Linear Algebra, Geometry and Groups

Lasse Rempe
(University of Liverpool, Spring Semester 2006)

Contents

CHAPTER 0

Introduction

In mathematics, we often encounter specific objects with certain interesting structures (say that of calculations in the real or complex numbers or the geometry of 3-dimensional space). Upon such encounters, mathematicians frequently take an abstract point of view by formulating the basic properties of these structures as axioms and then studying their consequences. This procedure has several advantages, e.g.

(a) it is not necessary to repeatedly develop the same theory each time we encounter an object with similar properties, but can rather apply general theorems;
(b) we will be able to clearly identify the consequences of our axioms, advancing our understanding of the objects under consideration;
(c) we might discover interesting new mathematics along the way.

This idea, which should be familiar to those who have already encountered rings and fields, will play an important role in this course. We will define *vector spaces*, inspired by the Euclidean spaces $\mathbb{R}^n$, and study some general properties of these spaces and their structure-preserving ("linear") maps. We will then turn our attention to some geometric objects in whose study linear algebra proves very useful. Finally, we will study *groups*, another abstract class of objects which is pervasive in mathematics and beyond.

0.1. Operations

We will be mostly concerned with structures arising from *operations*, such as addition and multiplication (as for real numbers), scalar multiplication (in $\mathbb{R}^n$), etc.

For example, we will be in the following situation: given a set V and "rules" for addition and scalar multiplication, we want to say what it means for V to be a *vector space*.

To do this in a mathematically correct way, we should say what a "rule for addition" actually is. Let us think about this for a moment. Clearly addition is an operation which takes two elements $v, w \in V$ and produces another element (the sum $v + w$).

In other words, $+$ is a *function* from the set of pairs $V \times V$ into V, or in short:

$$+ : V \times V \to V.$$

So, in the following, if we write

"let $+ : V \times V \to V$ be an operation,"

this means that $+$ is some arbitrary rule for addition which is defined for any two elements of V, yielding another element of V.

Another example: what kind of object is the standard division in the real numbers? The answer is simple: we can divide any real number by any *nonzero* real number, so $\div$ is

a function

$$\div : \mathbb{R} \times \underbrace{\mathbb{R}^*}_{=\mathbb{R}\setminus\{0\}} \to \mathbb{R}.$$

Of course, for any operation defined in this way, we will still keep writing $v + w$, $x \cdot y$, etc., instead of $+(v, w)$, $\cdot(x, y)$ etc. Also, when we define a product operation $\cdot$, we will often abbreviate ab instead of $a \cdot b$, as usual.

CHAPTER 1

Vector Spaces

Throughout this chapter — unless specifically noted otherwise — let $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$. (Those who know about such things may also think of $\mathbb{K}$ as any arbitrary *field*.) Most of the time, you can think of $\mathbb{K}$ as being $\mathbb{R}$.

1.1. Basic definitions

1.1.1. Definition (Vector Space)**.**
A vector space *(over $\mathbb{K}$) is a tuple $(V, +, \cdot)$, where V is a set and $+ : V \times V \to V$ and $\cdot : \mathbb{K} \times V \to V$ are operations, with the following properties.*

(V1) *(*additive associativity*)* $u + (v + w) = (u + v) + w$ *for all* $u, v, w \in V$*;*
(V2) *(*additive commutativity*)* $u + v = v + u$ *for al* $lu, v \in V$*;*
(V3) *(*neutral element of addition*) there is an element* $0 \in V$ *such that, for every* $v \in V$*,* $v + 0 = v$*;*
(V4) *(*additive inverses*) for every* $v \in V$*, there is an element* $-v \in V$ *such that* $v + (-v) = 0$*;*
(V5) *(*associativity of scalar multiplication*) for every* $\lambda, \mu \in \mathbb{K}$ *and* $v \in V$*, we have* $(\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot v)$*;*
(V6) *(*neutrality of $1 \in \mathbb{K}$*) for all* $v \in V$*,* $1 \cdot v = v$*;*
(V7) *(*distributivity I*) for all* $\lambda, \mu \in \mathbb{K}$*,* $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$*;*
(V8) *(*distributivity II*) for all* $\lambda \in \mathbb{K}$ *and* $v, w \in V$*,* $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$*.*

Remark.

(a) Usually, $\mathbb{K}$ and the rules of addition and multiplication are implicitly clear, and we will simply say that "V is a vector space".
(b) When $+$ and $\cdot$ are given by some explicit formula, one should check — as well as properties (V1) to (V8) — that these are indeed operations of the required form; i.e., for every $v, w \in V$ and $\lambda \in \mathbb{K}$, $v + w$ and $\lambda \cdot v$ should be defined and belong to V!
(c) In (V7) and (V8), there are two different operations of addition and multiplication involved: the usual operations of $\mathbb{K}$, and addition and scalar multiplication in V. Usually, this does not cause any confusion, but one should be aware of it.
(d) It is an easy exercise to see that the neutral and inverse elements are given by $0 = 0 \cdot v$ and $-v = (-1) \cdot v$.

1.1.2. Examples. (a) $\mathbb{R}^n$ with the usual operations is a real vector space (i.e., a vector space over $\mathbb{R}$). Similarly $\mathbb{C}^n$ is a complex vector space.

(b) $V = \{(x_1, x_2, x_3) : x_1, x_2, x_3 \in \mathbb{Z}\}$ (with the usual operations) is *not* a real vector space, since $\cdot$ is not an operation of the required form: e.g., $\frac{1}{2} \cdot (1, 0, 0) \notin V$. (Similarly, $\mathbb{R}^n$ is not a complex vector space.)

(c) The space $\mathbb{R}^{n\times m}$ of real n-by-m matrices is a real vector space. (Its zero element is the matrix all of whose entries are 0.)

(d) Let $\mathbb{R}^{\mathbb{R}}$ be the set of all functions $f : \mathbb{R} \to \mathbb{R}$, and let the operations $+ : V \times V \to V$ and $\cdot : \mathbb{R} \cdot V \to V$ be defined in the usual way. I.e., if $f, g \in \mathbb{R}^{\mathbb{R}}$ and $\lambda \in \mathbb{R}$, then the functions $f + g$ and λf are defined by

$$(f + g)(x) := f(x) + g(x) \quad \text{and} \quad (\lambda f)(x) = \lambda \cdot (f(x)).$$

Then V is a real vector space. (Its zero element is the function $f(x) = 0$.)

More generally, for *any* set A, the set $\mathbb{K}^A$ of functions $A \to \mathbb{K}$ is a vector space over $\mathbb{K}$.

(e) Let $\mathrm{Pol}(\mathbb{R}) \subset \mathbb{R}^{\mathbb{R}}$ denote the set of all *polynomials* with real coefficients, e.g., functions of the form

$$x \mapsto a_d x^d + a_{d-1} x^{d-1} + \cdots + a_1 x + a_0$$

with $a_i \in \mathbb{R}$. Then $\mathrm{Pol}(\mathbb{R})$ is a (real) vector space (with the same operations as in (d)). Similarly, the set $\mathrm{Pol}_d(\mathbb{R}) \subset \mathrm{Pol}(\mathbb{R})$ of all polynomials of degree at most d is a vector space.

The set of polynomials of degree *exactly* d is *not* a vector space. (Why not?)

(f) Suppose that V and W are vector spaces over $\mathbb{K}$. Then we can turn the set of pairs $V \times W$ into a vector space by setting

$$(v_1, w_1) + (v_2, w_2) := (v_1 + v_2, w_1 + w_2) \quad \text{and} \quad \lambda(v, w) := (\lambda v, \lambda w).$$

These examples suggest the following definition:

1.1.3. Definition (Subspace)**.**
Let $(V, +, \cdot)$ be a vector space. A set $W \subset V$ is called a subspace *of V if $(W, +|_{W\times W}, \cdot|_{\mathbb{K}\times W})$ is a vector space over $\mathbb{K}$.*

Remark. Note that both V and $\{0\}$ are subspaces of V.

In order to see whether a set W is a subspace, we don't have to check all properties of a vector space, since most of them automatically follow from the corresponding properties of V.

1.1.4. Proposition (Subspace criterion)**.**
Let V be a vector space and $W \subset V$. Then W is a subspace of V if and only if

(S1) $0 \in W$,
(S2) *for all $v, w \in W$, also $v + w \in W$, and*
(S3) *for all $\lambda \in \mathbb{K}$ and $v \in W$, also $\lambda v \in W$.*

PROOF. By (S2) and (S3), the operations $+|_{W\times W}$ and $\cdot|_{\mathbb{K}\times W}$ are indeed functions $W\times W \to W$ and $\mathbb{K}\times W$. The vector space conditions (V1), (V2) and (V5) to (V8) follow immediately from the corresponding conditions for V. Since $0 \in W$, property (V3) is also clear. Finally, we know (see exercise sheet 1) that $-v = (-1)\cdot v$ for every $v \in V$, and thus W also satisfies condition (V4) by (S3). ■

REMARK. It sometimes saves time to verify (S1), (S2) and (S3) at the same time by showing the condition

(S) $\lambda v_1 + \mu v_2 \in W$.

for all $v_1, v_2 \in W$ and $\lambda, \mu \in \mathbb{K}$.

Indeed, clearly every subspace W must satisfy (S). Conversely, (S1) follows from (S) by setting $\lambda = \mu = 0$, (S2) by setting $\lambda = \mu = 1$ and (S3) by setting $\mu = 0$.

1.1.5. Examples. (a) Let $V = \mathbb{R}^3$, and $W = \{(x, y, z) \in \mathbb{R}^3 : y + z = 0\}$. Then $(0, 0, 0) \in W$. Then

$$(\lambda y_1 + \mu y_2) + (\lambda z_1 + \mu z_2) = \lambda(y_1 + z_1) + \mu(y_2 + z_2) = 0 + 0 = 0.$$

So (S) is satisfied, and thus W is a subspace of V.

(b) Again, let $V = \mathbb{R}^3$. Then $W = \{(x, y, z) \in V : x, y, z \geq 0\}$ is *not* a subspace of V: e.g., $-1 \cdot (1, 1, 1) \notin W$. So (S3) is violated. (However, (S1) and (S2) still hold.)

(c) Let $V = \mathbb{R}^{2\times 2}$ be the space of real 2-by-2 matrices, and let W consist of those matrices whose bottom left entry is 0. We claim that W is a subspace of V.

- The zero element of V is the matrix $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, which clearly belongs to W. So (S1) is satisfied.
- If $A_1 = \begin{pmatrix} x_1 & y_1 \\ 0 & z_1 \end{pmatrix}$ and $A_2 = \begin{pmatrix} x_2 & y_2 \\ 0 & z_2 \end{pmatrix}$ are arbitrary elements of W, then
$$A_1 + A_2 = \begin{pmatrix} x_1 + x_2 & y_1 + y_2 \\ 0 & z_1 + z_2 \end{pmatrix} \in W.$$
So (S2) holds.
- Similarly, if A_1 as above is an arbitrary element of W and $\lambda \in \mathbb{R}$, then
$$\lambda A_1 = \begin{pmatrix} \lambda x_1 & \lambda y_1 \\ 0 & \lambda z_1 \end{pmatrix} \in W.$$

(Alternatively, we could have verified (S) by observing that

$$\lambda A_1 + \mu A_2 = \begin{pmatrix} \lambda x_1 + \mu x_2 & \lambda y_1 + \mu y_2 \\ 0 & \lambda z_1 + \mu z_2 \end{pmatrix} \in W.)$$

(d) $\mathrm{Pol}_d(\mathbb{R})$ is a subspace of $\mathrm{Pol}(\mathbb{R})$, which is a subspace of $\mathbb{R}^{\mathbb{R}}$.

1.1.6. Definition and Lemma (Operations on Subspaces)**.**
Let V be a vector space, and let W_1, W_2 be subspaces of V. Then

$$W_1 + W_2 = \{w_1 + w_2 : w_1 \in W_1 \text{ and } w_2 \in W_2\}$$

is a subspace of V.

If $\mathcal{W}$ is a set of subspaces of V, then

$$\bigcap \mathcal{W} := \bigcap_{W \in \mathcal{W}} W = \{v \in V : v \in W \text{ for all } W \in \mathcal{W}\}$$

is a subspace of V. (In particular, the intersection of two subspaces of V is itself a subspace of V.)

If W_1 and W_2 are subspaces of V such that $W_1 \cap W_2 = \{0\}$ and $W_1 + W_2 = V$, then we say that V is the direct sum of W_1 and W_2 *and write*

$$V = W_1 \oplus W_2.$$

PROOF. We leave the proof as an exercise. □

1.1.7. Lemma (Direct Sums)**.**
Let V be a vector space and W_1, W_2 be subspaces of V. Then $V = W_1 \oplus W_2$ if and only if, for every $v \in V$, there are unique *$w_1 \in W_1$ and $w_2 \in W_2$ such that $v = w_1 + w_2$.*

PROOF. For the if direction, suppose that the given condition holds; we must show that $W_1 + W_2 = V$ and $W_1 \cap W_2 = \{0\}$. The former claim is obvious, so let $v \in W_1 \cap W_2$. Then

$$\underbrace{0}_{\in W_1} + \underbrace{0}_{\in W_2} = 0 = \underbrace{v}_{\in W_1} + \underbrace{(-v)}_{\in W_2}.$$

Thus, by assumption, $v = 0$, as required.

For the "only if " direction, suppose that $V = W_1 \oplus W_2$ and let $v \in V$. Then, by definition, $V = W_1 + W_2$, so there are $w_1 \in W_1$ and $w_2 \in W_2$ with $w_1 + w_2 = v$. To prove uniqueness, suppose that $w_1' \in W_1$ and $w_2' \in W_2$ is another pair with $v = w_1' + w_2'$. Then

$$\underbrace{w_1 - w_1'}_{\in W_1} = (v - w_2) - (v - w_2') = (v - v) + (w_2' - w_2) = \underbrace{w_2' - w_2}_{\in W_2}.$$

Thus $w_1 - w_1' \in W_1 \cap W_2 = \{0\}$, and therefore $w_1 - w_1' = 0$. In other words, $w_1 = w_1'$, as required. ■

1.1.8. Examples. (a) Let $V = \mathbb{R}^2$, $W_1 := \{(x, y) : y = -x\}$ and $W_2 := \{(x, y) : x = y\}$. Then W_1 and W_2 are subspaces of V. We claim that $W_1 + W_2 = \mathbb{R}^2$. To prove this, given $(x, y) \in \mathbb{R}^2$), we will have to find elements $(\lambda, -\lambda) \in W_1$ and $(\mu, \mu) \in W_2$ such that

$$(x, y) = (\lambda, -\lambda) + (\mu, \mu).$$

In other words, we need to solve the linear system

$$\begin{aligned} \mu + \lambda &= x \quad \text{and} \\ \mu - \lambda &= y. \end{aligned}$$

This is easily done, and we see that

$$(x,y) = \underbrace{\left(\frac{x-y}{2}, \frac{y-x}{2}\right)}_{\in W_1} + \underbrace{\left(\frac{x+y}{2}, \frac{x+y}{2}\right)}_{\in W_2}.$$

Furthermore, if $(x,y) \in W_1 \cap W_2$, then $x = y = -x$; in other words, $x = y = 0$. So

$$W_1 \cap W_2 = \{(0,0)\},$$

and it follows that $V = W_1 \oplus W_2$.

(b) Let $V = \mathbb{R}^{\mathbb{R}}$. Let W_1 consist of all *odd* functions f, i.e., $f(x) = -f(-x)$ for all $x \in \mathbb{R}$. Let W_2 consist of all *even* functions, i.e. $f(x) = f(-x)$ for all $x \in \mathbb{R}$.

We claim that $V = W_1 \oplus W_2$. To prove that $V = W_1 + W_2$, let us define

$$f_1(x) := \frac{f(x) - f(-x)}{2} \quad \text{and} \quad f_2(x) := \frac{f(x) + f(-x)}{2}.$$

Then $f_1 \in W_1$, $f_2 \in W_2$ and $f = f_1 + f_2$. Furthermore,

$$\begin{aligned} W_1 \cap W_2 &= \{f : \mathbb{R} \to \mathbb{R}, f(x) = -f(-x) = -f(x) \text{ for all } x \in \mathbb{R}\} \\ &= \{x \mapsto 0\}. \end{aligned}$$

(c) Let $V = \mathbb{R}^{2\times 2}$, let W_1 denote the set of 2×2-matrices whose bottom left entry is 0, and let W_2 be the set of symmetric 2×2-matrices. We claim that

$$W_1 \cap W_2 = \left\{ \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} : x, y \in \mathbb{R}. \right\}.$$

Indeed, clearly every matrix of this form belongs to both W_1 and W_2. On the other hand, any matrix in $W_1 \cap W_2$ has a zero in the bottom left corner, and is symmetric, so also has a zero in the top right corner.

Furthermore, we claim that $W_1 + W_2 = V$. Indeed, let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in V$ be arbitrary. Then

$$A = \underbrace{\begin{pmatrix} a & b-c \\ 0 & d \end{pmatrix}}_{\in W_1} + \underbrace{\begin{pmatrix} 0 & c \\ c & 0 \end{pmatrix}}_{\in W_2}.$$

One of the things you should have noticed by now is that, in many of the examples, the vector spaces "looked very much alike": essentially they were determined by a finite number of real (or complex) numbers (entries for vectors or matrices, coefficients for polynomials). In the next section we will see that indeed all "n-dimensional" vector spaces (i.e., those whose elements are determined by n parameters) are "the same" in some way. In order to make precise what we mean by "the same", we have to introduce the notion of an *isomorphism* between vector spaces, which is a particular kind of *linear map*.

1.1.9. Definition (Linear maps)**.**
Let V, W be vector spaces and $\varphi : V \to W$ be a function. Then φ is called linear *if*

(L1) $\varphi(v+w) = \varphi(v) + \varphi(w)$ *for all* $v, w \in V$*, and*
(L2) $\varphi(\lambda v) = \lambda\varphi(v)$ *for all* $v \in V$ *and* $\lambda \in \mathbb{K}$*.*

Such a linear map is called an isomorphism *between* V *and* W *if it is* bijective*. That is,* φ *is* surjective *(*$\varphi(V) = W$*) and* injective *(no two different elements* $v, w \in V$ *have the same image under* φ*).*

V *and* W *are* isomorphic *if there exists an isomorphism between them.*

REMARK. Again, we could replace (L1) and (L2) by the single condition

(L) $\varphi(\lambda v_1 + \mu v_2) = \lambda\varphi(v_1) + \mu\varphi(v_2)$.

1.1.10. Examples. (a) The function $\varphi : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x + 2y$ is linear, since

$$\begin{aligned}\varphi(x_1 + x_2, y_1 + y_2) &= x_1 + x_2 + 2y_1 + 2y_2 \\ &= (x_1 + 2y_1) + (x_2 + 2y_2) = \varphi(x_1, y_1) + \varphi(x_2, y2) \quad \text{and} \\ \varphi(\lambda x, \lambda y) &= \lambda x + 2\lambda y = \lambda(x + 2y) = \lambda\varphi(x, y).\end{aligned}$$

(b) The function $\varphi : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (xy, y^2)$ is *not* linear, since

$$\varphi(2 \cdot (1, 1)) = (4, 4) \neq (2, 2) = 2 \cdot \varphi(1, 1).$$

(c) Let $A \in \mathbb{R}^{m\times n}$ be an m-by-n matrix, and define $\varphi : \mathbb{R}^m \to \mathbb{R}^n$ by

$$\varphi(x_1, \ldots, x_m) = A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}.$$

(using the usual matrix multiplication). Then φ is a linear map.

(d) The vector spaces $\mathbb{R}^{d+1}$ and $\mathrm{Pol}_d(\mathbb{R})$ are isomorphic: we claim that an isomorphism is given by the map

$$\varphi : \mathbb{R}^{d+1} \to \mathrm{Pol}_d(\mathbb{R}); (a_0, \ldots, a_d) \mapsto \sum_{j=0}^{d} a_j x^j.$$

Indeed,

$$\begin{aligned}\varphi(\lambda(a_0, \ldots, a_d) + \mu(b_0, \ldots, b_d)) &= \sum_{j=0}^{d} (\lambda a_j + \mu b_j) x^j \\ &= \lambda \sum_{j=0}^{d} a_j x^j + \mu \sum_{j=0}^{d} b_j x^j \\ &= \lambda\varphi(a_0, \ldots, a_d) + \mu\varphi(b_0, \ldots, b_d),\end{aligned}$$

so φ is linear. Clearly *phi* is surjective. Furthermore, φ is injective since two polynomials are equal if and only if their coefficients are the same.

(e) The vector space $\mathrm{Pol}(\mathbb{R})$ is *not* isomorphic to $\mathbb{R}^n$ for any $n \in \mathbb{N}$. This fact seems rather obvious, and could be proved by hand. However, in the next section, we will learn some general methods which will make the proof much easier, so we shall wait until then.

1.1.11. Lemma (Direct Sum)**.**
Let V be a vector space, and W_1 and W_2 subspaces of V. Then $V = W_1 \oplus W_2$ if and only if the linear function

$$\varphi : W_1 \times W_2 \to V; (w_1, w_2) \mapsto w_1 + w_2$$

is an isomorphism.

Proof. Exercise. □

1.1.12. Proposition.
Let $\varphi : V \to W$ be a linear map. Then $\ker(\varphi) := \{z \in V : \varphi(z) = 0\}$ *is a subspace of V (the* kernel *of φ), and* $\mathrm{Im}(\varphi) := \{\varphi(z) : z \in V\}$ *is a subspace of W (the* image *of φ).*

Proof. Let $v, w \in \ker(\varphi)$ and $\lambda, \mu \in \mathbb{K}$. Then

$$\varphi(\lambda v + \mu w) = \varphi(\lambda v) + \varphi(\mu w) = \lambda\varphi(v) + \mu\varphi(w) = 0 + 0 = 0.$$

So $\lambda v + \mu w \in \ker(\varphi)$, as required.

Similarly, let $w_1 = \varphi(v_1)$ and $w_2 = \varphi(v_2)$ be arbitrary elements of $\mathrm{Im}(\varphi)$, and let $\lambda, \mu \in \mathbb{R}$. Then

$$\lambda w_1 + \mu w_2 = \lambda\varphi(v_1) + \mu\varphi(v_2) = \varphi(\lambda v_1) + \varphi(\mu v_2) = \varphi(\lambda v_1 + \mu v_2) \in \mathrm{Im}(\varphi),$$

as required. ■

1.1.13. Proposition (Injectivity Criterion)**.**
Let $\varphi : V \to W$ be a linear map. Then φ is injective if and only if $\ker \varphi = \{0\}$.

Proof. If φ is injective, then 0 is the only element of V with $\varphi(0) = 0$, so $\ker V = \{0\}$.

Now suppose that φ is not injective. Then there are $v, w \in V$ with $v \neq w$ and $\varphi(v) = \varphi(w)$. By linearity,

$$\begin{aligned}\varphi(v - w) &= \varphi(v + (-w)) = \varphi(v) + \varphi((-1) \cdot w) \\ &= \varphi(v) + (-1)\varphi(w) = \varphi(v) + (-\varphi(w)) = 0,\end{aligned}$$

so $0 \neq v - w \in \ker \varphi$. ■

1.1.14. Examples. (a) Let $\varphi : \mathbb{R}^3 \to \mathbb{R}^2; (x, y, z) \mapsto (x + y, x + y)$. This map is linear (which is easy to check). Now

$$\begin{aligned}(x, y, z) \in \ker \varphi \quad &\iff \quad \varphi(x, y, z) = (0, 0) \quad \iff \quad (x + y, x + y) = (0, 0) \\ &\iff \quad x + y = 0 \quad \iff \quad x = -y \quad \iff \quad (x, y, z) = (x, -x, z).\end{aligned}$$

So $\ker(\varphi) = \{(x, -x, z) : x, z \in \mathbb{R}\}$. In particular, φ is not injective.

(b) Let us return to the isomorphism

$$\varphi : \mathbb{R}^{d+1} \to \mathrm{Pol}_d(\mathbb{R}); (a_0, \ldots, a_d) \mapsto \sum_{j=0}^{d} a_j x^j.$$

from a previous example. Let $(a_0, \ldots, a_d) \in \ker(\varphi)$, i.e. $\sum_{j=0}^{d} a_j x^j = 0$ for all $x \in \mathbb{R}$. A polynomial of degree d with nonzero coefficients can have at most d zeros, so it follows that $a_j = 0$ for all j. In other words, $\ker(\varphi) = \{(0, \ldots, 0)\}$, and thus φ is injective (which we already knew).

1.2. Bases of a Vector Space

1.2.1. Definition.
Let V be a vector space over $\mathbb{K}$, and let $A \subset V$.

(a) *A* linear combination *of vectors in A is an element of the form*

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

where $\lambda_j \in \mathbb{K}$ and $v_j \in A$.

(b) *The set* $\mathrm{span}(A) \subset V$ *of all linear combinations of vectors in A is a subspace of V.*

(c) *We say that A is a* spanning set *of V if $V = \mathrm{span}(A)$.*

(d) *The set A is* linearly independent *if no element of A can be written as a linear combination of other elements of A, or in other words, if*

$$0 = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$$

for $v_1, \ldots, v_n \in A$ is only possible if $\lambda_j = 0$ for all j. (Otherwise, A is linearly dependent*.)*

(e) *A linearly independent spanning set is called a* basis *of V.*

Remark.

(a) $\mathrm{span}(A)$ is the smallest subspace of V containing A.

(b) It is not obvious at all whether every vector space has a basis. In fact, this turns out to be equivalent to the *Axiom of Choice*. However, for *finite-dimensional* vector spaces, which are the ones we will be focussing on, we shall see that this question is much simpler.

1.2.2. Examples. (a) Let $V = \mathbb{R}^3$, and consider the set $A := \{(0, -1, 1), (1, 0, 0)\}$. Then

$$\begin{aligned} \mathrm{span}(A) &= \{\lambda(0, -1, 1) + \mu(1, 0, 0) : \lambda, \mu \in \mathbb{R}\} \\ &= \{(\mu, -\lambda, \lambda) : \lambda, \mu \in \mathbb{R}\} \\ &= \{(x, y, z) \in \mathbb{R}^3 : y + z = 0\} \end{aligned}$$

(which is a subspace of $\mathbb{R}^3$ which we have encountered before).

Also, A is linearly independent: if $(0, 0, 0) = \lambda(0, -1, 1) + \mu(1, 0, 0)$, then clearly we must have $\lambda = \mu = 0$.

(b) Let $V = \mathrm{Pol}(\mathbb{R})$, and consider $A := \{x^2 - 1, x^3 + 2, x^3 + 2x^2\}$. Then

$$2(x^2 - 1) + (x^3 + 2) - (x^3 + 2x^2) = 0,$$

so A is linearly dependent.

(c) Let $V = \mathbb{K}^n$, and set

$$e_j := (\underbrace{0, \ldots, 0}_{j-1 \text{ times}}, 1, 0 \ldots, 0);$$

i.e., e_j is the vector whose j-th component is 1, and all other entries are 0. Then e_j is a basis of $\mathbb{K}^n$. Indeed, clearly every vector $(x_1, \ldots, x_n) \in \mathbb{K}^n$ is a linear combination of the e_j:

$$(x_1, \ldots, x_n) = \sum_{j=1}^{n} x_j e_j.$$

Also, clearly $\sum_{j=1}^{n} x_j e_j = (0, \ldots, 0)$ if and only if each x_j is zero.

(d) Define $f_j \in \mathrm{Pol}(\mathbb{R})$ by $f_j(x) = x^j$. Then f_j is a basis of $\mathrm{Pol}(\mathbb{R})$.

(e) The set $(1, 0, -1), (0, 2, 1), (1, 0, 0)$ is a basis for $\mathbb{R}^3$. Indeed, if $(x, y, z) \in \mathbb{R}^3$, and $\lambda, \mu, \nu \in \mathbb{R}$, then

$$\begin{aligned} &(x, y, z) = \lambda(1, 0, -1) + \mu(0, 2, 1) + \nu(1, 0, 0) \quad \Longleftrightarrow \\ &x = \lambda + \nu, y = 2\mu \text{ and } z = -\lambda \quad \Longleftrightarrow \\ &\lambda = -z, \mu = 2/y \text{ and } \nu = x + z. \end{aligned}$$

The claim now follows easily.

(f) Let $V = \mathbb{R}^{2\times 2}$. Then

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

is a basis of V.

1.2.3. Remark (Adding and Removing Vectors)**.**
Let V be a vector space, and let $A \subset V$. Then

(a) *For every $v \in \mathrm{span}(A)$, we have* $\mathrm{span}(A \cup \{v\}) = \mathrm{span}(A)$.

(b) *For every $v \in V \setminus A$,*

$$\begin{aligned} &A \cup \{v\} \textit{is linearly dependent} \iff \\ &v \in \mathrm{span}(A) \textit{ or } A \textit{ is linearly dependent.} \end{aligned}$$

Proof. Write $v = \mu_1 v_1 + \cdots + \mu_n v_n$ with $v_j \in A$. Then any $w \in \mathrm{span}(A \cup \{v\})$ can be written as

$$\lambda_1 w_1 + \cdots + \lambda_m w_m + \lambda v = \lambda_1 w_1 + \cdots + \lambda_m w_m + \lambda\mu_1 v_1 + \cdots + \lambda\mu_n v_n \in \mathrm{span}(A)$$

(where $w_1, \ldots, w_m \in A$ and $\lambda_1, \ldots, \lambda_m, \lambda \in \mathbb{K}$).

For the second claim, note that the "if" direction is clear by definition.So suppose that $A \cup \{v\}$ is linearly independent, so that we can write

$$0 = \lambda_1 v_1 + \cdots + \lambda_n v_n + \mu v,$$

where $v_1, \ldots v_n$ are different elements of A and at least one of $\lambda_1, \ldots, \lambda_n, \mu$ is nonzero. If $\mu = 0$, then

$$0 = \lambda_1 v_1 + \cdots + \lambda_n v_n,$$

so A is linearly dependent. Otherwise, we can write

$$v = \frac{lambda_1}{\mu} v_1 + \cdots + \frac{lambda_n}{\mu} v_n \in \operatorname{span}(A).$$

1.2.4. Definition (Finite-Dimensional Spaces)**.**
A vector space V is finite-dimensional *if it has a finite spanning set $A = \{a_1, \ldots, a_n\}$.*

1.2.5. Examples. (a) $\mathbb{R}^n$ is a finite-dimensional vector space: the standard basis is a finite basis for $\mathbb{R}^n$.
(b) $\mathrm{Pol}_d(\mathbb{R})$ is a finite-dimensional vector space.
(c) $\mathbb{R}^{n\times m}$ is a finite-dimensional vector space.
(d) $\mathrm{Pol}(\mathbb{R})$ is *not* a finite-dimensional vector space. Indeed, suppose that $A \subset \mathrm{Pol}(\mathbb{R})$ is finite and let d be the largest degree among all polynomials in A. Then $x^{d+1} \notin \operatorname{span}(A)$, so $\operatorname{span}(A) \neq \mathrm{Pol}(\mathbb{R})$.

1.2.6. Lemma (Existence of Bases)**.**
Let V be a finite-dimensional vector space. Then every finite spanning set for V contains a basis. (In particular, V has a finite basis.) Furthermore, every linearly independent subset of V can be extended to a basis.

If V has a finite basis consisting of n elements, then V is isomorphic to $\mathbb{K}^n$.

Proof. Suppose, by contradiction, that A is a finite spanning set of V which does not contain a basis. We may assume that A is chosen to contain as few elements as posible. Then by definition, $\operatorname{span}(A) = V$ and A is linearly dependent; i.e., there is $v \in A$ such that $v \in \operatorname{span}(A \setminus \{v\})$. By Remark 1.2.3, it follows that $A' := A \setminus \{v\}$ is also a spanning set of V. Since A contains no basis, A' certainly also doesn't contain a basis; however, A' has one element less than A. This contradicts the choice of A.

Let A_0 be a linearly independent subset of V and $B = \{v_1, \ldots, v_n\}$ be a finite basis of V. If $B \subset \operatorname{span}(A)$, then clearly A is a spanning set, i.e. a basis. Otherwise, there is some j such that $v_j \notin \operatorname{span}(A)$. By Remark 1.2.3, the set $A_1 := A_0 \cup \{v_j\}$ is also linearly independent. If A_1 is a spanning set, then we are done, otherwise we can add another vector of B to A_1 to obtain a linearly independent set A_2, and so on. This process will yield a basis $A_m \supset A_0$ after at most n steps.

Finally, let $B = (v_1, \ldots, v_n)$ be a basis of V. Then

$$\varphi : \mathbb{R}^n \to A; (\lambda_1, \ldots, \lambda_n) \mapsto \lambda_1 v_1 + \cdots + \lambda_n v_n$$

is an isomorphism. Indeed, linearity of this map is easy to check. Furthermore, since A is a spanning set, φ is surjective. Since A is linearly independent, we have $\ker(\varphi) = \{(0, \dots, 0)\}$; i.e., φ is injective. ■

1.2.7. Proposition.
Let V be a finite-dimensional vector space. Then every basis of V has the same number n of elements, called the dimension *of V. (We write $\dim(V) := n$. Furthermore*

(a) *a spanning set A of V is a basis if and only if it has n elements, and*
(b) *a linearly independent subset of V is a basis if and only if it has n elements.*

Remark.

(a) The dimension of V is, informally speaking, the "number of parameters" required to describe an object of V.
(b) The dimension of the trivial vector space $\mathbb{K}^0 = \{0\}$ is 0.

Proof. Let n be the smallest number of elements in a basis of V. We need to show that V does not contain a basis containing more than n elements. By the previous lemma, V is isomorphic to $\mathbb{K}^n$. However, you should already know the fact that any linearly independent subset of $\mathbb{K}^n$ consists of at most n elements. (A more familiar way of stating this might be as follows: A system of n linear equations with more than n variables cannot have a unique solution.) This proves the first claim.

Now let A be a spanning set of n elements. Then A contains a basis by the previous lemma. Since any basis must have n elements, it follows that A itself is a basis. In the same way, we see that a linearly independent set of n elements is a basis. ■

1.2.8. Examples. (a) Let $V = \mathbb{R}^3$ and

$$A := \{\underbrace{(1,5,2)}_{v_1}, \underbrace{(-3,0,1)}_{v_2}, \underbrace{(2,5,1)}_{v_3}\}.$$

We claim that A is linearly independent. Indeed,

$$\begin{aligned}
&\lambda v_1 + \mu v_2 + \nu v_3 = 0 \iff \\
&\lambda - 3\mu + 2\nu = 0 \quad \text{and} \quad 5\lambda - 5\nu = 0 \quad \text{and} \quad 2\lambda + \mu + \nu = 0 \iff \\
&\lambda = -\nu, \quad \nu = 3\mu \quad \text{and} \quad \nu = \mu \iff \\
&\lambda = \mu = \nu = 0.
\end{aligned}$$

(b) Let $V := \mathrm{Pol}_3(\mathbb{R})$ and

$$A := \{\underbrace{x^3+1}_{v_1}, \underbrace{x^3}_{v_2}, \underbrace{x^2(x-1)}_{v_3}, \underbrace{x(x-2)(x-1)}_{v_4}\}.$$

Clearly $\{v_3, v_4\}$ is linearly independent. Furthermore, any element $f \in \mathrm{span}(v_3, v_4)$ must clearly satisfy $f(1) = 0$, so $v_2 \notin \mathrm{span}(v_3, v_4)$, and thus the three vectors v_2, v_3, v_4 are linearly independent. Similarly, any element f of the span of these

three vectors will satisfy $f(0) = 0$, so v_1 does not belong to this span, so A is linearly independent. In particular, A is a basis since $\dim(V) = 4$.

1.2.9. Corollary.
Any subspace W of a finite-dimensional vector space V is finite-dimensional and $\dim(W) \leq \dim(V)$. *If* $\dim W = \dim V$, *then* $W = V$.

PROOF. Let B_1 be a basis of W. Then B_1 extends to a basis of V, which must have $\dim(V)$ elements. ■

1.2.10. Proposition.
Let V be a vector space, and let W_1 and W_2 be subspaces of V.

(a) *If A_1 is a spanning set for W_1 and A_2 is a spanning set for W_2, then $A_1 \cup A_2$ is a spanning set for $W_1 + W_2$.*
(b) *Let B_1 is a basis of W_1 and B_2 is a basis of W_2. If $W_1 \cap W_2 = \{0\}$, then $B := B_1 \cup B_2$ is a basis of $W_1 + W_2$.*
(c) *If V is finite-dimensional, then*

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

PROOF. The first claim is left to the reader.

Let us prove the second claim. We already know from the first part that B is a spanning set of $W_1 + W_2$, so we need to shwow that B is linearly independent. Suppose that there are different elements $v_1, \ldots, v_n \in B_1$, $w_1, \ldots, w_m \in B_2$ and scalars $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_m \in \mathbb{K}$ such that

$$\underbrace{\lambda_1 v_1 + \cdots + \lambda_n}_{v} + \underbrace{\mu_1 w_1 + \cdots + \mu_m w_m}_{w} = 0.$$

Then $v \in W_1$ and $v = -w \in W_2$ So $v \in W_1 \cap W_2 = \{0\}$. Thus

$$0 = v = \lambda_1 v_1 + \cdots + \lambda_n \quad \text{and}$$
$$0 = w = \mu_1 w_1 + \cdots + \mu_m w_m.$$

By linear independence of B_1 and B_2, it follows that $\lambda_1 = \cdots = \lambda_n = \mu_1 = \cdots = \mu_m = 0$, as required.

Finally, let $B_\cap$ be a basis of $W_1 \cap W_2$. Then we can extend $B_\cap$ to a basis B_1 of W_1 and to a basis B_2 of W_2. It is easy to see that

$$W_1 \cap \operatorname{span}(B_2 \setminus B_\cap) = \{0\}.$$

So $B := B_1 \cup B_2 = B_1 \cup (B_2 \setminus B_\cap)$ is a basis of V, and we have

$$\dim(W_1 + W_2) = \#B = \#B_1 + (\#B_2 - \#B_\cap) = \\ \dim(W_1) + \dim(W_2) + \dim(W_1 \cap W_2).$$

■

1.2.11. Definition and Lemma.
Nullity and Rank Let V, W be vector spaces, and let V be finite-dimensional. Then $\ker(\varphi)$ *and* $\mathrm{Im}(\varphi)$ *are finite-dimensional.* $\dim(\ker(\varphi))$ *is called the* nullity *of φ, and* $\dim(\mathrm{Im}(\varphi))$ *is called the* rank *of φ.*

PROOF. $\ker(\varphi)$ is a subspace of V, and thus finite-dimensional. If A is a finite spanning set for V, then

$$\varphi(A) = \{\varphi(v) : v \in A\}$$

is a spanning set for $\mathrm{Im}(\varphi)$. (See Question 4 on problem sheet 2.) ■

1.2.12. Proposition (Rank and Nullity Theorem)**.**
If V is a finite-dimensional vector space, W is a vector space and $\varphi : V \to W$ is a linear map, then

$$\dim(V) = \mathrm{rank}(\varphi) + \mathrm{nullity}(\varphi).$$

PROOF. Let B_1 be a basis of $\ker(\varphi)$, and extend B_1 to a basis B of V. We set $B_2 := B \setminus B_1$ and $W := \mathrm{span}(B_2)$ (so B_2 is a basis of W). Then $V = W + \ker(\varphi)$ (in fact, V is the direct sum of these subspaces.)

We claim that the restricted map $\varphi|_W : W \to \mathrm{Im}(\varphi)$ is an isomorphism. Indeed, $\ker(\varphi|_W) = W \cap \ker(\varphi) = \{0\}$, so φ is injective. Furthermore, if $v \in V$, then we can write $v = w + u$ with $w \in W$ and $u \in \ker(\varphi)$, and we have

$$\varphi(v) = \varphi(w + u) = \varphi(w) + \varphi(u) = \varphi(w).$$

It follows that

$$\dim V = \#B = \#B_1 + \#B_2 = \mathrm{nullity}(\varphi) + \dim(W) = \mathrm{nullity}(\varphi) + \mathrm{rank}(\varphi).$$

That is what we set out to prove. ■

From now on, all vector spaces considered will be finite-dimensional, unless explicitly stated otherwise.

Index