Linear Algebra (Math 110), Summer 2012

George Melvin
University of California, Berkeley
(June 17, 2012 corrected version)

Abstract

These are notes for the upper division course 'Linear Algebra' (Math 110) taught at the University of California, Berkeley, during the summer session 2012. Students are assumed to have attended a first course in linear algebra (equivalent to UCB Math 54). The aim of this course is to provide an introduction to the study of finite dimensional vector spaces over fields of characteristic zero and linear morphisms between them and to provide an abstract understanding of several key concepts previously encountered. The main topics to be covered are: basics of vector spaces and linear morphisms, the Jordan canonical form and Euclidean/Hermitian spaces.

Contents

0 Preliminaries

In this preliminary section we will introduce some of the fundamental language and notation that will be adopted in this course. It is intended to be an informal introduction to the language of sets and functions and logical quantifiers.

0.1 Basic Set Theory

For most mathematicians the notion of a *set* is fundamental and essential to their understanding of mathematics. In a sense, everything in sight is a set (even functions can be considered as sets![1]) .

A vector space is an example of a *set with structure* so we need to ensure that we know what a set is and understand how to write down and describe sets using set notation.

Definition 0.1.1 (Informal Definition)**.** A *set* S is a collection of objects (or elements). We will denote the size of a set S by $|S|$; this will either be a natural number or infinite (we do discuss questions of cardinality of sets).

For example, we can consider the following sets:

[1] A function $f : A \to B$; $x \mapsto f(x)$ is the same data as providing a subset $\Gamma_f \subset A \times B$, where $\Gamma_f = \{(x, f(x)) \mid x \in A\}$, the *graph of* f. Conversely, if $C \subset A \times B$ is a subset such that, $\forall a \in A, \exists b \in B$ such that $(a, b) \in C$, and $(a, b) = (a, b') \in C \implies b = b'$, then C is the graph of some function.

- the set P of people in Etcheverry, room 3109, at 10.10am on 6/18/2012,
- the set B of all people in the city of Berkeley at 10.10am on 6/18/2012,
- the set $\mathbb{R}$ of all real numbers,
- the set A of all real numbers that are greater than or equal to π,
- the set $M_{m\times n}(\mathbb{R})$ of all $m \times n$ matrices with real entries,
- the set $\mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^m)$ of all $\mathbb{R}$-linear morphisms with domain $\mathbb{R}^n$ and codomain $\mathbb{R}^m$,
- the set $C(0, 1)$ of all real valued continuous functions with domain $(0, 1)$.

Don't worry if some of these words are new to you, we will define them shortly.

You will observe that there are some relations between these sets: for example,

- every person that is an object in the collection P is also an object in the collection B,
- every number that is an object of A is also an object of $\mathbb{R}$.

We say in this case that P (resp. A) is a *subset* of B (resp. $\mathbb{R}$), and write

$$P \subseteq B \text{ (resp. } A \subseteq \mathbb{R}).$$

Remark. In this class we will use the notations $\subseteq$ and $\subset$ interchangeably and make no distinction between them. On the blackboard I will write $\subseteq$ as this is a notational habit of mine whereas in these notes I shall usually write $\subset$ as it is a shorter command in LaTeX(the software I use to create these notes).

We can also write the following

$$P = \{x \in B \mid x \text{ is in Etcheverry, room 3109, at 10.10am on 6/18/2012}\},$$

or in words:

P is the set of those objects x in B such that x is in Etcheverry, room 3109, at 10.10am on 6/18/2012.

Here we have used

- the logical symbol '$\in$' which is to be translated as 'is a member of' or 'is an object in the collection',
- the vertical bar '$|$' which is to be translated as 'such that' or 'subject to the condition that'.

In general, we will write (sub)sets in the following way:

$$T = \{x \in S \mid \mathcal{P}\},$$

where $\mathcal{P}$ is some property or condition. In words, the above expression is translated as

T is the set of those objects x in the set S such that x satisfies the condition/property $\mathcal{P}$.

For example, we can write

$$A = \{x \in \mathbb{R} \mid x \geq \pi\}.$$

Definition 0.1.2. We will use the following symbols (or logical quantifiers) frequently:

- $\forall$ - translated as 'for all' or 'for every', (the *universal quantifier*)
- $\exists$ - translated as 'there exists' or 'there is, (the *existential quantifier*).

For example, the statement

'for every positive real number x, there exists some real number y such that $y^2 = x$',

can be written

$$\forall x \in \mathbb{R} \text{ with } x > 0, \exists y \in \mathbb{R} \text{ such that } y^2 = x.$$

Remark. Learning mathematics is difficult and can be made considerably more difficult if the basic language is not understood. If you ever encounter any notation that you do not understand please ask a fellow student or ask me and I will make sure to clear things up. I have spent many hours of my life staring blankly at a page due to misunderstood notation so I understand your pain in trying to get to grips with new notation and reading mathematics.

Notation. In this course we will adopt the following notational conventions:

- $\varnothing$, the *empty set* (ie the empty collection, or the collection of no objects),
- $[n] = \{1, 2, 3, \ldots, n\}$,
- $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$, the set of *natural numbers*,
- $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \ldots\}$, the set of *integers*,
- $\mathbb{Z}_{\geq a} = \{x \in \mathbb{Z} \mid x \geq a\}$, and similarly $\mathbb{Z}_{>a}, \mathbb{Z}_{\leq a}, \mathbb{Z}_{<a}$,
- $\mathbb{Q} = \{\frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0\}$, the set of *rational numbers*,
- $\mathbb{R}$, the set of *real numbers*,
- $\mathbb{C}$, the set of *complex numbers*.

Remark (Complex Numbers)**.** Complex numbers are poorly taught in most places so that most students have a fear and loathing of them. However, there is no need to be afraid! It really doesn't matter whether you consider imaginary numbers to be 'real' (or to exist in our domain of knowledge in this universe), all that matters is that you know their basic properties: a complex number $z \in \mathbb{C}$ is a'number' that can be expressed in the form

$$z = a + b\Delta, \ a, b \in \mathbb{R},$$

where, for now, Δ is just some symbol.

We can add and multiply the complex numbers $z = a + b\Delta, w = c + d\Delta \in \mathbb{C}$, as follows

$$z + w = (a + c) + (b + d)\Delta, \quad z.w = (ac - bd) + (bc + ad)\Delta.$$

If $z = a + b\Delta \in \mathbb{C}$ then the complex number $\tilde{z} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}\Delta$ satisfies

$$z.\tilde{z} = \tilde{z}.z = 1,$$

so that $\tilde{z}$ is the multiplicative inverse of z and we can therefore write $1/z = z^{-1} = \tilde{z}$. Hence, if $z = a + b\Delta, w = c + d\Delta \in \mathbb{C}$, then

$$z/w = z.w^{-1} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2}\Delta.$$

Of course, the number $i = 1.\Delta$ satisfies the property that $i^2 = -1$, so that $1.\Delta$ corresponds to the imaginary number i that you learned about in high school. However, as we will be using the letter i frequently for subscripts, we shall instead just write $\sqrt{-1}$ so that we will consider complex numbers to take the form

$$z = a + b\sqrt{-1}.$$

We have the following inclusions

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C},$$

so that, in particular, *every real number is also a complex number* (if $a \in \mathbb{R}$ then we consider $a = a.1 + 0.\sqrt{-1} \in \mathbb{C}$).

Definition 0.1.3 (Operations on Sets). • Suppose that S is a set and S_1, S_2 are subsets.

- the *union of* S_1 *and* S_2 is the set

$$S_1 \cup S_2 = \{x \in S \mid x \in S_1 \text{ or } x \in S_2\}.$$

- the *intersection of* S_1 *and* S_2 is the set

$$S_1 \cap S_2 = \{x \in S \mid x \in S_1 \text{ and } x \in S_2\}.$$

More generally, if $S_i \subset S, i \in J$, is a family of subsets of S, where J is some indexing set, then we can define

$$\bigcup_{i \in J} S_i = \{s \in S \mid s \in S_k, \text{ for some } k \in J\},$$

and

$$\bigcap_{i \in J} S_i = \{s \in S \mid s \in S_k, \forall k \in J\}.$$

• Let A, B be sets.

- the *Cartesian product of* A *and* B is the set

$$A \times B = \{(a, b) \mid a \in A, b \in B\},$$

so that the elements of $A \times B$ are *ordered pairs* (a, b), with $a \in A, b \in B$. In particular, it is not true that $A \times B = B \times A$.

Moreover, if $(a, b), (a', b') \in A \times B$ and $(a, b) = (a', b')$, then we must necessarily have $a = a'$ *and* $b = b'$.

For example, consider the following subsets of $\mathbb{R}$:

$$A = \{x \in \mathbb{R} \mid 0 < x < 2\},\ B = \{x \in \mathbb{R} \mid x > 1\},\ C = \{x \in \mathbb{R} \mid x < 0\}.$$

Then,

$$A \cup B = (0, \infty),\ A \cap B = (1, 2],\ A \cap C = \varnothing,\ A \cup B \cup C = \{x \in \mathbb{R} \mid x \neq 0\}.$$

Also, we have

$$A \times C = \{(x, y) \mid 0 < x < 2,\ y < 0\}.$$

0.2 Functions

Functions allow us to talk about certain relationships that exist between sets and allow us to formulate certain operations we may wish to apply to sets. You should already know what a function is but the notation to be introduced may not have been encountered before.

Definition 0.2.1. Let A, B be sets and suppose we have a function $f : A \to B$. We will write the information of the function f as follows:

$$f : A \to B,\ x \mapsto f(x),$$

where $x \mapsto f(x)$ is to be interpreted as providing the data of the function, ie, x is the input of the function and $f(x)$ is the output of the function. Moreover,

- A is called the *domain of* f,
- B is called the *codomain of* f.

For example, if we consider the function $|.| : \mathbb{R} \to [0, \infty)$, the 'absolute value' function, then we write

$$|.| : \mathbb{R} \to [0, \infty),\ x \mapsto \begin{cases} x, & \text{if } x \geq 0, \\ -x, & \text{if } x < 0. \end{cases}$$

The codomain of $|.|$ is $[0, \infty)$ and the domain of $|.|$ is $\mathbb{R}$.

Definition 0.2.2. Let $f : A \to B$ be a function.

- we say that f is *injective* if the following condition is satisfied:

$$\forall x, y \in A, \text{ if } f(x) = f(y) \text{ then } x = y,$$

- we say that f is *surjective* if the following condition is satisfied:

$$\forall y \in B, \exists x \in A \text{ such that } f(x) = y,$$

- we say that f is *bijective* if f is both injective and surjective.

It should be noted that the injectivity of f can also be expressed as the following (logically equivalent) condition:

$$\text{if } x, y \in A,\ x \neq y, \text{ then } f(x) \neq f(y).$$

Also, the notion of bijectivity can be expressed in the following way:

$$\forall y \in B, \text{ there is a unique } x \in A \text{ such that } f(x) = y.$$

Hence, if a function is bijective then there exists an inverse function $g : B \to A$ such that

$$\forall x \in A,\ g(f(x)) = x, \text{ and } \forall y \in B,\ f(g(y)) = y.$$

The goal of the first half of this course is an attempt to try and understand the 'linear functions' whose domain and codomain are vector spaces. We will investigate if there is a way to represent the function in such a way that all desirable information we would like to know about the function is easy to obtain. In particular, we will provide (finite) criteria that allow us to determine if a function is injective/surjective (cf. Theorem **??**).

Remark. These properties of a function can be difficult to grasp at first. Students tend to find that injectivity is the hardest attribute of a function to comprehend. The next example is an attempt at providing a simple introduction to the concept of injectivity/surjectivity of functions.

Example 0.2.3. Consider the set P described above (so an object in P is a person in Etcheverry, room 3109, at 10.10am on 6/18/2012) and let $\mathcal{C}$ denote the set of all possible cookie ice cream sandwiches available at C.R.E.A.M. on Telegraph Avenue (for example, vanilla ice cream on white chocolate chip cookies). Consider the following function

$$f : P \to \mathcal{C}\ ;\ x \mapsto f(x) = x\text{'s favourite cookie ice cream sandwich.}$$

In order for f to define a function we are assuming that nobody who is an element of P is indecisive so that they have precisely one favourite cookie ice cream sandwich.[2]

So, for example,

$$f(\textit{George}) = \text{ banana walnut ice cream on chocolate chip cookies.}$$

What does it mean for f to be

- injective? Let's go back to the definition: we require that for any two people $x, y \in P$, if $f(x) = f(y)$ then $x = y$, ie, if any two people in P have the same favourite cookie ice cream sandwich then those two people must be the same person. Or, what is the same, no two people in P have the same favourite cookie ice cream sandwich.

[2] Why are we making this assumption?

- surjective? Again, let's go back to the definition: we require that, if $y \in \mathcal{C}$ then there exists some $x \in P$ such that $f(x) = y$, ie, for any possible cookie ice cream sandwich y available at C.R.E.A.M. there must exist some person $x \in P$ for which y is x's favourite cookie ice cream sandwich.

There are a couple of things to notice here:

1. in order for f to be surjective, we must necessarily have at least as many objects in P as there are objects in $\mathcal{C}$. That is
$$f \text{ surjective} \implies |P| \geq |\mathcal{C}|.$$
2. in order for f to be injective, there must necessarily be more objects in $\mathcal{C}$ as there are in P. That is
$$f \text{ injective} \implies |P| \leq |\mathcal{C}|.$$
3. if P and $\mathcal{C}$ have the same number of objects then f is injective if and only if f is surjective.

You should understand and provide a short proof as to why these properties hold true.

The fact that these properties are true is dependent on the fact that both P and $\mathcal{C}$ are *finite sets*. We will see a generalisation of these properties to finite dimensional vector spaces and linear morphisms between them: here we replace the 'size' of a vector space by its dimension (a linear algebra measure of 'size').

We will not include a basic lemma that will be useful throughout these notes. Its proof is left to the reader.

Lemma 0.2.4. *Let $f : R \to S$ and $g : S \to T$ be two functions.*

- *If f and g are both injective, then $g \circ f : R \to T$ is injective. Moreover, if $g \circ f$ is injective then f is injective.*
- *If f and g are both surjective, then $g \circ f : R \to T$ is surjective. Moreover, if $g \circ f$ is surjective then g is surjective.*
- *If f and g are bijective, then $g \circ f : R \to T$ is bijective.*

1 Vector Spaces & Linear Morphisms

This chapter is intended as a reintroduction to results that you have probably seen before in your previous linear algebra course. However, we will adopt a slightly more grown-up viewpoint and discuss some subtleties that arise when we are considering infinite dimensional vector spaces. Hopefully most of the following results are familiar to you - don't forget or disregard any previous intuition you have gained and think of the following approach as supplementing those ideas you are already familiar with.

1.1 Fields

[[1] p. 1-3]

In your previous linear algebra course (eg. Math 54) you will have mostly worked with column (or row) vectors with real or complex coefficients. However, most of linear algebra does not require that we work only with real or complex entries, only that the set of '*scalars*' we use satisfy some nice properties.

For those of you who have taken an Abstract Algebra course (eg. Math 113) you may have already been introduced to the notion of a *ring* or a *field*. What follows is a very brief introduction to (number) fields.

Definition 1.1.1 (Number Field)**.** A nonempty set $\mathbb{K}$[3] is called an *number field* if

1. $\mathbb{Z} \subset \mathbb{K}$,
2. there are well-defined notions of addition, subtraction, multiplication and division that obey all the usual laws of arithmetic.

Note that, by 1. we have that $\mathbb{K}$ contains every integer $x \in \mathbb{Z}$. Therefore, by 2., since we must be able to divide through by nonzero x, we necessarily have $\mathbb{Q} \subset \mathbb{K}$.

Example 1.1.2. 1. $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ are all examples of number fields. However, $\mathbb{Z}$ is not a number field since 2 does not have a multiplicative inverse in $\mathbb{Z}$.

2. Consider the set
$$\mathbb{Q}(\sqrt{2}) = \{a + b\Delta \mid a, b \in \mathbb{Q}\},$$
where we consider Δ as some symbol. Define an addition on $\mathbb{Q}(\sqrt{2})$ as follows: for $z = a + b\Delta, w = c + d\Delta \in \mathbb{Q}(\sqrt{2})$, define

$$z + w = (a + c) + (b + d)\Delta \in \mathbb{Q}(\sqrt{2}), \quad -z = -a + (-b)\Delta,$$

$$z.w = (ac + 2bd) + (ad + bc)\Delta, \quad z^{-1} = \frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2}\Delta.$$

Note that $a^2 - 2b^2 \neq 0$, for any $a, b \in \mathbb{Q}$ such that $(a, b) \neq (0, 0)$, since $\sqrt{2}$ is irrational.

Then, it is an exercise to check that all the usual rules of arithmetic (eg. Rules 1-9 on p.1 of [1]) hold for this addition and multiplication just defined. Moreover, it is easy to check that

$$(1.\Delta)^2 = 2,$$

so that we can justify swapping the symbol $\sqrt{2}$ for Δ.

Be careful though: we do not care about the actual value of $\sqrt{2}$ (=1.414...) as a real number, we are only interested in the algebraic properties of this number, namely that it squares to 2, and as such only consider $\sqrt{2}$ as a symbol such that $\sqrt{2}^2 = 2$ in our number field $\mathbb{Q}(\sqrt{2})$.

You should think of $\mathbb{Q}(\sqrt{2})$ in the same way as $\mathbb{C}$: for our purposes of arithmetic, we do not care whether $\sqrt{2}$ is a 'real' number or not, we only care about its basic algebraic properties as a symbol.

[3]We adopt the letter $\mathbb{K}$ for a (number) field following the Germans. In German the word for a '(number) field' is *'(zahlen) korps'*. The notation $\mathbb{Z}$ for the integers also comes from the German word *zahlen*, meaning 'number'. Most of the basics of modern day algebra was formulated and made precise by German mathematicians, the foremost of whom being C. F. Gauss, D. Hilbert, R. Dedekind, E. Noether, E. Steinitz and many, many others.

Why should we care about $\mathbb{Q}(\sqrt{2})$ when we can just think about $\mathbb{R}$? Most of modern number theory is concerned with the study of number fields and there is some sense in which the deep structure of number fields is related to seemingly unrelated areas of mathematics such as real analysis.

3. Let $p > 0$ be a nonsquare integer, so that there does not exist $x \in \mathbb{Z}$ such that $x^2 = p$. Then, we can form the number field $\mathbb{Q}(\sqrt{p})$ in a similar manner as above.

4. The p-adic numbers $\mathbb{Q}_p$: let p be a prime number. Then, you may have heard of the *p-adic numbers*: this is a number field that is obtained from $\mathbb{Q}$ in a similar way that $\mathbb{R}$ can be obtained from $\mathbb{Q}$ (via Cauchy sequences; this is Math 104 material). Essentially, a p-adic number $a \in \mathbb{Q}_p$ can be considered as a formal power series

$$a = \sum_{i \geq m} a_i p^i = a_m p^m + a_{m+1} p^{m+1} + \cdots, \quad a_i \in \{0, 1, \dots, p-1\}, m \in \mathbb{Z},$$

where we do not care about the fact that this 'sum' does not converge and add and multiply as you would as if you were in high school (remembering to reduce coefficients modulo p). We will not talk about this number field again and if you are interested in learning more simply Google 'p-adic numbers' and there will be plenty information available online.

5. The field of rational polynomials $\mathbb{Q}(t)$: here we have

$$\mathbb{Q}(t) = \{p/q \mid p, q \in \mathbb{Q}[t],\ q \neq 0\},$$

where $\mathbb{Q}[t]$ is the set of polynomials with rational coefficients. For example,

$$\frac{3 - \frac{5}{3}t^8}{t^2 + \frac{2}{7}t^{167}} \in \mathbb{Q}(t).$$

Again, $\mathbb{Q}(t)$ is a number field and arises in algebraic geometry, that area of mathematics concerned with solving systems of polynomials equations (it's very hard!).

Remark. 1. The definition of 'number field' given above is less general than you might have seen: in general, a *field* $\mathbb{K}$ is a nonempty set for which there are well-defined notions of addition, subtraction, multiplication and division (and obeying all the usual laws of arithmetic) without the extra requirement that $\mathbb{Z} \subset \mathbb{K}$; this is the definition given in [1]. The definition we have given in Definition 1.1.1 defines a *field of characteristic zero*.

2. In this course we will only be concerned with 'linear algebra over number fields', meaning the scalars we consider will have to take values in a number field $\mathbb{K}$ as defined in Definition 1.1.1. Moreover, most of the time we will take $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ and when we come to discuss Euclidean (resp. Hermitian) spaces we must necessarily have $\mathbb{K} = \mathbb{R}$ (resp. $\mathbb{K} = \mathbb{C}$).

In grown-up mathematical language we will be studying '**vector spaces over characteristic zero fields**' (or, in an even more grown-up language, $\mathbb{K}$-modules, where $\mathbb{K}$ is a characteristic zero field).

From now on, $\mathbb{K}$ will always denote a number field.

1.2 Vector Spaces

1.2.1 Basic Definitions

[p.31-33, [1]]

Definition 1.2.1 (Vector Space)**.** A $\mathbb{K}$-*vector space* (or *vector space over* $\mathbb{K}$) is a triple (V, α, σ), where V is a nonempty set and

$$\alpha : V \times V \to V\ ;\ (u, v) \mapsto \alpha(u, v), \quad \sigma : \mathbb{K} \times V \to V\ ;\ (\lambda, v) \mapsto \sigma(\lambda, v),$$

are two functions called *addition* and *scalar multiplication*, and such that the following axioms are imposed:

(VS1) α is associative: for every $u, v, w \in V$ we have

$$\alpha(u, \alpha(v, w)) = \alpha(\alpha(u, v), w);$$

(VS2) α is commutative: for every $u, v \in V$ we have

$$\alpha(u, v) = \alpha(v, u);$$

(VS3) there exists an element $0_V \in V$ such that, for every $v \in V$, we have

$$\alpha(0_V, v) = \alpha(v, 0_V) = v.$$

We call 0_V a (in fact, *the*[4]) *zero element* or *zero vector* of V;

(VS4) for every $v \in V$ there exists an element $\hat{v} \in V$ such that

$$\alpha(v, \hat{v}) = \alpha(\hat{v}, v) = 0_V.$$

We call $\hat{v}$ the[5] *negative of* v and denote it $-v$;

(VS5) for every $\lambda, \mu \in \mathbb{K}$, $v \in V$, we have

$$\sigma(\lambda + \mu, v) = \alpha(\sigma(\lambda, v), \sigma(\mu, v));$$

(VS6) for every $\lambda, \mu \in \mathbb{K}$, $v \in V$, we have

$$\sigma(\lambda\mu, v) = \sigma(\lambda, \sigma(\mu, v));$$

(VS7) for every $\lambda \in \mathbb{K}, u, v \in V$, we have

$$\sigma(\lambda, \alpha(u, v)) = \alpha(\sigma(\lambda, u), \sigma(\lambda, v));$$

(VS8) for every $v \in V$ we have

$$\sigma(1, v) = v.$$

In case α and σ satisfy the above axioms so that (V, α, σ) is a vector space (over $\mathbb{K}$) we will usually write

$$\boxed{\begin{array}{c} \alpha(u, v) = u + v, \quad u, v, \in V, \\ \sigma(\lambda, v) = \lambda \cdot v, \ \text{ or simply } \ \sigma(\lambda, v) = \lambda v, \quad \lambda \in \mathbb{K}, v \in V. \end{array}}$$

If (V, α, σ) is a vector space over $\mathbb{K}$ then we will call an element $x \in V$ a *vector* and an element $\lambda \in \mathbb{K}$ a *scalar*.

If $v \in V$ is such that

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

for some vectors $v_1, ..., v_n \in V$ and scalars $\lambda_1, ..., \lambda_n \in \mathbb{K}$, then we say that v is a *linear combination* of the vectors $v_1, ..., v_n$.

[4] See Proposition 1.2.3.

[5] We shall see (Proposition 1.2.4) that the negative of v is unique, so that this notation is well-defined.

Remark 1.2.2. The given definition of a vector space might look cumbersome given the introduction of the functions α and σ. However, it is important to realise that these defined notions of addition and scalar multiplication tell us how we are to 'add' vectors and how we are to 'scalar multiply' vectors by scalars in $\mathbb{K}$; in particular, a nonempty set V may have many different ways that we can define a vector space structure on it, ie, it may be possible that we can obtain two distinct $\mathbb{K}$-vector spaces (V, α, σ) and (V, α', σ') which have the same underlying set but different notions of addition and scalar multiplication. In this case, it is important to know which 'addition' (ie, which function α or α') we are discussing, or which 'scalar multiplication' (ie, which function σ or σ') we are discussing.

In short, the definition of a vector space is the data of providing a nonempty set together with the rules we are using for 'addition' and 'scalar multiplication'.

Notation. • Given a $\mathbb{K}$-vector space (V, α, σ) we will usually know which notions of addition and scalar multiplication we will be discussing so we will often just write V instead of the triple (V, α, σ), the functions α, σ being understood *a priori*.

• We will also frequently denote an arbitrary vector space (V, α, σ) by V, even when we don't know what α, σ are explicitly. Again, we are assuming that we have been given 'addition' and 'scalar multiplication' functions *a priori*.

Proposition 1.2.3. *Let (V, α, σ) be a $\mathbb{K}$-vector space. Then, a zero vector $0_V \in V$ is unique.*

Proof: Suppose there is some element $z \in V$ such that z satisfies the same properties as 0_V, so that, for every $v \in V$, we have $z + v = v + z = v$. Then, in particular, we have

$$0_V = z + 0_V = z,$$

where the first equality is due to the characterising properties of z as a zero vector (Axiom VS3), and the second equality is due to the characterising property of 0_V as a zero vector (Axiom VS3). Hence, $z = 0_V$ so that a zero vector is unique. $\square$

Proposition 1.2.4 (Uniqueness of negatives)**.** *Let (V, α, σ) be a $\mathbb{K}$-vector space. Then, for every $v \in V$, the element $\hat{v}$ that exists by Axiom VS4 is unique.*

Proof: Let $v \in V$ and suppose that $w \in V$ is such that $w + v = v + w = 0_V$. Then,

$$\begin{aligned} w = w + 0_V = w + (v + \hat{v}), &\quad \text{by defining property of } \hat{v}, \\ = (w + v) + \hat{v}, &\quad \text{by Axiom VS1}, \\ = 0_V + \hat{v}, &\quad \text{by assumed property of } w, \\ = \hat{v}, &\quad \text{by Axiom VS3}. \end{aligned}$$

Hence, $w = \hat{v}$ and the negative of v is unique. $\square$

Proposition 1.2.5 (Characterising the zero vector)**.** *Let (V, α, σ) be a $\mathbb{K}$-vector space. Then, for every $v \in V$ we have $0 \cdot v = 0_V$. Moreover, $\lambda \cdot 0_V = 0_V$, for every $\lambda \in \mathbb{K}$. Conversely, if $\lambda \cdot v = 0_V$ with $v \neq 0_V$, then $\lambda = 0 \in \mathbb{K}$.*

Proof: Let $v \in V$. Then, noting the trivial fact that $0 = 0 + 0 \in \mathbb{K}$, we have

$$\begin{aligned} 0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v, &\quad \text{by Axiom VS5}, \\ \implies 0_V = 0 \cdot v + (-0 \cdot v) = (0 \cdot v + 0 \cdot v) + (-0 \cdot v) = 0 \cdot v + (0 \cdot v + (-0 \cdot v)), &\quad \text{using Axiom VS1}, \\ \implies 0_V = 0 \cdot v + 0_V = 0 \cdot v, &\quad \text{using Axioms VS3 and VS4}. \end{aligned}$$

Furthermore, in a similar way (using Axioms VS3, VS4 and VS7) we can show that $\lambda \cdot 0_V = 0_V$, for every $\lambda \in \mathbb{K}$.[6]

[6] Do this as an exercise!

Conversely, suppose that $v \neq 0_V$ is a vector in V and $\lambda \in \mathbb{K}$ is such that $\lambda v = 0_V$. Assume that $\lambda \neq 0$; we aim to provide a contradiction. Then, λ^{-1} exists and we have

$$\begin{aligned} v = 1 \cdot v = (\lambda^{-1}\lambda) \cdot v, &\quad \text{using Axiom VS8,} \\ = \lambda^{-1} \cdot (\lambda \cdot v), &\quad \text{by Axiom VS6,} \\ = \lambda^{-1} \cdot 0_V, &\quad \text{using our assumption,} \\ = 0_V, &\quad \text{by the result just proved.} \end{aligned}$$

But this contradicts our assumption that v is nonzero. Hence, our initial assumption that $\lambda \neq 0$ cannot hold so that $\lambda = 0 \in \mathbb{K}$.

□

The following examples will be fundamental for the rest of the course so make sure that you acquaint yourself with them as they will be used frequently throughout class and on homework/exams. As such, if you are having trouble understanding them then please ask a fellow student for help or feel free to send me an email and I will help out as best I can.

I have only provided the triple (V, α, σ) in each example, you should *define the zero vector in each example and the negative of an arbitrary given vector $v \in V$*. Also, you should check that the Axioms VS1-VS8 hold true.

Example 1.2.6. 1. For $n \in \mathbb{N}$, consider the $\mathbb{K}$-vector space $(\mathbb{K}^n, \alpha, \sigma)$, where

$$\mathbb{K}^n = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mid x_1, \dots, x_n \in \mathbb{K} \right\},$$

$$\alpha\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right) = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}, \quad \text{and} \quad \sigma\left(\lambda, \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right) = \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{bmatrix}.$$

We will usually denote this vector space simply $\mathbb{K}^n$. It is of fundamental importance in all that follows.

We will denote by $e_i \in \mathbb{K}^n$, the column vector that has a 1 in the i^{th} entry and 0s elsewhere, and call it the i^{th} *standard basis vector*.

2. Let S be an arbitrary nonempty set. Then, define the $\mathbb{K}$-vector space $(\mathbb{K}^S, \alpha, \sigma)$, where

$$\mathbb{K}^S = \{\text{functions } f : S \to \mathbb{K}\},$$

$$\alpha(f, g) : S \to \mathbb{K}\ ;\ s \mapsto f(s) + g(s) \in \mathbb{K}, \quad \text{and} \quad \sigma(\lambda, f) : S \to \mathbb{K}\ ;\ s \mapsto \lambda f(s).$$

That is, we have defined the sum of two functions $f, g \in \mathbb{K}^S$ to be the new function $\alpha(f, g) : S \to \mathbb{K}$ such that $\alpha(f, g)(s) = f(s) + g(s)$ (here the addition is taking place inside the number field $\mathbb{K}$).

Since this example can be confusing at first sight, I will give you the zero vector $0_{\mathbb{K}^S}$ and the negative of a vector $f \in \mathbb{K}^S$: since the elements in $\mathbb{K}^S$ are functions we need to ensure that we define a function

$$0_{\mathbb{K}^S} : S \to \mathbb{K},$$

satisfying the properties required of Axiom VS3. Consider the function

$$0_{\mathbb{K}^S} : S \to \mathbb{K}\ ;\ s \mapsto 0,$$

that is, $0_{\mathbb{K}^S}(s) = 0 \in \mathbb{K}$, for every $s \in S$. Let's show that this function just defined satisfies the properties required of a zero vector in $(\mathbb{K}^S, \alpha, \sigma)$. So, let $f \in \mathbb{K}^S$, ie,

$$f : S \to \mathbb{K}\ ;\ s \mapsto f(s).$$

Then, we have, for every $s \in S$,

$$\alpha(f, 0_{\mathbb{K}^S})(s) = f(s) + 0_{\mathbb{K}^S}(s) = f(s) + 0 = f(s),$$

so that $\alpha(f, 0_{\mathbb{K}^S}) = f$. Similarly, we have $\alpha(0_{\mathbb{K}^S}, f) = f$. Hence, $0_{\mathbb{K}^S}$ satisfies the properties required of the zero vector in $\mathbb{K}^S$ (Axiom VS3).

Now, let $f \in \mathbb{K}^S$. We define a function $-f \in \mathbb{K}^S$ as follows:

$$-f : S \to \mathbb{K} \; ; \; s \mapsto -f(s) \in \mathbb{K}.$$

Then, $-f$ satsfies the properties required of the negative of f (Axiom VS4).

For every $s \in S$, we define the *characteristic function of s, $e_s \in \mathbb{K}^S$*, where

$$e_s(t) = \begin{cases} 0, & \text{if } t \neq s, \\ 1, & \text{if } t = s. \end{cases}$$

For example, if $S = \{1, 2, 3, 4\}$ then

$$\mathbb{K}^S = \{\text{functions } f : \{1, 2, 3, 4\} \to \mathbb{K}\}.$$

What is a function $f : \{1, 2, 3, 4\} \to \mathbb{K}$? To each $i \in \{1, 2, 3, 4\}$ we associate a scalar $f(i) \in \mathbb{K}$, which we can also denote $f_i \overset{def}{=} f(i)$. This choice of notation should lead you to think there is some kind of similarity between $\mathbb{K}^4$ and $\mathbb{K}^{\{1,2,3,4\}}$; indeed, these two vector spaces are *isomorphic*, which means they are essentially the same (in the world of linear algebra).

For example, we have the characteristic function $e_2 \in \mathbb{K}^{\{1,2,3,4\}}$, where

$$e_2(1) = e_2(3) = e_2(4) = 0, \quad e_2(2) = 1.$$

3. Let $m, n \in \mathbb{N}$ and consider the sets $[m] = \{1, \dots, m\}, [n] = \{1, \dots, n\}$. Then, we have the set

$$[m] \times [n] = \{(x, y) \mid x \in [m], y \in [n]\}.$$

Then, we define the set of *$m \times n$ matrices with entries in $\mathbb{K}$* to be

$$Mat_{m,n}(\mathbb{K}) \overset{def}{=} \mathbb{K}^{[m]\times[n]}.$$

Hence, an *$m \times n$ matrix A* is a function $A : [m] \times [n] \to \mathbb{K}$, ie, a matrix is completely determined by the values $A(i, j) \in \mathbb{K}$, for $(i, j) \in [m] \times [n]$. We will denote such an $m \times n$ matrix in the usual way:

$$A \equiv \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix},$$

where we have $A_{ij} \overset{def}{=} A(i, j)$. Thus, we add and scalar multiply $m \times n$ matrices 'entry-wise'.

We will denote the zero vector $0_{Mat_{m,n}(\mathbb{K})} \in Mat_{m,n}(\mathbb{K})$ by $0_{m,n}$.

4. This is a generalisation of the previous examples. Let (V, β, τ) be a vector space and S any nonempty set. Define the $\mathbb{K}$-vector space (V^S, α, σ) where

$$V^S \overset{def}{=} \{\text{functions } f : S \to V\},$$

$$\alpha(f, g) : S \to V \; ; \; s \mapsto \beta(f(s), g(s)), \quad \text{and} \quad \sigma(\lambda, f) : S \to V \; ; \; s \mapsto \tau(\lambda, f(s)).$$

This might look a bit confusing to you so let's try and make things a bit clearer: denote the addition afforded by β as $\hat{+}$, so that if $u, v \in V$ then the addition defined in V is denoted $u \hat{+} v$. If we denote the addition we have defined by α in V^S as $\tilde{+}$, then the previous definition states that

$$\forall f, g \in V^S, \text{ define } f \tilde{+} g (= \alpha(f, g)) \in V^S \text{ by } (f \tilde{+} g)(s) = f(s) \hat{+} g(s) (= \beta(f(s), g(s))) \in V.$$

5. Consider the set

$$\mathbb{K}[t] \stackrel{def}{=} \{a_0 + a_1 t + ... + a_m t^m \mid a_m \in \mathbb{K}, m \in \mathbb{Z}_{\geq 0}\},$$

of polynomials with coefficients in $\mathbb{K}$. Define the $\mathbb{K}$-vector space $(\mathbb{K}[t], \alpha, \sigma)$ where, for

$$f = a_0 + a_1 t + ... + a_m t^m, \; g = b_0 + b_1 t + ... + b_n t^n \in \mathbb{K}[t],$$

with $m \leq n$, say, we have

$$\alpha(f, g) = (a_0 + b_0) + (a_1 + b_1)t + ... + (a_m + b_m)t^m + b_{m+1}t^{m+1} + ... + b_n t^n.$$

Also, we have, for $\lambda \in \mathbb{K}$,

$$\sigma(\lambda, f) = (\lambda a_0) + (\lambda a_1)t + ... + (\lambda a_m)t^m.$$

That is, for any two polynomials $f, g \in \mathbb{K}[t]$ we simply add and scalar multiply them 'coefficient-wise'.

For $n \in \mathbb{Z}_{\geq 0}$ we define the vector spaces $(\mathbb{K}_n[t], \alpha_n, \sigma_n)$ where,

$$\mathbb{K}_n[t] = \{f = a_0 + a_1 t + ... + a_m t^m \in \mathbb{K}[t] \mid m \leq n\},$$

$$\alpha_n(f, g) = \alpha(f, g), \quad \text{and} \quad \sigma_n(\lambda, f) = \sigma(\lambda, f).$$

Here $\mathbb{K}_n[t]$ is the set of all polynomials with coefficients in $\mathbb{K}$ of degree at most n; the fact that we have defined the addition and scalar multiplication in $\mathbb{K}_n[t]$ via the addition and scalar multiplication of $\mathbb{K}[t]$ is encoded in the notion of a *vector subspace*.

$(*)$ There is an important (and subtle) point to make here: a vector $f \in \mathbb{K}[t]$ (or $\mathbb{K}_n[t]$) is just a <u>formal expression</u>

$$f = a_0 + a_1 t + ... + a_n t^n, \quad a_0, ..., a_n \in \mathbb{K}.$$

This means that we are only considering a polynomial as a formal symbol that we add and scalar multiply according to the rules we have given above. In particular, **two polynomials** $f, g \in \mathbb{K}[t]$ **are equal, so that** $f = g \in \mathbb{K}[t]$**, if and only if they are equal coefficient-wise**.

This might either seem obvious or bizarre to you. I have included this comment as most students are used to seeing polynomials considered as *functions*

$$f : \mathbb{K} \to \mathbb{K} \; ; \; t \mapsto f(t) = \sum_{i=0}^{n} a_i t^i,$$

and I am saying that we do not care about this interpretation of a polynomial, we are only concerned with its formal (linear) algebraic properties. This is why I will write 'f' instead of '$f(t)$' for a polynomial. You might wonder why we even bother writing a polynomial as

$$f = a_0 + a_1 t + ... + a_n t^n,$$

when we don't care about the powers of t that are appearing, we could just write

$$f = (a_0, ..., a_n),$$

and 'represent' the polynomial f by this row vector (or, even a column vector). Well, the $\mathbb{K}$-vector space of polynomials has a further property, it is an example of a $\mathbb{K}$-algebra: for those of you who will take Math 113, this means $\mathbb{K}[t]$ is not only a $\mathbb{K}$-vector space, it is also has the structure of an algebraic object called a *ring*. This extra structure arises from the fact that we can multiply polynomials together (in the usual way) and in this context we do care about the powers of t that appear.

6. All the previous examples of $\mathbb{K}$-vector spaces have underlying sets that are *infinite*. What happens if we have a $\mathbb{K}$-vector space (V, α, σ) whose underlying set V is *finite*?

First, we give an example of a $\mathbb{K}$-vector space containing one element: we define the[7] *trivial $\mathbb{K}$-vector space* to be $(\underline{Z}, \alpha, \sigma)$ where $\underline{Z} = \{0_{\underline{Z}}\}$ and

$$\alpha : \underline{Z} \times \underline{Z} \to \underline{Z} \ ; \ (0_{\underline{Z}}, 0_{\underline{Z}}) \mapsto 0_{\underline{Z}}, \quad \text{and} \quad \sigma : \mathbb{K} \times \underline{Z} \to \underline{Z} \ ; \ (\lambda, 0_{\underline{Z}}) \mapsto 0_{\underline{Z}}.$$

This defines a structure of a $\mathbb{K}$-vector space on $\underline{Z}$.[8]

Now, recall that we are assuming that $\mathbb{K}$ is a number field so that $\mathbb{Z} \subset \mathbb{K}$ and $\mathbb{K}$ must therefore be infinite. Also, let's write (as we will do for the rest of these notes) λv instead of $\sigma(\lambda, v)$, for $v \in V, \lambda \in \mathbb{K}$.

Since V defines a vector space then we must have the zero element $0_V \in V$ (Axiom VS3). Suppose that there exists a nonzero vector $w \in V$ (ie $w \neq 0_V$); we aim to provide a contradiction, thereby showing that V *must contain exactly one element*. Then, since $\lambda w \in V$, for every $\lambda \in \mathbb{K}$, and $\mathbb{K}$ is infinite we must necessarily have distinct scalars $\mu_1, \mu_2 \in \mathbb{K}$ such that $\mu_1 w = \mu_2 w$ (else, all the λw's are distinct, for all possible $\lambda \in \mathbb{K}$. Since there are an infinite number of these we can't possibly have V finite). Furthermore, we assume that both μ_1 and μ_2 are nonzero scalars[9]. Hence, we have

$$w = 1w = (\mu_1^{-1}\mu_1)w = \mu_1^{-1}(\mu_1 w) = \mu_1^{-1}(\mu_2 w) = (\mu_1^{-1}\mu_2)w,$$

where we have used Axiom VS6 for the third and fifth equalities and our assumption for the fourth equality.

Hence, adding $-(\mu_1^{-1}\mu_2)w$ to both sides of this equation gives

$$w + (-(\mu_1^{-1}\mu_2)w) = 0_V \implies (1 - \mu_1^{-1}\mu_2)w = 0_V, \text{ by Axiom VS5.}$$

Therefore, by Proposition 1.2.5, we must have

$$1 - \mu_1^{-1}\mu_2 = 0 \in \mathbb{K} \implies \mu_1 = \mu_2,$$

contradicting the fact that μ_1 and μ_2 are distinct. Hence, our intial assumption - that there exists a nonzero vector $w \in V$ - cannot hold true, so that $V = \{0_V\}$.

Any $\mathbb{K}$-vector space (V, α, σ) for which V is a finite set must be a trivial $\mathbb{K}$-vector space.

7. The set of complex numbers $\mathbb{C}$ is a $\mathbb{R}$-vector space with the usual addition and scalar multiplication (scalar multiply $z \in \mathbb{C}$ by $x \in \mathbb{R}$ as $xz \in \mathbb{C}$). Moreover, both $\mathbb{R}$ and $\mathbb{C}$ are $\mathbb{Q}$-vector spaces with the usual addition and scalar multiplication.

However, $\mathbb{R}$ is not a $\mathbb{Q}(\sqrt{-1})$-vector space, where $\mathbb{Q}(\sqrt{-1}) = \{a + b\sqrt{-1} \mid a, b \in \mathbb{Q}\}$ with the arithmetic laws defined in a similar way as we have defined for $\mathbb{C}$.

Moreover, $\mathbb{Q}$ is not a $\mathbb{R}$-vector space nor a $\mathbb{C}$-vector space; $\mathbb{R}$ is not a $\mathbb{C}$-vector space.

Remark 1.2.7. From now on we will no longer denote a particular $\mathbb{K}$-vector space that appears in Example 1.2.6 as a triple (V, α, σ) but only denote its underlying set V, the operations of addition and scalar multiplication begin understood to be those appearing above. We will also write $u + v$ (resp. λv) instead of $\alpha(u, v)$ (resp. $\sigma(\lambda, v)$) for these examples.

[7] Technically, we should say 'a' instead of 'the' as any one-point set defines a $\mathbb{K}$-vector space and all of them are equally 'trivial'. However, we can show that all of these trivial $\mathbb{K}$-vector spaces are *isomorphic* (a notion to be defined in the next section) so that, for the purposes of linear algebra, they are all (essentially) the same. For our purposes we will not need to care about this (extremely subtle) distinction.

[8] Exercise: check this. Note that this is the *only* possible $\mathbb{K}$-vector space structure we can put on $\underline{Z}$. Moreover, if you think about it, you will see that *any* set with one element defines a $\mathbb{K}$-vector space.

[9] Why? Try and prove this as an exercise.

1.2.2 Subspaces

[p.42, [1]]

Definition 1.2.8 (Subspace)**.** Let (V, α, σ) be a $\mathbb{K}$-vector space and $U \subset V$ a nonempty subset. Then, we say that U is a *vector subspace of V* if the following properties hold:

(SUB1) $0_V \in U$,

(SUB2) for every $u, v \in U$, we have $\alpha(u, v) \in U$, (*closed under addition*)

(SUB3) for every $\lambda \in \mathbb{K}, u \in U$, we have $\sigma(\lambda, u) \in U$. (*closed under scalar multiplication*)

In fact, we can subsume these three properties into the following single property

(SUB) for every $u, v \in U, , \mu, \lambda \in \mathbb{K}$, we have $\alpha(\sigma(\mu, u), \sigma(\lambda, v)) \in U$ (ie, $\mu u + \lambda v \in U$).

In this case, U can be considered as a $\mathbb{K}$-vector space in its own right: we have a triple $(U, \alpha_{|U}, \sigma_{|U})$ where $\alpha_{|U}$ (resp. $\sigma_{|U}$) denote the functions α (resp. σ) *restricted to U*.[10] Notice that we need to ensure that U is closed under addition (and scalar multiplication) in order that the functions $\alpha_{|U}$ and $\sigma_{|U}$ are well-defined.

Example 1.2.9. Recall the examples from Example 1.2.6 and our conventions adopted thereafter (Remark 1.2.7).

0. There are always two obvious subspaces of a $\mathbb{K}$-vector space V: namely, V is a subspace of itself, and the subset $\{0_V\} \subset V$ is a subspace of V called the *zero subspace*. We call these subspaces the *trivial subspaces of V*. All other subspaces are called *nontrivial*.

1. Consider the $\mathbb{Q}$-vector space $\mathbb{Q}^3$ and the subset

$$U = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{Q}^3 \mid x_1 + x_2 - 2x_3 = 0 \right\}.$$

Then, U is a $\mathbb{Q}$-vector space.

How can we confirm this? We need to show that U satisfies the Axiom SUB from Definition 1.2.8. So, let $u, v \in U$ and $\mu, \lambda \in \mathbb{Q}$. Thus,

$$u = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \text{with} \quad x_1 + x_2 - 2x_3 = 0, \quad \text{and}$$

$$v = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}, \quad \text{with} \quad y_1 + y_2 - 2y_3 = 0.$$

Then,

$$\mu u + \lambda v = \begin{bmatrix} \mu x_1 \\ \mu x_2 \\ \mu x_3 \end{bmatrix} + \begin{bmatrix} \lambda y_1 \\ \lambda y_2 \\ \lambda y_3 \end{bmatrix} = \begin{bmatrix} \mu x_1 + \lambda y_1 \\ \mu x_1 + \lambda y_1 \\ \mu x_3 + \lambda y_3 \end{bmatrix},$$

and to show that $\mu u + \lambda v = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \in U$ we must show that $z_1 + z_2 - 2z_3 = 0$: indeed, we have

$$\begin{aligned} z_1 + z_2 - 2z_3 &= (\mu x_1 + \lambda y_1) + (\mu x_2 + \lambda y_2) - 2(\mu x_3 + \lambda y_3), \\ &= \mu(x_1 + x_2 - 2x_3) + \lambda(y_1 + y_2 - 2y_3), \\ &= 0 + 0 = 0. \end{aligned}$$

[10] Here is an important but subtle distinction: the functions α and $\alpha_{|U}$ are not the <u>same</u> functions. Recall that when we define a function we must also specify its domain and codomain. The functions α and $\alpha_{|U}$ are defined as

$$\alpha : V \times V \to V \ ; \ (u, v) \mapsto \alpha(u, v), \quad \alpha_{|U} : U \times U \to U \ ; \ (u', v') \mapsto \alpha(u', v'),$$

So, technically, even though α and $\alpha_{|U}$ are defined by the same 'rule', they have different (co)domains so should be considered as different functions. The same reasoning holds for σ and $\sigma_{|U}$ (how are these functions defined?)

Hence, U is a vector subspace of $\mathbb{Q}^3$.

2. Consider the subset

$$U = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mid x_1 = 1 \right\} \subset \mathbb{C}^2,$$

of the $\mathbb{C}$-vector space $\mathbb{C}^2$. Then, U is not a vector subspace of $\mathbb{C}^2$.

How can we show that a given subset E of a $\mathbb{K}$-vector space V is not a vector subspace? We must show that E does not satisfy *all* of the Axioms SUB1-3 from Definition 1.2.8, so we need to show that *at least one* of these axioms fails to hold. If you are given some subset (eg. $U \subset \mathbb{C}^2$ above) and want to determine that it is not a subspace, the first thing to check is whether the zero vector is an element of this subset: for us, this means checking to see if $0_{\mathbb{C}^2} \in U$. This is easy to check: we have

$$0_{\mathbb{C}^2} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

and since the first entry of $0_{\mathbb{C}^2} \neq 1$ it is not an element in the set U. Hence, Axiom SUB1 does not hold for the subset $U \subset \mathbb{C}^2$ so that U is not a subspace of $\mathbb{C}^2$.

However, it is possible for a subset $E \subset V$ of a vector space to contain the zero vector and still not be a subspace.[11]

3. This is an example that requires some basic Calculus.

Consider the $\mathbb{R}$-vector space $\mathbb{R}^{(0,1)}$ consisting of all $\mathbb{R}$-valued functions

$$f : (0,1) \to \mathbb{R},$$

and the subset

$$C_{\mathbb{R}}(0,1) = \{f \in \mathbb{R}^{(0,1)} \mid f \text{ is continuous}\}.$$

Then, it is a fact proved in Math 1A that $C_{\mathbb{R}}(0,1)$ is a vector subspace of $\mathbb{R}^{(0,1)}$: namely, the (constant) zero function is continuous, the sum of two continuous functions is again a continuous function and a scalar multiple of a continuous function is a continuous function.

4. Consider the $\mathbb{Q}$-vector space $Mat_3(\mathbb{Q})$ of 3×3 matrices with $\mathbb{Q}$-entries. Then, for a 3×3 matrix A define the *trace of A* to be

$$\mathrm{tr}(A) = \mathrm{tr}\left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \right) = a_{11} + a_{22} + a_{33} \in \mathbb{Q}.$$

Denote

$$sl_3(\mathbb{Q}) = \{A \in Mat_3(\mathbb{Q}) \mid \mathrm{tr}(A) = 0\},$$

the set of 3×3 matrices with trace zero. Then, $sl_3(\mathbb{Q})$ is a subspace of $Mat_3(\mathbb{Q})$. Let's check the Axioms SUB1-3 from Definition 1.2.8 (or, equivalently, you can just check Axiom SUB):

SUB1: recall that the zero vector in $Mat_3(\mathbb{Q})$ is just the zero matrix

$$0_{Mat_3(\mathbb{Q})} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus, it is trivial to see that this matrix has trace zero so that $0_{Mat_3(\mathbb{Q})} \in sl_3(\mathbb{Q})$.

[11] For example, consider the subset $\mathbb{Q} \subset \mathbb{R}$ of the $\mathbb{R}$-vector space $\mathbb{R}$.

SUB2: let $A, B \in sl_3(\mathbb{Q})$ be two matrices with trace zero, so that

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad \text{with} \quad a_{11} + a_{22} + a_{33} = 0, \quad \text{and}$$

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}, \quad \text{with} \quad b_{11} + b_{22} + b_{33} = 0.$$

Then,

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} \\ a_{31} + b_{31} & a_{32} + b_{32} & a_{33} + b_{33} \end{bmatrix},$$

and

$$(a_{11} + b_{11}) + (a_{22} + b_{22}) + (a_{33} + b_{33}) = (a_{11} + a_{22} + a_{33}) + (b_{11} + b_{22} + b_{33}) = 0 + 0 = 0,$$

so that $A + B \in sl_3(\mathbb{Q})$.

SUB3: let $A, \in sl_3(\mathbb{Q}), \lambda \in \mathbb{Q}$. Then,

$$\lambda A = \begin{bmatrix} \lambda a_{11} & \lambda a_{12} & \lambda a_{13} \\ \lambda a_{21} & \lambda a_{22} & \lambda a_{23} \\ \lambda a_{31} & \lambda a_{32} & \lambda a_{33} \end{bmatrix},$$

and

$$\lambda a_{11} + \lambda a_{22} + \lambda a_{33} = \lambda(a_{11} + a_{22} + a_{33}) = \lambda.0 = 0.$$

Hence, $sl_3(\mathbb{Q})$ is a subspace of the $\mathbb{Q}$-vector space $Mat_3(\mathbb{Q})$.

5. Consider the subset $\mathrm{GL}_3(\mathbb{Q}) \subset Mat_3(\mathbb{Q})$, where

$$\mathrm{GL}_3(\mathbb{Q}) = \{A \in Mat_3(\mathbb{Q}) \mid \det(A) \neq 0\}.$$

Here, $\det(A)$ denotes the determinant of A that you should have already seen in Math 54 (or an equivalent introductory linear algebra course). Then, $\mathrm{GL}_3(\mathbb{Q})$ is not a vector subspace.

Again, we need to show that at least one of Axioms SUB1-3 does not hold: we will show that Axiom SUB2 does not hold. Consider the 3×3 identity matrix $I_3 \in Mat_3(\mathbb{Q})$. Then, $I_3 \in \mathrm{GL}_3(\mathbb{Q})$ and $-I_3 \in \mathrm{GL}_3(\mathbb{Q})$. However,

$$\det(I_3 + (-I_3)) = \det(0_{Mat_3(\mathbb{Q})}) = 0,$$

so that $\mathrm{GL}_3(\mathbb{Q})$ is not closed under addition, therefore is not a subspace of $Mat_3(\mathbb{Q})$.

Note that we could have also shown that $0_{Mat_3(\mathbb{Q})} \notin \mathrm{GL}_3(\mathbb{Q})$.[12]

6. This example generalises Examples 1 and 4 above.[13] Consider the subset

$$U = \left\{ \underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{K}^n \mid A\underline{x} = \underline{0} \in \mathbb{K}^m \right\},$$

where A is an $m \times n$ matrix with entries in $\mathbb{K}$. Then, U is a vector subspace of $\mathbb{K}^n$.

In the next section we will see that we can interpret U as the *kernel* of a linear transformation

$$T_A : \mathbb{K}^n \to \mathbb{K}^m \; ; \; \underline{x} \mapsto A\underline{x},$$

[12] Here, the symbol $\notin$ should be translated as 'is not a member of' or 'is not an element of'.

[13] Once we have (re)considered linear transformations in the next section you should explain why we are generalising those particular examples.

defined by the matrix A. As such, we will leave the proof that U is a subspace until then.

Note here an important point: if $\underline{b} \neq \underline{0} \in \mathbb{K}^m$ then the subset

$$W = \left\{ \underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{K}^n \mid A\underline{x} = \underline{b} \in \mathbb{K}^m \right\} \subset \mathbb{K}^n,$$

is *not* a vector subspace of $\mathbb{K}^n$. For example, $\underline{0}_{\mathbb{K}^n} \notin W$. This is the generalisation of Example 2 above.

So, a slogan might be something like

$$\boxed{\text{'}\textbf{subspaces are kernels of linear maps}\text{'.}}$$

In fact, this statement is more than just a slogan[14]:

Theorem. *Let V be a $\mathbb{K}$-vector space and $U \subset V$ a subspace. Then, there exists a $\mathbb{K}$-vector space W and a $\mathbb{K}$-linear morphism $\pi : V \to W$ such that $U = \ker \pi$.*

We will now provide some constructions that allow us to form new subspaces from old subspaces.

Definition 1.2.10 (Operations on Subspaces)**.** Let V be a $\mathbb{K}$-vector space, $U, W \subset V$ two subspaces of V. Then,

- the *sum of U and W* is the subspace[15]

$$U + W = \{u + w \in V \mid u \in U, w \in W\},$$

- the *intersection of U and W* is the subspace[16]

$$U \cap W = \{v \in V \mid v \in U \text{ and } v \in W\},$$

Moreover, these notions can be extended to arbitrary families of subspace $(U_j)_{j \in J}$, with each $U_j \subset V$ a subspace of V.

We say that the sum of U and W is a *direct sum*, if $U \cap W = \{0_V\}$ is the zero subspace of V. In this case we write

$$U \bigoplus W, \quad \text{instead of} \quad U + W. \quad \text{(cf. p.45, [1])}$$

Proposition 1.2.11. *Let V be a $\mathbb{K}$-vector space and $U, W \subset V$ vector subspaces. Then, $V = U \oplus W$ if and only if, for every $v \in V$ there exists unique $u \in U$ and unique $w \in W$ such that $v = u + w$.*

Proof: ($\Rightarrow$) Suppose that $V = U \oplus W$. By definition, this means that $V = U + W$ and $U \cap W = \{0_V\}$. Hence, for every $v \in V$ we have $u \in U, w \in W$ such that $v = u + w$ (dy the definition of the sum $U + W$). We still need to show that this expression is unique: if there are $u' \in U, w' \in W$ such that

$$u' + w' = v = u + w,$$

then we have

$$u' - u = w - w',$$

and the LHS of this equation is a vector in U (since it's a subspace) and the RHS is a vector in W (since it's a subspace). Hence, if we denote this vector y (so $y = u' - u = w - w'$) then we have $y \in U$ and $y \in W$ so that $y \in U \cap W$, by definition. Therefore, as $U \cap W = \{0_V\}$, we have $y = 0_V$ so that

$$u' - u = 0_V, \text{ and } w - w' = 0_V,$$

giving $u = u'$ and $w = w'$ and the uniqueness is verified.

[14]The proof of the Theorem requires the notion of a *quotient space*.
[15]You will prove this for homework.
[16]You will prove this for homework.

($\Leftarrow$) Conversely, suppose that every vector $v \in V$ can be expressed uniquely as $v = u + w$ for $u \in U$ and $w \in W$. Then, the existence of this expression for each $v \in V$ is simply the statement that $V = U + W$. Moreover, let $x \in U \cap W$, so that $x \in U$ and $x \in W$. Thus, there are $u \in U$ and $w \in W$ (namely, $u = x$ and $w = x$) such that

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

and since U and W are subspaces (so that $0_V \in U, W$) we find, by the uniqueness of an expression for $x \in U \cap W \subset V$, that $u = 0_V = w$. Hence, $x = 0_V$ and $U \cap W = \{0_V\}$. $\square$

References

[1] Shilov, Georgi E., *Linear Algebra*, Dover Publications 1977.