 0$ and $u, v_1, \ldots, v_m \in \mathcal{A} \cup \{u\}$ this proves that $\mathcal{A} \cup \{u\}$ is linearly dependent.

Now assume that $\mathcal{A} \cup \{u\}$ is linearly dependent. Then there exist $m \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_m \in \mathbb{F}$ and distinct vectors $v_1, \ldots, v_m \in \mathcal{A} \cup \{u\}$ such that

$$\alpha_1 v_1 + \cdots + \alpha_m v_m = 0_{\mathcal{V}} \quad \text{and} \quad \alpha_k \neq 0 \ \text{ for some } \ k \in \{1, \ldots, m\}.$$

Since $\mathcal{A}$ is linearly independent it is not possible that $v_1, \ldots, v_m \in \mathcal{A}$. Thus, $u \in \{v_1, \ldots, v_m\}$. Hence $u = v_j$ for some $j \in \{1, \ldots, m\}$. Again, since $\mathcal{A}$ is linearly independent $\alpha_j = 0$ is not possible. Thus $\alpha_j \neq 0$ and consequently

$$u = v_j = -\frac{1}{\alpha_j} \sum_{\substack{i=1 \\ i \neq j}}^{m} \alpha_i \, v_i. \qquad \square$$

Proposition 1.10. *Let $\mathcal{B}$ be a nonempty subset of $\mathcal{V}$. Then $\mathcal{B}$ is linearly independent if and only if for every $u \in \mathcal{B}$ we have $u \notin \operatorname{span}\bigl(\mathcal{B} \setminus \{u\}\bigr)$. Equivalently, $\mathcal{B}$ is linearly dependent if and only if there exists $u \in \mathcal{B}$ such that $u \in \operatorname{span}\bigl(\mathcal{B} \setminus \{u\}\bigr)$.*

Proof. We first prove the implication:

$$\mathcal{B} \quad \text{linearly independent} \quad \Rightarrow \quad \forall\, u \in \mathcal{B} \quad u \notin \operatorname{span}\bigl(\mathcal{B} \setminus \{u\}\bigr).$$

Assume that $\mathcal{B}$ is linearly independent. Let $u \in \mathcal{B}$ be arbitrary. Then $\mathcal{B} \setminus \{u\}$ is linearly independent by Proposition 1.8. Now, with $\mathcal{A} = \mathcal{B} \setminus \{u\}$, since $\mathcal{B} = \mathcal{A} \cup \{u\}$ is linearly independent, Proposition 1.9 yields that $u \notin \operatorname{span}\bigl(\mathcal{B} \setminus \{u\}\bigr)$.

To prove the converse of the displayed implication we will prove the contrapositive of the converse. (In mathematical logic the contrapositive of the converse is called the inverse of the starting implication. Consequently, the converse and the inverse of an implication are equivalent.) That is we prove:

$$\mathcal{B} \quad \text{linearly dependent} \quad \Rightarrow \quad \exists\, u \in \mathcal{B} \text{ such that } u \in \operatorname{span}\bigl(\mathcal{B} \setminus \{u\}\bigr).$$

Assume that $\mathcal{B}$ is linearly dependent. Then there exist $m \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_m \in \mathbb{F}$ and distinct vectors $v_1, \ldots, v_m \in \mathcal{B}$ such that

$$\sum_{j=1}^{m} \alpha_j v_j = 0_{\mathcal{V}} \quad \text{and} \quad \alpha_k \neq 0 \ \text{ for some } \ k \in \{1, \ldots, m\}.$$

Consequently,

$$v_k = -\frac{1}{\alpha_k} \sum_{\substack{j=1 \\ j \neq k}}^{m} \alpha_j v_j,$$

and thus $v_k \in \operatorname{span}\bigl(\mathcal{B} \setminus \{v_k\}\bigr)$. $\square$

The following equivalence will sometimes be helpful.

Lemma 1.11. *Let $\mathcal{B}$ be a nonempty subset of $\mathcal{V}$ and $u \in \mathcal{B}$. Then*

$$\operatorname{span}\bigl(\mathcal{B} \setminus \{u\}\bigr) = \operatorname{span}(\mathcal{B}) \qquad \Leftrightarrow \qquad u \in \operatorname{span}\bigl(\mathcal{B} \setminus \{u\}\bigr).$$

With this lemma Proposition 1.10 can be restated as follows.

Corollary 1.12. *Let $\mathcal{B}$ be a nonempty subset of $\mathcal{V}$. Then $\mathcal{B}$ is linearly independent if and only if*

$$\forall u \in \mathcal{B} \quad \operatorname{span}\bigl(\mathcal{B} \setminus \{u\}\bigr) \subsetneq \operatorname{span}(\mathcal{B}) \quad \textit{(strict inclusion)}.$$

2. Finite dimensional vector spaces. Bases

Definition 2.1. A vector space $\mathcal{V}$ over $\mathbb{F}$ is *finite-dimensional* if there exists a finite subset $\mathcal{A}$ of $\mathcal{V}$ such that $\mathcal{V} = \operatorname{span}(\mathcal{A})$. A vector space which is not finite-dimensional is said to be *infinite-dimensional.* $\diamondsuit$

Since the empty set is finite and since $\operatorname{span}(\emptyset) = \{0_{\mathcal{V}}\}$, the trivial vector space $\{0_{\mathcal{V}}\}$ is finite-dimensional.

Definition 2.2. A linearly independent spanning set is called a *basis* of $\mathcal{V}$. $\diamondsuit$

The next theorem shows that each finite-dimensional vector space has a basis.

Theorem 2.3. *Let $\mathcal{V}$ be a finite-dimensional vector space over $\mathbb{F}$. Then $\mathcal{V}$ has a basis.*

Proof. If $\mathcal{V}$ is a trivial vector space its basis is the empty set. Let $\mathcal{V} \neq \{0_{\mathcal{V}}\}$ be a finite-dimensional vector space. Let $\mathcal{A}$ be a finite subset of $\mathcal{V}$ such that $\mathcal{V} = \operatorname{span}(\mathcal{A})$. Let $p = \#\mathcal{A}$. Set

$$\mathbb{K} = \bigl\{k \in \mathbb{N} : \exists \mathcal{C} \subseteq \mathcal{A} \text{ such that } \ k = \#\mathcal{C} \ \text{ and } \ \operatorname{span}(\mathcal{C}) = \mathcal{V}\bigr\}.$$

Since $p \in \mathbb{K}$, $\mathbb{K}$ is a nonempty set of positive integers. By the Well Ordering Axiom $\mathbb{K}$ has a minimum. Set $n = \min \mathbb{K}$. By the definition of $\mathbb{K}$ there exists $\mathcal{B} \subseteq \mathcal{V}$ such that $\#\mathcal{B} = n$ and $\operatorname{span}(\mathcal{B}) = \mathcal{V}$. Since $n = \min \mathbb{K}$ we have $n - 1 \notin \mathbb{K}$. Let $u \in \mathcal{B}$ be arbitrary. Then $\#(\mathcal{B} \setminus \{u\}) = n - 1$ and consequently

$$\operatorname{span}\bigl(\mathcal{B} \setminus \{u\}\bigr) \subsetneq \mathcal{V} = \operatorname{span}(\mathcal{B}). \quad \text{(strict inclusion)}$$

Corollary 1.12 implies that $\mathcal{B}$ is linearly independent. Thus $\mathcal{B}$ is a basis for $\mathcal{V}$. $\square$

The second proof of Theorem 2.3. If $\mathcal{V}$ is a trivial vector space its basis is the empty set. Let $\mathcal{V} \neq \{0_{\mathcal{V}}\}$ be a finite-dimensional vector space. Let $\mathcal{A}$ be a finite subset of $\mathcal{V}$ such that $\mathcal{V} = \operatorname{span}(\mathcal{A})$. Let $p = \#\mathcal{A}$. Set

$$\mathbb{K} = \bigl\{\#\mathcal{C} : \mathcal{C} \subseteq \mathcal{A} \ \text{ and } \ \mathcal{C} \text{ is linearly independent}\bigr\}.$$

We first prove that $1 \in \mathbb{K}$. Since $\mathcal{V} \neq \{0_{\mathcal{V}}\}$ there exists $v \in \mathcal{A}$ such that $v \neq 0_{\mathcal{V}}$. Set $\mathcal{C} = \{v\}$. Then clearly $\mathcal{C} \subseteq \mathcal{A}$ and $\mathcal{C}$ is linearly independent. Thus $\#\mathcal{C} = 1 \in \mathbb{K}$.

If $\mathcal{C} \subseteq \mathcal{A}$, then $\#\mathcal{C} \leq \#\mathcal{A} = p$. Thus $\mathbb{K} \subseteq \{0, 1, \ldots, p\}$. As a subset of a finite set the set $\mathbb{K}$ is finite. Thus $\mathbb{K}$ has a maximum. Set $n = \max \mathbb{K}$. Since $n \in \mathbb{K}$ there exists $\mathcal{B} \subseteq \mathcal{A}$ such that $\mathcal{B}$ is linearly independent and $n = \#\mathcal{B}$.

Next we will prove that $\operatorname{span}(\mathcal{B}) = \mathcal{V}$. In fact we will prove that $\mathcal{A} \subseteq \operatorname{span}(\mathcal{B})$. If $\mathcal{B} = \mathcal{A}$, then this is trivial. So Assume that $\mathcal{B} \subsetneq \mathcal{A}$ and let $u \in \mathcal{A} \setminus \mathcal{B}$ be arbitrary. Then

$$\#\bigl(\mathcal{B} \cup \{u\}\bigr) = n + 1 \quad \text{and} \quad \mathcal{B} \cup \{u\} \subseteq \mathcal{A}.$$

Since $n = \max \mathbb{K}$, $n+1 \notin \mathbb{K}$. Therefore $\mathcal{B} \cup \{u\}$ is linearly dependent. By Proposition 1.9 $u \in \operatorname{span}(\mathcal{B})$. Hence $\mathcal{A} \subseteq \operatorname{span}(\mathcal{B})$. By Proposition 1.4, $\mathcal{V} = \operatorname{span}(\mathcal{A}) \subseteq \operatorname{span}(\mathcal{B})$. Since $\operatorname{span}(\mathcal{B}) \subseteq \mathcal{V}$ is obvious, we proved that $\operatorname{span}(\mathcal{B}) = \mathcal{V}$. This proves that $\mathcal{B}$ is a basis of $\mathcal{V}$. □

The third proof of Theorem 2.3. We will reformulate Theorem 2.3 so that we can use the Mathematical induction. Let n be a nonnegative integer. Denote by $P(n)$ the following statement: If $\mathcal{V} = \operatorname{span}(\mathcal{A})$ and $\#\mathcal{A} = n$, then there exists linearly independent set $\mathcal{B} \subseteq \mathcal{A}$ such that $\mathcal{V} = \operatorname{span}(\mathcal{B})$.

First we prove that $P(0)$ is true. Assume that $\mathcal{V} = \operatorname{span}(\mathcal{A})$ and $\#\mathcal{A} = 0$. Then $\mathcal{A} = \emptyset$. Since $\emptyset$ is linearly independent we can take $\mathcal{B} = \mathcal{A} = \emptyset$.

Now let k be an arbitrary nonnegative integer and assume that $P(k)$ is true. That is we assume that the following implication is true: If $\mathcal{U} = \operatorname{span}(\mathcal{C})$ and $\#\mathcal{C} = k$, then there exists linearly independent set $\mathcal{D} \subseteq \mathcal{C}$ such that $\mathcal{U} = \operatorname{span}(\mathcal{D})$. This is the inductive hypothesis.

Next we will prove that $P(k+1)$ is true. Assume that $\mathcal{V} = \operatorname{span}(\mathcal{A})$ and $\#\mathcal{A} = k+1$. Let $u \in \mathcal{A}$ be arbitrary. Set $\mathcal{C} = \mathcal{A} \setminus \{u\}$. Then $\#\mathcal{C} = k$. Set $\mathcal{U} = \operatorname{span}(\mathcal{C})$. The inductive hypothesis $P(k)$ applies to the vector space $\mathcal{U}$. Thus we conclude that there exists a linearly independent set $\mathcal{D} \subseteq \mathcal{C}$ such that $\mathcal{U} = \operatorname{span}(\mathcal{D})$.

We distinguish two cases: Case 1. $u \in \mathcal{U} = \operatorname{span}(\mathcal{C})$ and Case 2. $u \notin \mathcal{U} = \operatorname{span}(\mathcal{C})$. In Case 1 we have $\mathcal{A} \subseteq \operatorname{span}(\mathcal{C})$. Therefore, by Proposition 1.4, $\mathcal{V} = \operatorname{span}(\mathcal{A}) \subseteq \mathcal{U} \subseteq \mathcal{V}$. Thus $\mathcal{V} = \mathcal{U}$ and we can take $\mathcal{B} = \mathcal{D}$ in this case. In Case 2, $u \notin \mathcal{U} = \operatorname{span}(\mathcal{D})$. Since $\mathcal{D}$ is linearly independent Proposition 1.9 yields that $\mathcal{D} \cup \{u\}$ is linearly independent. Set $\mathcal{B} = \mathcal{D} \cup \{u\}$. Since $\mathcal{U} = \operatorname{span}(\mathcal{C}) = \operatorname{span}(\mathcal{D}) \subseteq \operatorname{span}(\mathcal{B})$ we have that $\mathcal{C} \subseteq \operatorname{span}(\mathcal{B})$. Clearly $u \in \operatorname{span}(\mathcal{B})$. Consequently, $\mathcal{A} \subseteq \operatorname{span}(\mathcal{B})$. By Proposition 1.4 $\mathcal{V} = \operatorname{span}(\mathcal{A}) \subseteq \operatorname{span}(\mathcal{B}) \subseteq \mathcal{V}$. Thus $\mathcal{V} = \operatorname{span}(\mathcal{B})$. As proved earlier $\mathcal{B}$ is linearly independent and $\mathcal{B} \subseteq \mathcal{A}$. This proves $P(k+1)$ and completes the proof. □

Notice that in the proof of Theorem 2.3 we have proved the following proposition.

Proposition 2.4. *Let $\mathcal{V}$ be a vector space over $\mathbb{F}$ and let $\mathcal{A} \subseteq \mathcal{V}$ be a finite subset of $\mathcal{V}$ such that $\mathcal{V} = \operatorname{span}(\mathcal{A})$. Then there exists a basis $\mathcal{B}$ for $\mathcal{V}$ such that $\mathcal{B} \subseteq \mathcal{A}$.*

3. Dimension

Theorem 3.1 (The Steinitz Exchange Lemma)**.** *Let $\mathcal{V}$ be a vector space over $\mathbb{F}$. Let $\mathcal{A}$ and $\mathcal{B}$ be finite subsets of $\mathcal{V}$ such that $\mathcal{A}$ spans $\mathcal{V}$ and $\mathcal{B}$ is linearly independent. Then $\#\mathcal{B} \leq \#\mathcal{A}$ and there exists $\mathcal{C} \subseteq \mathcal{A}$ such that $\#\mathcal{C} = \#\mathcal{A} - \#\mathcal{B}$ and $\mathcal{B} \cup \mathcal{C}$ spans $\mathcal{V}$.*

Proof. Let $\mathcal{A} \subseteq \mathcal{V}$ be a finite spanning set for $\mathcal{V}$ such that $\#\mathcal{A} = p$.

The proof is by mathematical induction on $m = \#\mathcal{B}$. Since the empty set is linearly independent the statement is true for $m = 0$. The statement is trivially true in this case. (You should do a proof of the case $m = 1$ as an exercise.)

Now let k be an arbitrary nonnegative integer and assume that the following statement (the inductive hypothesis) is true: If $\mathcal{D} \subseteq \mathcal{V}$ is a linearly independent set such that $\#\mathcal{D} = k$, then $k \leq p$ and there exists $\mathcal{E} \subseteq \mathcal{A}$ such that $\mathcal{E} = p - k$ and $\mathcal{D} \cup \mathcal{E}$ is a spanning set for $\mathcal{V}$.

To prove the inductive step we will prove the following statement: If $\mathcal{B} \subseteq \mathcal{V}$ is a linearly independent set such that $\mathcal{B} = k+1$, then $k+1 \leq p$ and there exists $\mathcal{C} \subseteq \mathcal{A}$ such that $\#\mathcal{C} = p - k - 1$ and $\mathcal{B} \cup \mathcal{C}$ is a spanning set for $\mathcal{V}$.

Assume that $\mathcal{B} \subseteq \mathcal{V}$ is a linearly independent set such that $\#\mathcal{B} = k + 1$. Let $u \in \mathcal{B}$ be arbitrary. Set $\mathcal{D} = \mathcal{B} \setminus \{u\}$. Since $\mathcal{B} = \mathcal{D} \cup \{u\}$ is linearly independent, by Proposition 1.10 we have $u \notin \operatorname{span}(\mathcal{D})$. Also, $\mathcal{D}$ is linearly independent and $\#\mathcal{D} = k$. The inductive hypothesis implies that $k \leq p$ and there exists $\mathcal{E} \subseteq \mathcal{A}$ such that $\#\mathcal{E} = p - k$ and $\mathcal{D} \cup \mathcal{E}$ is a spanning set for $\mathcal{V}$. Since $\mathcal{D} \cup \mathcal{E}$ is a spanning set for $\mathcal{V}$ and $u \in \mathcal{V}$, u can be written as a linear combination of vectors in $\mathcal{D} \cup \mathcal{E}$. But, as we noticed earlier, $u \notin \operatorname{span}(\mathcal{D})$. Thus, $\mathcal{E} \neq \emptyset$. Hence, $p - k = \#\mathcal{E} \geq 1$. Consequently, $k+1 \leq p$ is proved. Since $u \in \operatorname{span}\bigl(\mathcal{D} \cup \mathcal{E}\bigr)$, there exist $i, j \in \mathbb{N}$ and $u_1, \ldots, u_i \in \mathcal{D}$ and $v_1, \ldots, v_j \in \mathcal{E}$ and $\alpha_1, \ldots, \alpha_i, \beta_1, \ldots, \beta_j \in \mathbb{F}$ such that

$$u = \alpha_1 u_1 + \cdots + \alpha_i u_i + \beta_1 v_1 + \cdots + \beta_j v_j.$$

(If $\mathcal{D} = \emptyset$, then $i = 0$ and the vectors from $\mathcal{D}$ are not present in the above linear combination.) Since $u \notin \operatorname{span}(\mathcal{D})$ at least one of $\beta_1, \ldots, \beta_j \in \mathbb{F}$ is nonzero. But, by dropping v-s with zero coefficients we can assume that all $\beta_1, \ldots, \beta_j \in \mathbb{F}$ are nonzero. Then

$$v_1 = \frac{1}{\beta_1}\bigl(u - \alpha_1 u_1 - \cdots - \alpha_i u_i - \beta_2 v_2 - \cdots - \beta_j v_j\bigr).$$

Now set $\mathcal{C} = \mathcal{E} \setminus \{v_1\}$. Then $\#\mathcal{C} = p-k-1$. Notice that $u, u_1, \ldots, u_i \in \mathcal{B}$ and $v_2, \ldots, v_j \in \mathcal{C}$; so the last displayed equality implies that $v_1 \in \operatorname{span}\bigl(\mathcal{B} \cup \mathcal{C}\bigr)$. Since $\mathcal{E} = \mathcal{C} \cup \{v_1\}$ and $\mathcal{D} \subseteq \mathcal{B}$, it follows that $\mathcal{D} \cup \mathcal{E} \subseteq \operatorname{span}\bigl(\mathcal{B} \cup \mathcal{C}\bigr)$. Therefore,

$$\mathcal{V} = \operatorname{span}\bigl(\mathcal{D} \cup \mathcal{E}\bigr) \subseteq \operatorname{span}\bigl(\mathcal{B} \cup \mathcal{C}\bigr).$$

Hence, $\operatorname{span}\bigl(\mathcal{B} \cup \mathcal{C}\bigr) = \mathcal{V}$ and the proof is complete. □

The following corollary is a direct logical consequence of the Steinitz exchange lemma. It is in fact a partial contrapositive of the lemma.

Corollary 3.2. *Let $\mathcal{B}$ be a finite subset of $\mathcal{V}$. If $\mathcal{V}$ is a finite-dimensional vector space over $\mathbb{F}$, then there exists $p \in \mathbb{N}$ such that $\#\mathcal{B} > p$ implies $\mathcal{B}$ is linearly dependent.*

Proof. Assume that $\mathcal{B}$ is a finite subset of $\mathcal{V}$ and $\mathcal{V}$ is a finite-dimensional vector space over $\mathbb{F}$. Then there exists a finite subset $\mathcal{A}$ of $\mathcal{V}$ such that $\mathcal{V} = \operatorname{span}(\mathcal{A})$. Set $p = \#\mathcal{A}$. Then the Steinitz exchange lemma yields the following implication: If $\mathcal{B}$ is linearly independent, then $\#\mathcal{B} \leq p$. The contrapositive of the last implication is the claim of the corollary. □

Corollary 3.3. *Let $\mathcal{V}$ be a finite-dimensional vector space over $\mathbb{F}$. If $\mathcal{C}$ is an infinite subset of $\mathcal{V}$, then $\mathcal{C}$ is linearly dependent.*

Proof. Let $p \in \mathbb{N}$ be a number whose existence has been proved in Corollary 3.2. Let $\mathcal{C}$ be an infinite subset of $\mathcal{V}$. Since $\mathcal{C}$ is infinite it has a finite subset $\mathcal{A}$ such that $\#\mathcal{A} = p + 1$. Corollary 3.2 yields that $\mathcal{A}$ is linearly dependent. Since $\mathcal{A} \subseteq \mathcal{C}$, by Proposition 1.8, $\mathcal{C}$ is linearly dependent. □

Theorem 3.4. *Let $\mathcal{V}$ be a finite-dimensional vector space and let $\mathcal{B}$ and $\mathcal{C}$ be bases of $\mathcal{V}$. Then both $\mathcal{B}$ and $\mathcal{C}$ are finite sets and $\#\mathcal{B} = \#\mathcal{C}$.*

Proof. Let $\mathcal{B}$ and $\mathcal{C}$ be bases of $\mathcal{V}$. Since both $\mathcal{B}$ and $\mathcal{C}$ are linearly independent Corollary 3.3 implies that they are finite. Now we can apply the Steinitz exchange lemma to the finite spanning set $\mathcal{B}$ and the finite linearly independent set $\mathcal{C}$. We conclude that $\#\mathcal{C} \leq \#\mathcal{B}$. Applying again the Steinitz exchange lemma to the finite spanning set $\mathcal{C}$ and the finite linearly independent set $\mathcal{B}$ we conclude that $\#\mathcal{B} \leq \#\mathcal{C}$. Thus $\#\mathcal{B} = \#\mathcal{C}$. □

Definition 3.5. The *dimension* of a finite-dimensional vector space is the number of vectors in its basis. The dimension of a vector space $\mathcal{V}$ is denoted by $\dim \mathcal{V}$. ♢

The following corollary restates a part of Theorem 3.1 in terms of the dimension.

Corollary 3.6. *Let $\mathcal{V}$ be a finite-dimensional vector space over $\mathbb{F}$. Let $\mathcal{A}$ and $\mathcal{B}$ be finite subsets of $\mathcal{V}$. The following statements hold.*

(a) *If $\operatorname{span}(\mathcal{A}) = \mathcal{V}$, then $\#\mathcal{A} \geq \dim \mathcal{V}$.*
(b) *If $\mathcal{B}$ is linearly independent, then $\#\mathcal{B} \leq \dim \mathcal{V}$.*
(c) *If $\#\mathcal{A} < \dim \mathcal{V}$, then $\operatorname{span}(\mathcal{A}) \subsetneq \mathcal{V}$.*
(d) *$\#\mathcal{B} > \dim \mathcal{V}$, then $\mathcal{B}$ is linearly dependent.*

Proposition 3.7. *Let $\mathcal{V}$ be a finite-dimensional vector space over $\mathbb{F}$ and let $\mathcal{B}$ he a finite subset of $\mathcal{V}$. Then any two of the following three statements imply the remaining one.*

(a) *$\#\mathcal{B} = \dim \mathcal{V}$.*
(b) *$\operatorname{span}(\mathcal{B}) = \mathcal{V}$.*
(c) *$\mathcal{B}$ is linearly independent.*

Proof. The easiest implication is: (b) and (c) imply (a). This is the definition of the dimension.

Next we prove the implication (a) and (b) imply (c). Assume (a) and (b). If $\mathcal{B}$ is an empty set, then by definition it is linearly independent, that is (c) holds in this case. Assume now that $\mathcal{B} \neq \emptyset$. Let $u \in \mathcal{B}$ be arbitrary. Then $\#\big(\mathcal{B} \setminus \{u\}\big) < \dim \mathcal{V}$, so Corollary 3.6(c) yields $\operatorname{span}(\mathcal{B} \setminus \{u\}) \subsetneq \mathcal{V}$. Hence, for every $u \in \mathcal{B}$ we have that $\operatorname{span}(\mathcal{B} \setminus \{u\}) \subsetneq \operatorname{span}(\mathcal{B})$, which, by Proposition 1.10, implies that $\mathcal{B}$ is linearly independent.

Now assume (a) and (c). Let $\mathcal{A}$ be a basis of $\mathcal{V}$. By the Steinitz exchange lemma there exists $\mathcal{C} \subseteq \mathcal{A}$ such that $\#\mathcal{C} = \#\mathcal{A} - \#\mathcal{B} = 0$ and $\operatorname{span}\bigl(\mathcal{B} \cup \mathcal{C}\bigr) = \mathcal{V}$. Since $\mathcal{C} = \emptyset$, (b) follows. □

Remark 3.8. Notice that Corollary 3.6(a) and Proposition 3.7 imply that a finite spanning set for $\mathcal{V}$ is a basis if and only if it has the smallest possible cardinality. Similarly, Corollary 3.6(b) and Proposition 3.7 imply that in a finite-dimensional vector space a linearly independent subset is a basis if and only if it has the largest possible cardinality. ♢

In the following proposition we characterize infinite-dimensional vector spaces.

Proposition 3.9. *Let $\mathcal{V}$ be a vector space over $\mathbb{F}$. Set $\mathcal{A}_0 = \emptyset$. The following statements are equivalent.*

(a) *The vector space $\mathcal{V}$ over $\mathbb{F}$ is infinite-dimensional.*
(b) *For every $n \in \mathbb{N}$ there exists linearly independent set $\mathcal{A}_n \subseteq \mathcal{V}$ such that $\#(\mathcal{A}_n) = n$ and $\mathcal{A}_{n-1} \subsetneq \mathcal{A}_n$.*
(c) *There exists an infinite linearly independent subset of $\mathcal{V}$.*

Proof. We first prove (a)⇒(b). Assume (a). For $n \in \mathbb{N}$, denote by $P(n)$ the following statement:

$$\text{There exists linearly independent set } \mathcal{A}_n \subseteq \mathcal{V} \text{ such that } \#(\mathcal{A}_n) = n \text{ and } \mathcal{A}_{n-1} \subsetneq \mathcal{A}_n.$$

We will prove that $P(n)$ holds for every $n \in \mathbb{N}$. Mathematical induction is a natural tool here. Since the space $\{0_\mathcal{V}\}$ is finite-dimensional, we have $\mathcal{V} \neq \{0_\mathcal{V}\}$. Therefore there exists $v \in \mathcal{V}$ such that $v \neq 0_\mathcal{V}$. Set $\mathcal{A}_1 = \{v\}$ and the proof of $P(1)$ is complete. Let $k \in \mathbb{N}$ and assume that $P(k)$ holds. That is assume that there exists linearly independent set $\mathcal{A}_k \subseteq \mathcal{V}$ such that $\#(\mathcal{A}_k) = k$. Since $\mathcal{V}$ is an infinite-dimensional, $\operatorname{span}(\mathcal{A}_k)$ is a proper subset of $\mathcal{V}$. Therefore there exists $u \in \mathcal{V}$ such that $u \notin \operatorname{span}(\mathcal{A}_k)$. Since $\mathcal{A}_k$ is also linearly independent, Proposition 1.9 implies that $\mathcal{A}_k \cup \{u\}$ is linearly independent. Set $\mathcal{A}_{k+1} = \mathcal{A}_k \cup \{u\}$. Then, since $\#(\mathcal{A}_{k+1}) = k+1$ and $\mathcal{A}_k \subset \mathcal{A}_{k+1}$, the statement $P(k+1)$ is proved. This proves (b).

Now we prove (b)⇒(c). Assume (b) and set $\mathcal{C} = \cup\bigl\{\mathcal{A}_n : n \in \mathbb{N}\bigr\}$. Then $\mathcal{C}$ is infinite. To prove that $\mathcal{C}$ is linearly independent, let $m \in \mathbb{N}$ be arbitrary and let $v_1, \dots, v_m$ be distinct vectors in $\mathcal{C}$ and let $\alpha_1, \dots, \alpha_m \in \mathbb{F}$ such that

$$\alpha_1 v_1 + \cdots + \alpha_m v_m = 0_\mathcal{V}.$$

By the definition of $\mathcal{C}$, for every $k \in \{1, \dots, m\}$ there exists $n_k \in \mathbb{N}$ such that $v_k \in \mathcal{A}_{n_k}$. Set $q = \max\bigl\{n_k : k \in \{1, \dots, m\}\bigr\}$. By the inclusion property of the sequence $\mathcal{A}_n$, we have $\mathcal{A}_{n_k} \subseteq \mathcal{A}_q$ for all $k \in \{1, \dots, m\}$. Therefore, $v_k \in \mathcal{A}_q$ for all $k \in \{1, \dots, m\}$. Since the set $\mathcal{A}_q$ is linearly independent we conclude that $\alpha_k = 0_\mathbb{F}$ for all $k \in \{1, \dots, m\}$. This proves (c).

The implication (c)⇒(a) is a partial contrapositive of Corollary 3.3. This completes the proof. □

4. Subspaces

Proposition 4.1. *Let $\mathcal{U}$ be a subspace of $\mathcal{V}$. If $\mathcal{U}$ is infinite-dimensional, then $\mathcal{V}$ is infinite-dimensional. Equivalently, if $\mathcal{V}$ is finite-dimensional, then $\mathcal{U}$ is finite-dimensional. (In plain English, every subspace of a finite-dimensional vector space is finite-dimensional.)*

Proof. Assume that $\mathcal{U}$ is infinite-dimensional. Then, by the sufficient part of Proposition 3.9, for every $n \in \mathbb{N}$ there exists $\mathcal{A} \subseteq \mathcal{U}$ such that $\#\mathcal{A} = n$ and $\mathcal{A}$ is linearly independent. Since $\mathcal{U} \subseteq \mathcal{V}$, we have that for every $n \in \mathbb{N}$ there exists $\mathcal{A} \subseteq \mathcal{V}$ such that $\#\mathcal{A} = n$ and $\mathcal{A}$ is linearly independent. Now by the necessary part of Proposition 3.9 we conclude that $\mathcal{V}$ is infinite-dimensional. □

Theorem 4.2. *Let $\mathcal{V}$ be a finite-dimensional vector space and let $\mathcal{U}$ be a subspace of $\mathcal{V}$. Then there exists a subspace $\mathcal{W}$ of $\mathcal{V}$ such that $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$.*

Proof. Let $\mathcal{B}$ be a basis of $\mathcal{V}$ and let $\mathcal{A}$ a basis of $\mathcal{U}$. By Proposition 4.1, the Steinitz exchange lemma applies to the finite spanning set $\mathcal{B}$ and the finite linearly independent set $\mathcal{A}$. Consequently, there exists $\mathcal{C} \subseteq \mathcal{B}$ such that $\#\mathcal{C} = \#\mathcal{B} - \#\mathcal{A}$ and such that $\operatorname{span}(\mathcal{A} \cup \mathcal{C}) = \mathcal{V}$. Applying the Steinitz exchange lemma again to the linearly independent set $\mathcal{B}$ and the spanning set $\mathcal{A} \cup \mathcal{C}$ we conclude that $\#(\mathcal{A} \cup \mathcal{C}) \geq \#\mathcal{B}$. Since clearly $\#(\mathcal{A} \cup \mathcal{C}) \leq \#\mathcal{A} + \#\mathcal{C} = \#\mathcal{B}$ we have $\#(\mathcal{A} \cup \mathcal{C}) = \#\mathcal{A} + \#\mathcal{C} = \#\mathcal{B} = \dim \mathcal{V}$. Now the statement (a) and (b) imply (c) from Proposition 3.7 yields that $\mathcal{A} \cup \mathcal{C}$ is a basis of $\mathcal{V}$. Set $\mathcal{W} = \operatorname{span}(\mathcal{C})$. Then, since $\mathcal{A} \cup \mathcal{C}$ is a basis of $\mathcal{V}$, $\mathcal{V} = \mathcal{U} + \mathcal{W}$. It is not difficult to show that $\mathcal{U} \cap \mathcal{W} = \{0_{\mathcal{V}}\}$. Thus $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$. This proves the theorem. □

Lemma 4.3. *Let $\mathcal{V}$ be a finite-dimensional vector space and let $\mathcal{U}$ and $\mathcal{W}$ be subspaces of $\mathcal{V}$ such that $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$. Then $\dim \mathcal{V} = \dim \mathcal{U} + \dim \mathcal{W}$.*

Proof. Let $\mathcal{A}$ and $\mathcal{B}$ be basis of $\mathcal{U}$ and $\mathcal{W}$ respectively. Using $\mathcal{V} = \mathcal{U} + \mathcal{W}$, it can be proved that $\mathcal{A} \cup \mathcal{B}$ spans $\mathcal{V}$. Using $\mathcal{U} \cap \mathcal{W} = \{0_{\mathcal{V}}\}$, it can be shown that $\mathcal{A} \cup \mathcal{B}$ is linearly independent and $\mathcal{A} \cap \mathcal{B} = \emptyset$. Therefore $\mathcal{A} \cup \mathcal{B}$ is a basis of $\mathcal{V}$ and consequently $\dim \mathcal{V} = \#(\mathcal{A} \cup \mathcal{B}) = \#\mathcal{A} + \#\mathcal{B} = \dim \mathcal{U} + \dim \mathcal{V}$. □

Theorem 4.4. *Let $\mathcal{V}$ be a finite-dimensional vector space and let $\mathcal{U}$ and $\mathcal{W}$ be subspaces of $\mathcal{V}$ such that $\mathcal{V} = \mathcal{U} + \mathcal{W}$. Then*

$$\dim \mathcal{V} = \dim \mathcal{U} + \dim \mathcal{W} - \dim(\mathcal{U} \cap \mathcal{W}).$$

Proof. Since $\mathcal{U} \cap \mathcal{W}$ is a subspace of $\mathcal{U}$, Theorem 4.2 implies that there exists a subspace $\mathcal{U}_1$ of $\mathcal{U}$ such that

$$\mathcal{U} = \mathcal{U}_1 \oplus (\mathcal{U} \cap \mathcal{W}) \quad \text{and} \quad \dim \mathcal{U} = \dim \mathcal{U}_1 + \dim(\mathcal{U} \cap \mathcal{W}).$$

Similarly, there exists a subspace $\mathcal{W}_1$ of $\mathcal{W}$ such that $\mathcal{W} = \mathcal{W}_1 \oplus (\mathcal{U} \cap \mathcal{W})$ and $\dim \mathcal{W} = \dim \mathcal{W}_1 + \dim(\mathcal{U} \cap \mathcal{W})$. Next we will prove that $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}_1$. Let $v \in \mathcal{V}$ be arbitrary. Since $\mathcal{V} = \mathcal{U} + \mathcal{W}$ there exist $u \in \mathcal{U}$ and $w \in \mathcal{W}$ such that $v = u + w$. Since $\mathcal{W} = \mathcal{W}_1 \oplus (\mathcal{U} \cap \mathcal{W})$ there exist $w_1 \in \mathcal{W}_1$ and

$x \in \mathcal{U} \cap \mathcal{W}$ such that $w = w_1 + x$. Then $v = u + w_1 + x = (u + x) + w_1$. Since $u + x \in \mathcal{U}$ this proves that $\mathcal{V} = \mathcal{U} + \mathcal{W}_1$. Clearly $\mathcal{U} \cap \mathcal{W}_1 \subseteq \mathcal{U} \cap \mathcal{W}$ and $\mathcal{U} \cap \mathcal{W}_1 \subseteq \mathcal{W}_1$. Thus,

$$\mathcal{U} \cap \mathcal{W}_1 \subseteq \big(\mathcal{U} \cap \mathcal{W}\big) \cap \mathcal{W}_1 = \{0_{\mathcal{V}}\}.$$

Hence, $\mathcal{U} \cap \mathcal{W}_1 = \{0_{\mathcal{V}}\}$. This proves $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}_1$. By Lemma 4.3, $\dim \mathcal{V} = \dim \mathcal{U} + \dim \mathcal{W}_1 = \dim \mathcal{U} + \dim \mathcal{W} - \dim\big(\mathcal{U} \cap \mathcal{V}\big)$. This completes the proof. □

Combining the previous theorem and Lemma 4.3 we get the following corollary.

Corollary 4.5. *Let $\mathcal{V}$ be a finite-dimensional vector space and let $\mathcal{U}$ and $\mathcal{W}$ be subspaces of $\mathcal{V}$ such that $\mathcal{V} = \mathcal{U} + \mathcal{W}$. Then the sum $\mathcal{U} + \mathcal{W}$ is direct if and only if* $\dim \mathcal{V} = \dim \mathcal{U} + \dim \mathcal{W}$.

The previous corollary holds for any number of subspaces of $\mathcal{V}$. The proof is by mathematical induction on the number of subspaces.

Proposition 4.6. *Let $\mathcal{V}$ be a finite-dimensional vector space and let $\mathcal{U}_1$, …, $\mathcal{U}_m$ be subspaces of $\mathcal{V}$ such that $\mathcal{V} = \mathcal{U}_1 + \cdots + \mathcal{U}_m$. Then the sum $\mathcal{U}_1 + \cdots + \mathcal{U}_m$ is direct if and only if* $\dim \mathcal{V} = \dim \mathcal{U}_1 + \cdots + \dim \mathcal{U}_m$.

5. Example

Let $\mathbb{F}[x]$ be the vector space of all polynomials with coefficients in $\mathbb{F}$. We consider $\mathbb{F}[x]$ as a subspace of $\mathbb{F}^{\mathbb{F}}$. In fact,

$$\mathbb{F}[x] = \operatorname{span}\big(\{1\} \cup \big\{x^n : n \in \mathbb{N}\big\}\big).$$

Here, 1 stands for the constant polynomial whose range is $\{1\}$.

First we deduce some useful formulas with power functions.

For all $n \in \mathbb{N}$ and all $x, y \in \mathbb{C}$ we have

$$x^{n+1} - y^{n+1} = (x - y) \sum_{k=0}^{n} x^k y^{n-k}. \tag{5.1}$$

The proof of (5.1) is an exercise in summation. We calculate

$$\begin{aligned}
(x - y) \sum_{k=0}^{n} x^k y^{n-k} &= \sum_{k=0}^{n} x^{k+1} y^{n-k} - \sum_{k=0}^{n} x^k y^{n-k+1} \\
&= x^{n+1} + \sum_{k=0}^{n-1} x^{k+1} y^{n-k} - \sum_{k=1}^{n} x^k y^{n-k+1} - y^{n+1} \\
&= x^{n+1} - y^{n+1} + \sum_{j=1}^{n} x^j y^{n-j+1} - \sum_{k=1}^{n} x^k y^{n-k+1} \\
&= x^{n+1} - y^{n+1}.
\end{aligned}$$

A generalization of the formula (5.1) is as follows. For $n \in \mathbb{N}$ and $\alpha_0, \dots, \alpha_{n+1} \in \mathbb{C}$ set

$$p(x) = \sum_{k=0}^{n+1} \alpha_k x^k.$$

Then for all $x, y \in \mathbb{C}$ we have

$$p(x) - p(y) = (x - y) \sum_{j=0}^{n} x^j \sum_{k=j}^{n} \alpha_{k+1} y^{k-j}, \tag{5.2}$$

or, equivalently,

$$p(x) - p(y) = (x - y) \sum_{k=0}^{n} \alpha_{k+1} \sum_{j=0}^{k} x^j y^{k-j}. \tag{5.3}$$

The proof of (5.2) is an exercise in summation. We calculate

$$\begin{aligned}
&(x-y)\sum_{j=0}^{n} x^j \sum_{k=j}^{n} \alpha_{k+1} y^{k-j} \\
&\quad = \sum_{j=0}^{n} x^{j+1} \sum_{k=j}^{n} \alpha_{k+1} y^{k-j} - \sum_{j=0}^{n} x^j \sum_{k=j}^{n} \alpha_{k+1} y^{k-j+1} \\
&\quad = \sum_{k=0}^{n} \alpha_{k+1} \sum_{j=0}^{k} x^{j+1} y^{k-j} - \sum_{j=0}^{n} \alpha_{k+1} \sum_{j=0}^{k} x^j y^{k-j+1} \\
&\quad = \sum_{k=0}^{n} \alpha_{k+1} \left(\sum_{j=0}^{k} x^{j+1} y^{k-j} - \sum_{j=0}^{k} x^j y^{k-j+1} \right) \\
&\quad = \sum_{k=0}^{n} \alpha_{k+1} \left(x^{k+1} - y^{k+1} + \sum_{j=0}^{k-1} x^{j+1} y^{k-j} - \sum_{j=1}^{k} x^j y^{k-j+1} \right) \\
&\quad = \sum_{k=0}^{n} \alpha_{k+1} \left(x^{k+1} - y^{k+1} + \sum_{j=1}^{k} x^{l} y^{k-l+1} - \sum_{j=1}^{k} x^j y^{k-j+1} \right) \\
&\quad = \sum_{k=0}^{n} \alpha_{k+1} \left(x^{k+1} - y^{k+1} \right) \\
&\quad = p(x) - p(y).
\end{aligned}$$

The formula in (5.2) gives a factorization of a polynomial $p(x)$ if $y = x_0$ is a zero of $p(x)$, that is if $p(x_0) = 0$. Substituting $y = x_0$ in (5.2) and using $p(x_0) = 0$ yields

$$p(x) = (x - x_0) \sum_{j=0}^{n} x^j \sum_{k=j}^{n} \alpha_{k+1} x_0^{k-j}. \tag{5.4}$$

Notice that on the right-hand side of (5.4) is a product of the linear factor $x - x_0$ and the polynomial

$$\begin{aligned} q(x) &= \sum_{j=0}^{n} x^j \sum_{k=j}^{n} \alpha_{k+1} x_0^{k-j} \\ &= \left(\sum_{k=0}^{n} \alpha_{k+1} x_0^{k} \right) + \cdots + (\alpha_n + \alpha_{n+1} x_0)\, x^{n-1} + \alpha_{n+1} x^n. \end{aligned}$$

If $\alpha_{n+1} \neq 0$, that is if $p(x)$ is a polynomial of degree $n+1$, then the polynomial $q(x)$ is a polynomial of degree n.

In conclusion, we have proved the following factorization theorem

Theorem 5.1 (4.6 p.122 in the textbook)**.** *Let $n \in \mathbb{N}$ and let $\alpha_0, \ldots, \alpha_{n+1} \in \mathbb{C}$ with $\alpha_{n+1} \neq 0$. Set*

$$p(x) = \sum_{k=0}^{n+1} \alpha_k x^k \in \mathbb{C}[x].$$

Let $x_0 \in \mathbb{C}$. Then $p(x_0) = 0$ if and only if there exists a polynomial $q(x)$ of degree n with the leading coefficient $\alpha_{n+1} \neq 0$ such that

$$p(x) = (x - x_0) q(x).$$

Proposition 5.2 (4.8 p.123 in the textbook)**.** *The following statement holds for all $n \in \mathbb{N}$ and for all $(\alpha_0, \ldots, \alpha_n) \in \mathbb{F}^{n+1}$:*

$$\alpha_n \neq 0 \quad \Rightarrow \quad \#\bigl\{x \in \mathbb{F} \ :\ \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n = 0\bigr\} \leq n. \tag{5.5}$$

Proof. We use the Mathematical Induction. For $n \in \mathbb{N}$ we consider the following propositional function of n:

$$P(n): \quad \forall (\alpha_0, \ldots, \alpha_n) \in \mathbb{F}^{n+1} \quad \text{(5.5) holds.}$$

Base case. $P(1)$ reads: For all $(\alpha_0, \alpha_1) \in \mathbb{F}^2$ the following implication holds:

$$\alpha_1 \neq 0 \quad \Rightarrow \quad \#\bigl\{x \in \mathbb{F} \ :\ \alpha_0 + \alpha_1 x = 0\bigr\} \leq 1.$$

Assume $\alpha_1 \neq 0$. It is straightforward to verify that

$$\bigl\{x \in \mathbb{F} \ :\ \alpha_0 + \alpha_1 x = 0\bigr\} = \bigl\{-\alpha_0/\alpha_1\bigr\}.$$

Since $\#\bigl\{-\alpha_0/\alpha_1\bigr\} = 1$, $P(1)$ holds. This proves the base case.

Inductive step. Let $m \in \mathbb{N}$ be arbitrary. We prove

$$P(m) \quad \Rightarrow \quad P(m+1).$$

Assume $P(m)$. (This is the **Inductive hypothesis**.) That is we assume the following: For all $(\beta_0, \ldots, \beta_m) \in \mathbb{F}^{m+1}$ the following implication holds:

$$\beta_m \neq 0 \quad \Rightarrow \quad \#\bigl\{x \in \mathbb{F} \ :\ \beta_0 + \beta_1 x + \cdots + \beta_m x^m = 0\bigr\} \leq m.$$

Next we will prove $P(m+1)$. That is, we will prove: For all $(\alpha_0, \dots, \alpha_{m+1}) \in \mathbb{F}^{m+2}$ the following implication holds:

$$\alpha_{m+1} \neq 0 \quad \Rightarrow \quad \#\{x \in \mathbb{F} \, : \, \alpha_0 + \alpha_1 x + \cdots + \alpha_{m+1} x^{m+1} = 0\} \leq m+1.$$

Here is a proof. Let $(\alpha_0, \dots, \alpha_{m+1}) \in \mathbb{F}^{m+2}$ be arbitrary. Assume $\alpha_{m+1} \neq 0$. We continue a proof by cases.

Case 1: $\forall x \in \mathbb{F} \quad \alpha_0 + \cdots + \alpha_{m+1} x^{m+1} \neq 0$.

In this case

$$\{x \in \mathbb{F} \, : \, \alpha_0 + \alpha_1 x + \cdots + \alpha_{m+1} x^{m+1} = 0\} = \emptyset.$$

Hence, $P(m+1)$ holds.

Case 2: $\exists x_0 \in \mathbb{F}$ such that $\alpha_0 + \cdots + \alpha_{m+1} x_0^{m+1} = 0$.

By the factorization theorem, Theorem 5.1, there exist $(\beta_0, \dots, \beta_m) \in \mathbb{F}^{m+1}$ such that $\beta_m = \alpha_{m+1} \neq 0$ and

$$\alpha_0 + \cdots + \alpha_{m+1} x^{m+1} = (x - x_0)\big(\beta_0 + \cdots + \beta_m x^m\big). \tag{5.6}$$

Since for all $\alpha, \beta \in \mathbb{F}$ we have $\alpha\beta = 0$ if and only if $\alpha = 0$ or $\beta = 0$, (5.6) implies that

$$\alpha_0 + \cdots + \alpha_{m+1} x^{m+1} = 0 \quad \Leftrightarrow \quad x = x_0 \ \vee \ \beta_0 + \cdots + \beta_m x^m = 0.$$

Consequently,

$$\begin{aligned} \{x \in \mathbb{F} \, : \, \alpha_0 + \alpha_1 x + \cdots + \alpha_{m+1} x^{m+1} = 0\} \\ = \{x_0\} \cup \{x \in \mathbb{F} \, : \, \beta_0 + \beta_1 x + \cdots + \beta_m x^m = 0\}, \end{aligned}$$

and therefore, first using properties of counting with finite sets and then the **Inductive hypothesis** we have

$$\begin{aligned} \#\{x \in \mathbb{F} \, : \, \alpha_0 + \alpha_1 x + \cdots + \alpha_{m+1} x^{m+1} = 0\} \\ \leq 1 + \#\{x \in \mathbb{F} \, : \, \beta_0 + \beta_1 x + \cdots + \beta_m x^m = 0\} \\ \leq 1 + m. \end{aligned}$$

This completes the proof of $P(m+1)$. □

Theorem 5.3. *Let D be an infinite subset of $\mathbb{F}$. The set of monomials*

$$\mathcal{M} = \{1\} \cup \{x^n : n \in \mathbb{N}\}$$

is a linearly independent set in $\mathbb{F}^D$.

Proof. We need to prove the following implication: for all $n \in \mathbb{N}$ and all $(\alpha_0, \dots, \alpha_n) \in \mathbb{F}^{n+1}$ the following implication holds:

$$\forall x \in D \quad \alpha_0 + \cdots + \alpha_n x^n = 0 \quad \Rightarrow \quad \forall k \in \{0, \dots, n\} \ \ \alpha_k = 0.$$

Let $n \in \mathbb{N}$ and $(\alpha_0, \dots, \alpha_n) \in \mathbb{F}^{n+1}$ be arbitrary. Let us prove the contrapositive of the preceding implication:

$$\exists k \in \{0, \dots, n\} \ \text{ s.t. } \ \alpha_k \neq 0 \quad \Rightarrow \quad \exists x \in D \ \text{ s.t. } \ \alpha_0 + \cdots + \alpha_n x^n \neq 0.$$

Assume that there exists $k \in \{0, \dots, n\}$ such that $\alpha_k \neq 0$. Set $l = \max\{k \in \{0, \dots, n\} : \alpha_k \neq 0\}$. We consider two cases. Case 1: $l = 0$. Then

$$\forall x \in D \quad \alpha_0 + \cdots + \alpha_n x^n = \alpha_0 \neq 0,$$

so, the contrapositive is proved in this case. Case 2: $l \geq 1$. Then by Proposition 5.2 there exists a subset A of $\mathbb{F}$ such that $\#A \leq l$ and

$$\forall x \in \mathbb{F} \setminus A \ \ \alpha_0 + \cdots + \alpha_l x^l \neq 0.$$

Since $D \subseteq \mathbb{F}$ is infinite and A is finite, the set $D \setminus A$ is nonempty and

$$\forall x \in D \setminus A \ \ \alpha_0 + \cdots + \alpha_l x^l \neq 0.$$

This proves the contrapositive in Case 2 and the proof is complete. □

6. Problems

Problem 6.1. Consider the vector space $\mathcal{V} = \mathbb{R}^{\mathbb{N}}$ of all real valued functions defined on $\mathbb{N}$ with the values in $\mathbb{R}$ over the scalar field $\mathbb{R}$. Simply says, this is the vector space of all real sequences. Consider the special sequences in $\mathbb{R}^{\mathbb{N}}$. For arbitrary $n \in \mathbb{N}$ defined the sequence $\phi_n \in \mathbb{R}^{\mathbb{N}}$ by

$$\forall k \in \mathbb{N} \quad \phi_n(k) = \begin{cases} 1, & \text{if} \quad k = n, \\ 0, & \text{if} \quad k \neq n. \end{cases}$$

Consider the set $\mathcal{A}$ of all such special sequences. That is

$$\mathcal{A} = \{g \in \mathbb{R}^{\mathbb{N}} : \exists n \in \mathbb{N} \ \text{ such that } \ g = \phi_n\} = \{\phi_n \in \mathbb{R}^{\mathbb{N}} : n \in \mathbb{N}\}.$$

(i) Prove that

$$\operatorname{span}(\mathcal{A}) = \{f \in \mathbb{R}^{\mathbb{N}} : \exists m \in \mathbb{N} \ \text{ such that } \ \forall k \in \mathbb{N} \ \ k > m \ \Rightarrow \ g(k) = 0\}.$$

(ii) Prove that the set $\mathcal{A}$ is linearly independent.

◊

Problem 6.2. Prove that $\mathcal{V}$ is finite dimensional if and only if all linearly independent subsets of $\mathcal{V}$ are finite. (Give a complete proof without citing propositions in this section. You can, of course, use ideas utilized in the proofs of this section.) ◊

Problem 6.3 (This is a challenging problem)**.** Let $\mathcal{V}$ be a finite-dimensional nonzero vector space $\mathcal{V}$ over $\mathbb{F}$. Let $n = \dim \mathcal{V}$ and let $\{v_1, \dots, v_n\}$ be a basis of $\mathcal{V}$. Let $\mathcal{U}$ and $\mathcal{W}$ be subspaces of $\mathcal{V}$ such that $\mathcal{V}$ is a direct sum of $\mathcal{U}$ and $\mathcal{W}$, that is

$$\mathcal{V} = \mathcal{U} \oplus \mathcal{W}.$$

Let

$$v_k = u_k + w_k, \quad \text{where} \quad u_k \in \mathcal{U}, \ \ w_k \in \mathcal{W} \ \text{ for all } \ k \in \{1, \dots, n\}.$$

Prove that there exist subsets $\mathbb{A}$ and $\mathbb{B}$ of $\{1, \dots, n\}$ such that:

$$\{1, \dots, n\} = \mathbb{A} \cup \mathbb{B} \quad \text{and} \quad \mathbb{A} \cap \mathbb{B} = \emptyset$$

and

$$\{u_k : k \in \mathbb{A}\} \text{ is a basis for } \mathcal{U} \quad \text{and} \quad \{w_k : k \in \mathbb{B}\} \text{ is a basis for } \mathcal{W}.$$

♢

Problem 6.4. This problem concerns the vector space $\mathbb{R}^{2\times 2}$ over $\mathbb{R}$. This vector space consists of all 2×2 matrices with entries from $\mathbb{R}$.

(a) Denote by $\mathcal{S}$ the subset of $\mathbb{R}^{2\times 2}$ which consists of all matrices A such that $\mathsf{A}^2 = 0$.

  (i) Is $\mathcal{S}$ a subspace of the vector space $\mathbb{R}^{2\times 2}$? Justify your answer.

  (ii) Describe all subspaces of $\mathbb{R}^{2\times 2}$ which contain $\mathcal{S}$. For each such subspace give a basis.

  (iii) Is there a two-dimensional subspace contained in $\mathcal{S}$?

  (iv) Each three-dimensional vector space over $\mathbb{R}$ can be identified with the three dimensional Euclidian space. In this way the language of Euclidian geometry can be used to describe subsets of a three-dimensional vector space over $\mathbb{R}$. I hope that in (aii) you found a three-dimensional subspace that contains $\mathcal{S}$. Use this fact to give a geometric description of the set $\mathcal{S}$. (Hint: Try to find a basis of a subspace from (aii) with respect to which the corresponding equation for $\mathcal{S}$ will be very simple.)

(b) Let $\mathsf{J} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Consider the set $\mathcal{S}$ of all matrices X in $\mathbb{R}^{2\times 2}$ such that $\mathsf{XJX} = \mathsf{0}$.

  (i) Is $\mathcal{S}$ a subspace of the vector space $\mathbb{R}^{2\times 2}$? Give detailed explanation of your answer.

  (ii) What is the dimension of the smallest subspace of $\mathbb{R}^{2\times 2}$ which contains $\mathcal{S}$? Give a basis of this subspace. Is this subspace uniquely determined?

  (iii) Is there a 2-dimensional subspace of $\mathbb{R}^{2\times 2}$ that is contained in $\mathcal{S}$?

  (iv) Use parts (bi), (bii) and (biii) of this problem to give a geometric description of the set $\mathcal{S}$. What is the equation of the set $\mathcal{S}$ with respect to the basis from (bii)? Try to find another basis for the subspace from (bii) with respect to which the corresponding equation of $\mathcal{S}$ will be very simple?

♢

Problem 6.5. Consider the vector space $\mathbb{R}^{\mathbb{R}}$ of all real valued functions defined on $\mathbb{R}$. This vector space is considered over the field $\mathbb{R}$. The purpose of this exercise is to study some special subspaces of the vector space $\mathbb{R}^{\mathbb{R}}$. Let $\omega \in \mathbb{R}$ be arbitrary. Consider the set

$$\mathcal{S}_\omega := \left\{ f \in \mathbb{R}^{\mathbb{R}} : \exists a, b \in \mathbb{R} \ \text{ such that } \ f(t) = a \sin(\omega t + b) \ \ \forall t \in \mathbb{R} \right\}.$$

(a) Do you see exceptional values for ω for which the set $\mathcal{S}_\omega$ is particularly simple? State them and explain why they are special. Here I used plural just in case that there are several special cases. However, it is possible that there is only one special case for ω.

(b) Prove that $\mathcal{S}_\omega$ is a subspace of $\mathbb{R}^\mathbb{R}$. (Hint: Except for the special case, this problem should be solved by writing the set $\mathcal{S}_\omega$ as a span of two linearly independent famous functions. One should use only basic trigonometry and polar coordinates.)

(c) For each $\omega \in \mathbb{R}$ find a basis for $\mathcal{S}_\omega$. Plot the function $\omega \mapsto \dim \mathcal{S}_\omega$ with $\omega \in \mathbb{R}$.

(d) For all $\psi, \omega \in \mathbb{R}$ calculate $\dim(\mathcal{S}_\psi \cap \mathcal{S}_\omega)$.

(e) Find all $\psi, \omega \in \mathbb{R}$ for which $\mathcal{S}_\psi \cup \mathcal{S}_\omega$ is a subspace of $\mathbb{R}^\mathbb{R}$.

(f) For all $\psi, \omega \in \mathbb{R}$ calculate $\dim\bigl(\mathcal{S}_\psi + \mathcal{S}_\omega\bigr)$.

$\Diamond$

Problem 6.6. Consider the vector space $\mathcal{P}_2 = \mathbb{R}[x]_{\leq 2}$ of all polynomials with real coefficients of degree at most 2. We consider $\mathcal{P}_2$ as a subspace of $\mathbb{R}^\mathbb{R}$.

Let $s \in \mathbb{R}$. We say that $p \in \mathcal{P}_2$ has a *vertex at* s if the following condition is satisfied

$$\bigl(\forall x \in \mathbb{R} \ \ p(x) \leq p(s)\bigr) \quad \vee \quad \bigl(\forall x \in \mathbb{R} \ \ p(s) \leq p(x)\bigr)$$

Notice that under this definition a constant polynomial has a vertex at every real number s.

Let $s \in \mathbb{R}$. Denote by $\mathcal{V}_s$ the subset of $\mathcal{P}_2$ which consists of all polynomials that have a vertex at s. In set-builder notation

$$\mathcal{V}_s = \bigl\{p \in \mathcal{P}_2 : p \text{ has a vertex at } s\bigr\}.$$

Let $t \in \mathbb{R}$. We say that $p \in \mathcal{P}_2$ has a *zero at* t if $p(t) = 0$. Notice that under this definition the zero polynomial has a zero at every real number t.

Let $t \in \mathbb{R}$. Denote by $\mathcal{Z}_t$ the subset of $\mathcal{P}_2$ which consists of all polynomials that have zero at t. In set-builder notation

$$\mathcal{Z}_t = \bigl\{p \in \mathcal{P}_2 : p(t) = 0\bigr\}.$$

(a) Let $t \in \mathbb{R}$ be an arbitrary (fixed) number. Prove that $\mathcal{Z}_t$ is a subspace of $\mathbb{P}_2$. Determine $\dim \mathcal{Z}_t$.

(b) Let $s \in \mathbb{R}$ be an arbitrary (fixed) number. Prove that $\mathcal{V}_s$ is a subspace of $\mathbb{P}_2$. Determine $\dim \mathcal{V}_s$.

(c) Let $s, t \in \mathbb{R}$ be given such that $s \neq t$. Describe the polynomials in each of the subspaces $\mathcal{Z}_s \cap \mathcal{Z}_t$, $\mathcal{V}_s \cap \mathcal{Z}_t$ and $\mathcal{V}_s \cap \mathcal{V}_t$. Determine the dimension for each of these subspaces.

(d) Let $s, t \in \mathbb{R}$ be given such that $s \neq t$. Find $u, v \in \mathbb{R}$ such that the equality $\mathcal{Z}_s \cap \mathcal{Z}_u = \mathcal{V}_v \cap \mathcal{Z}_t$ holds.

(e) Is the following statement true or false: For every one-dimensional subspace $\mathcal{U}$ of $\mathcal{P}_2$ there exists $t \in \mathbb{R}$ such that $\mathcal{U} \oplus \mathcal{Z}_t = \mathcal{P}_2$.

(f) Is the following statement true or false: There exists an one-dimension subspace $\mathcal{U}$ of $\mathcal{P}_2$ such that for all $t \in \mathbb{R}$ we have $\mathcal{U} \oplus \mathcal{Z}_t = \mathcal{P}_2$.

(g) Is the following statement true or false: For every one-dimensional subspace $\mathcal{U}$ of $\mathcal{P}_2$ there exists $s \in \mathbb{R}$ such that $\mathcal{U} \oplus \mathcal{V}_s = \mathcal{P}_2$.

(h) Is the following statement true or false: There exists an one-dimension subspace $\mathcal{U}$ of $\mathcal{P}_2$ such that for all $s \in \mathbb{R}$ we have $\mathcal{U} \oplus \mathcal{V}_s = \mathcal{P}_2$.

$\Diamond$