1.4 Linear Morphisms, Part I

We have given an introduction to vector spaces and we have introduced the fundamental ideas of linear (in)dependence and spans. In this section we will consider the relationships that can exist between distinct vector spaces and which respect the 'linear algebraic' structure of vector spaces: this is thel notion of a *linear morphism* between vector spaces.

Definition 1.4.1. Let V and W be $\mathbb{K}$-vector spaces.

• A function
$$f : V \to W \; ; \; v \mapsto f(v),$$
is called a $\mathbb{K}$-*linear morphism between* V *and* W if the following properties hold:

(LIN1) for every $u, v \in V$, we have
$$f(u+v) = f(u) + f(v);$$
where the '$+$' on the LHS of this equation is addition in V and the '$+$' on the RHS of this equation is addition in W,

(LIN2) for every $u \in V$, $\lambda \in \mathbb{K}$, we have
$$f(\lambda v) = \lambda f(v);$$
where the scalar multiplication on the LHS of this equation is occuring in V and on the RHS of this equation it is occuring in W.

In fact, we can subsume both of these properties into

(LIN) for every $u, v \in V$, $\lambda \in \mathbb{K}$, we have
$$f(u + \lambda v) = f(u) + \lambda f(v),$$
where the scalar multiplication on the LHS of this equation is occuring in V and on the RHS of this equation it is occuring in W.

• For given $\mathbb{K}$-vector spaces V and W we denote the set of all $\mathbb{K}$-linear morphisms by
$$\mathrm{Hom}_{\mathbb{K}}(V, W) = \{f : V \to W \mid f \text{ linear}\}.$$

• The set of all $\mathbb{K}$-linear morphisms from a $\mathbb{K}$-vector space V to itself is denoted
$$\mathrm{End}_{\mathbb{K}}(V) \stackrel{def}{=} \mathrm{Hom}_{\mathbb{K}}(V, V).$$

A vector $f \in \mathrm{End}_{\mathbb{K}}(V)$ is called an *endomorphism* of V. For every $\mathbb{K}$-vector space V there exists the *identity morphism of* V, denoted $\mathrm{id}_V \in \mathrm{End}_{\mathbb{K}}(V)$. See the upcoming examples (Example 1.4.8,).

• We will use the adjectives 'injective', 'surjective' and 'bijective' to describe linear morphisms that satisfy the corresponding conditions.

• A bijective linear morphism will be called an *isomorphism*.

The set of all bijective $\mathbb{K}$-linear morphisms from a $\mathbb{K}$-vector space V to itself is denoted
$$\mathrm{GL}_{\mathbb{K}}(V) = \{f \in \mathrm{End}_{\mathbb{K}}(V) \mid f \text{ is bijective}\}.$$

We will see that, in the world of linear algebra, $\mathbb{K}$-vector spaces that are isomorphic have the same linear algebraic properties (and, therefore, can be regarded as 'the same').

Notation. You may have seen the phrases '**linear map**', '**linear transformation**' or '**linear function**': these all mean the same thing, namely, a function satisfying (LIN) above. We are using the word 'morphism' to emphasise the fact that a linear morphism is a function that 'changes' one vector space to another. This is also the fancy grown-up word that certain mathematicians use (myself included) in daily parlance.

Remark 1.4.2. We will see in a later section (Theorem **??**) that, for $f \in \mathrm{End}_{\mathbb{K}}(V)$, with V a *finite dimensional* $\mathbb{K}$-vector space

$$\text{`}f \text{ injective'} \implies \text{`}f \text{ surjective'} \implies \text{'}f \text{ bijective'} \implies \text{`}f \text{ injective'},$$

so that all of these notions are equivalent for finite-dimensional $\mathbb{K}$-vector spaces.

Lemma 1.4.3. *Let $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ be a $\mathbb{K}$-linear morphism between the $\mathbb{K}$-vector spaces V and W. Then, $f(0_V) = 0_W$.*

Proof: We have

$$f(0_V) = f(0_V + 0_V) = f(0_V) + f(0_V), \quad \text{by LIN1},$$

and subtracting $f(0_V)$ from both sides of this equation we obtain

$$0_W = f(0_V).$$

□

Definition 1.4.4. Let V, W be $\mathbb{K}$-vector spaces and $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$. Then,

- the *kernel of f* is the subset

$$\ker f = \{v \in V \mid f(v) = 0_W\} \subset V,$$

- the *image of f* is the subset

$$\mathrm{im} f = \{w \in W \mid w = f(v), \text{ for some } v \in V\} \subset W.$$

Proposition 1.4.5. *Let $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$, for $\mathbb{K}$-vector spaces V, W. Then,*

- $\ker f$ *is a subspace of V,*
- $\mathrm{im} f$ *is a subspace of W.*

Proof: Left to the reader. □

Definition 1.4.6. Let V, W be $\mathbb{K}$-vector spaces. Then, we will define the structure of a $\mathbb{K}$-vector space on the set $\mathrm{Hom}_{\mathbb{K}}(V, W)$: define the $\mathbb{K}$-vector space $(\mathrm{Hom}_{\mathbb{K}}(V, W), \alpha, \sigma)$ where

$$\alpha : \mathrm{Hom}_{\mathbb{K}}(V, W) \times \mathrm{Hom}_{\mathbb{K}}(V, W) \to \mathrm{Hom}_{\mathbb{K}}(V, W) \ ; \ (f, g) \mapsto (\alpha(f, g) : V \to W \ ; \ v \mapsto f(v) + g(v)),$$

$$\sigma : \mathbb{K} \times \mathrm{Hom}_{\mathbb{K}}(V, W) \to \mathrm{Hom}_{\mathbb{K}}(V, W) \ ; \ (\lambda, f) \mapsto (\sigma(\lambda, f) : V \to W \ ; \ v \mapsto \lambda f(v)).$$

As usual we will write

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

Whenever we discuss $\mathrm{Hom}_{\mathbb{K}}(V, W)$ as a $\mathbb{K}$-vector space, it will be this $\mathbb{K}$-vector space structure that we mean.

There are a couple of things that need to be checked to ensure that the above defintion of $\mathbb{K}$-vector space on $\mathrm{Hom}_{\mathbb{K}}(V, W)$ makes sense:

1. You need to check that the new functions $f + g$ and λf that we have defined, for $f, g \in \mathrm{Hom}_{\mathbb{K}}(V, W), \lambda \in \mathbb{K}$, are actually elements in $\mathrm{Hom}_{\mathbb{K}}(V, W)$, that is, that they are $\mathbb{K}$-linear morphisms.
2. The zero vector $0_{\mathrm{Hom}_{\mathbb{K}}(V,W)} \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ is the $\mathbb{K}$-linear morphism

$$0_{\mathrm{Hom}_{\mathbb{K}}(V,W)} : V \to W \ ; \ v \mapsto 0_W.$$

3. Given $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ we define the negative of f to be the $\mathbb{K}$-linear morphism

$$-f : V \to W \; ; \; v \mapsto -f(v),$$

where $-f(v)$ is the negative (in W) of the vector $f(v)$, for each $v \in V$.

Remark 1.4.7. 1. The fact that $\mathrm{Hom}_{\mathbb{K}}(V, W)$ has the structure of a $\mathbb{K}$-vector space will be important when we come to consider the Jordan canonical form. In that case, we will be considering the $\mathbb{K}$-vector space $\mathrm{End}_{\mathbb{K}}(V)$ and using some of its basic linear algebraic structure to deduce important properties of $\mathbb{K}$-linear morphisms $f : V \to V$.

2. We can consider $\mathrm{Hom}_{\mathbb{K}}(V, W) \subset W^V$ as a subset of the $\mathbb{K}$-vector space of (arbitrary) functions (recall Example 1.2.6)

$$W^V = \{f : V \to W\}.$$

In fact, $\mathrm{Hom}_{\mathbb{K}}(V, W) \subset W^V$ is a vector subspace.

However, the condition of $\mathbb{K}$-linearity that we have imposed on the functions is very strong and there are far 'fewer' $\mathbb{K}$-linear functions than there are arbitrary functions. For example, we will see in a proceeding section that $\mathrm{Hom}_{\mathbb{K}}(V, W)$ is finite-dimensional, whereas W^V is infinite-dimensional (assuming $W \neq Z$, the trivial vector space introduced in Example 1.2.6, 6[21]).

3. It is **not** true that $\mathrm{GL}_{\mathbb{K}}(V)$ is a vector subspace of $\mathrm{End}_{\mathbb{K}}(V)$, for any $\mathbb{K}$-vector space V that is not the trivial $\mathbb{K}$-vector space $\underline{Z}$ with one element (cf. Example 1.2.6). For example, the zero vector $0_{\mathrm{End}_{\mathbb{K}}(V)} \notin \mathrm{GL}_{\mathbb{K}}(V)$ since $0_{\mathrm{End}_{\mathbb{K}}(V)} : V \to V$ is not an injective function: if $v \in V$ is nonzero in V then

$$0_{\mathrm{End}_{\mathbb{K}}(V)}(v) = 0_V = 0_{\mathrm{End}_{\mathbb{K}}(V)}(0_V),$$

where we have used Lemma 1.4.3 for the RHS equality.

We will now give some basic examples of $\mathbb{K}$-linear morphisms. Most of these should be familiar from your first linear algebra class and, as such, you should feel pretty at ease with showing that the given functions are linear.

Example 1.4.8. 1. Consider the function

$$f : \mathbb{Q}^4 \to \mathbb{Q}^2 \; ; \; \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \mapsto \begin{bmatrix} x_1 - x_4 \\ x_3 + \frac{2}{7}x_1 \end{bmatrix}.$$

Then, f is $\mathbb{Q}$-linear.

2. The function

$$f : \mathbb{R}^2 \mapsto \mathbb{R}^3 \; ; \; \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mapsto \begin{bmatrix} x_1^3 + 2x_2^2 \\ -x_1 + \sqrt{2}x_2 \\ x_1 \end{bmatrix},$$

is not $\mathbb{R}$-linear. For example, if it were, then we must have (recall the definition of e_i from Example 1 1.2.6)

$$f(-e_2) = \begin{bmatrix} 2 \\ -\sqrt{2} \\ 0 \end{bmatrix},$$

whereas

$$-f(e_2) = -\begin{bmatrix} 2 \\ \sqrt{2} \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ -\sqrt{2} \\ 0 \end{bmatrix}.$$

Hence,

$$f(-e_2) \neq -f(e_2),$$

[21] What happens when $W = Z$?

so that Axiom LIN2 does not hold.

The problem we have here is the appearance of 'nonlinear' terms x_2^2 etc. In general, we must only have single powers of x_i appearing as in Example 1.

3. In general, a function

$$f : \mathbb{K}^n \to \mathbb{K}^m \; ; \; \underline{x} \mapsto f(\underline{x}),$$

is a $\mathbb{K}$-linear morphism if and only there exists an $m \times n$ matrix $A \in Mat_{m\times n}(\mathbb{K})$ with entries in $\mathbb{K}$ such that

$$f(\underline{x}) = A\underline{x}, \quad \text{for every } \underline{x} \in \mathbb{K}^n.$$

You should have already seen this result from your first linear algebra class.

Conversely, given $A \in Mat_{m,n}(\mathbb{K})$ we define the $\mathbb{K}$-linear morphism

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

This notation will reappear through these notes.

4. Let V be a $\mathbb{K}$-vector space. Then, the *identity morphism of V* is the $\mathbb{K}$-linear morphism

$$\mathrm{id}_V : V \to V \; ; \; v \mapsto v.$$

It is easy to see that $\mathrm{id}_V \in GL_{\mathbb{K}}(V)$, ie, that id_V is an isomorphism.

5. Let V be a $\mathbb{K}$-vector space and $U \subset V$ be a vector subspace. Then, there is a $\mathbb{K}$-linear morphism

$$i_U : U \to V \; ; \; u \mapsto u,$$

called the *inclusion morphism of U*. It is trivial to verify that this is $\mathbb{K}$-linear. Moreover, i_U is an *injective* morphism, for any subspace $U \subset V$.

6. Let V be a $\mathbb{K}$-vector space and suppose that there are subspaces $U, W \subset V$ such that $V = U \oplus W$. Then, define the *projection morphisms onto U and W* as follows:

$$p_U : V \to U \; ; \; v = u + w \mapsto u,$$

$$p_W : V \to W \; ; \; v = u + w \mapsto w.$$

These morphisms are *surjective*.

Note that these functions are well-defined because $V = U \oplus W$ and so every $v \in V$ can be uniquely written as $v = u + w$ (by Proposition 1.2.11). Therefore, we need not worry about whether p_U, p_W are functions.[22]

7. The following are examples from calculus: consider the $\mathbb{R}$-vector space $C_{\mathbb{R}}[0,1]$ of continuous functions $f : [0,1] \to \mathbb{R}$. Then, the function

$$\int_0^1 : C_{\mathbb{R}}[0,1] \to \mathbb{R} \; ; \; f \mapsto \int_0^1 f(x)dx,$$

is $\mathbb{R}$-linear. This should be well-known to all.

If we denote by $C^1(0,1) \subset C_{\mathbb{R}}(0,1)$ the set of all continuous functions $f : (0,1) \to \mathbb{R}$ that are differentiable, then we have an $\mathbb{R}$-linear map[23]

$$\frac{d}{dx} : C^1(0,1) \to C_{\mathbb{R}}(0,1) \; ; \; f \mapsto \frac{df}{dx},$$

[22] If we did not have the uniqueness property, and only knew that $V = U = W$, then it could be possible that $v = u + w = u' + w$ with $u \neq u' \in U$. Then, $p_U(v)$ could equal either u or u', so that p_U can't be a function (recall that a function $f : S \to W$ must assign a unique value $f(s)$, to every $s \in S$).

[23] It is not necessarily true that a function that can be differentiated once can be differentiated *twice*. It is actually surprisingly hard to find such an example but if you take Math 104 you should see the following example of such a function

$$f : \mathbb{R} \to \mathbb{R} \; ; \; x \mapsto \begin{cases} x^2 \sin(x^{-1}), & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

which is just the '*derivative with respect to x*' morphism. It is $\mathbb{R}$-linear.

8. This example exhibits a subtlety that we shall come back to in later sections: recall the set of natural numbers $\mathbb{N}$. Define a function

$$T : \mathbb{K}^{\mathbb{N}} \to \mathbb{K}^{\mathbb{N}} \ ; \ (i \mapsto f(i)) \mapsto \left(i \mapsto \begin{cases} 0, & \text{if } i = 1, \\ f(i-1), & \text{if } i \neq 1. \end{cases} \right).$$

That is, if we represent a function $(f : \mathbb{N} \to \mathbb{K} \ ; \ i \mapsto f(i)) \in \mathbb{K}^{\mathbb{N}}$ by an infinite sequence

$$(f_i) = (f_1, f_2, f_3, ...),$$

where $f_i \stackrel{def}{=} f(i)$, then

$$T((f_i)) = (0, f_1, f_2, f_3, ...).$$

So, T is the '*shift to the right by one place*' function defined on infinite sequences of numbers in $\mathbb{K}$.

Then, it is relatively straightforward to see that T is $\mathbb{K}$-linear and is injective. However, T is **not surjective**: thus, we have an example of an injective linear endomorphism of a $\mathbb{K}$-vector space that is not surjective. As we will see in an upcoming section, this is impossible if $\mathbb{K}^{\mathbb{N}}$ were finite-dimensional (cf. Theorem **??**). Hence, this implies that $\mathbb{K}^{\mathbb{N}}$ is an infinite dimensional $\mathbb{K}$-vector space.

We now recall an important result that allows us to characterise when $\mathbb{K}$-linear morphisms are injective. In practice, whenever you want to show that a morphism is injective you should use the following

Lemma 1.4.9 (Characterising injective linear morphisms)**.** *Let V, W be $\mathbb{K}$-vector spaces, $f : V \to W$ a $\mathbb{K}$-linear morphism. Then, f is injective if and only if* $\ker f = \{0_V\}$.

Proof: ($\Rightarrow$) Suppose that f is injective. Let $v \in \ker f$; we want to show that $v = 0_V$. Now, since $v \in \ker f$, then $f(v) = 0_W$, by the definition of $\ker f$. Furthermore, by Lemma 1.4.3, we know that $f(0_V) = 0_W$. Hence, as f is injective then

$$f(v) = f(0_V) \implies v = 0_V,$$

so that $\ker f = \{0_V\}$.

($\Leftarrow$) Conversely, suppose that $\ker f = \{0_V\}$. We must show that f is injective: therefore, we need to show that, whenever $f(v) = f(w)$, for some $v, w \in V$, then we necessarily have $v = w$. So suppose that there are $v, w \in V$ with $f(v) = f(w)$. Then

$$f(v) = f(w) \implies f(v) - f(w) = 0_W \implies f(v - w) = 0_W, \text{ since } f \text{ linear},$$

so that $v - w \in \ker f = \{0_V\}$. Hence, $v = w$. Therefore, f must be an injective function. □

Remark 1.4.10. In this section, we have given a (re)introduction to linear morphisms (or linear maps, transformations, whatever) and stated some basic properties and examples. However, in practice it is usually pretty difficult to prove certain things about linear morphisms (for example, injectivity, surjectivity etc.) in a direct manner.

In order to make questions easier to understand and solve we will most often *represent* a linear morphism using a *matrix representation*. This will be done in the proceeding sections. However, it should be noted that this approach to attacking problems only works for finite-dimensional vector spaces and the morphisms between them (infinite matrices are difficult to manipulate!).

We finish this section with some important facts that we will use throughout the remainder of these notes.

Theorem 1.4.11 (Invariance of Domain)**.** *Suppose that there exists an isomorphism*

$$f : \mathbb{K}^n \to \mathbb{K}^m.$$

Then, $n = m$.

Proof: This is an exercise in row-reduction and one which you should already be familiar with.

Recall that for any linear morphism $f : \mathbb{K}^n \to \mathbb{K}^m$, there is a matrix A_f called the *standard matrix associated to* f such that

$$\text{for every } \underline{x} \in \mathbb{K}^n, \ f(\underline{x}) = A_f \underline{x}.$$

A_f is defined to be the $m \times n$ matrix whose i^{th} column is the column vector $f(e_i)$, where e_i is the i^{th} standard basis vector of $\mathbb{K}^n$ (Example 1.2.6).
Then, it will be an exercise to show the following:

- f is injective if and only if A_f has a pivot in every column, and
- f is surjective if and only if A_f has a pivot in every row.

Therefore, since we are assuming that f is an isomorphism it must, by definition, be a bijective morphism. Hence, it is both injective and surjective. By the preceding comments we must therefore have a pivot in every column and every row. The only way that this can happen is if $n = m$. □

We will see later, after the introduction of bases for vector spaces, that the converse if also true: namely, **if $n = m$ then $\mathbb{K}^n$ and $\mathbb{K}^m$ are isomorphic.**

Proposition 1.4.12. *Let V, W be $\mathbb{K}$-vector spaces, $E \subset V$ a subset of V. Let $f : V \to W$ be an isomorphism from V to W and denote $f(E) = \{f(e) \mid e \in E\}$, the* image set *of E.*[24] *Then,*

- *E is linearly independent if and only if $f(E)$ is linearly independent.*
- *E spans V if and only if $f(E)$ spans W.*

Proof: Left to the reader. □

1.5 Bases, Dimension

In this section we will introduce the notion a *basis* of a $\mathbb{K}$-vector space. We will provide several equivalent approaches to the definition of a basis and see that the size of a basis is an invariant[25] of a $\mathbb{K}$-vector space which we will call its *dimension*. You should have already seen the words *basis* and *dimension* in your previous linear algebra course so do not abandon what you already know! We are just simply going to provide some interesting(?) ways we can think about a basis; in particular, these new formulations will allow us to extend our results to infinite dimensional vector spaces.

First, we must introduce a (somewhat annoying) idea to keep us on the straight and narrow when we are considering bases, that of an *ordered set*.

Definition 1.5.1 (Ordered Set)**.** An *ordered set* is a nonempty set S for which we have provided a 'predetermined ordering' on S.

Remark 1.5.2. 1. This definition might seem slightly confusing (and absurd); indeed, it is both of these things as I have not rigorously defined what a 'predetermined ordering' is. Please don't dwell too much on this as we will only concern ourselves with orderings of finite sets (for which it is easy to provide an ordering) or the standard ordering of $\mathbb{N}$. An ordered set is literally just a nonempty set S whose elements have been (strictly) ordered in some way.

For example, suppose that $S = [3] = \{1, 2, 3\}$. We usually think of S has having its natural ordering $(1, 2, 3)$. However, when we consider this ordering we are actually considering the ordered set $(1, 2, 3)$ and not the set S... Confused? I thought so. We could also give the objects in S the ordering $(2, 1, 3)$ and when we do this we have a defined a *different* ordered set to $(1, 2, 3)$.

[24] This is not necessarily the same as the *image of f*, $\mathrm{im} f$, introduced before.

[25] In mathematics, when we talk of an *invariant* we usually mean an atrribute or property of an object that remains unchanged whenever that object is transformed to another via an isomorphism (in an appropriate sense). For example, you may have heard of the *genus* of a (closed) geometric surface: this is an invariant of a surface that counts the number of 'holes' that exist within a (closed) surface. Perhaps you have heard or read the phrase that a mathematician thinks a coffee mug and a donut are indistuingishable. This is because we can continuously deform a donut into a coffee mug, and vice versa. This continuous deformation can be regarded as an 'isomorphism' in the world of (closed) geometric surfaces.

If you are still confused, do not worry. Here is another example: consider the set

$$S = \{\text{Evans Hall, Doe Library, Etcheverry Hall}\}.$$

Now, there is no predetermined way that we can order this set: I might choose the ordering

$$(\text{Evans Hall, Etcheverry Hall, Doe Library}),$$

whereas you might think it better to choose an ordering

$$(\text{Doe Library, Etcheverry Hall, Evans Hall}).$$

Of course, neither of these choices of orderings is 'right' and we are both entitled to our different choices. However, these ordered sets **are different**.

The reason we require this silly idea is when we come to consider coordinates (with respect to a given ordered basis). Then, it will be extremely important that we declare an ordering of a basis and that we are consistent with this choice.

2. Other examples of ordered sets include $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ and $\mathbb{R}$ with their usual orderings. We can also order $\mathbb{C}$ in an ordering called a *lexicographic ordering*: here we say that $z = a_1 + b_1\sqrt{-1} < w = a_2 + b_2\sqrt{-1}$ if and only if either, $a_1 < a_2$, or, $a_1 = a_2$ *and* $b_1 < b_2$. Think of this as being similar to the way that words are ordered in the dictionary, except now we consider only 'words' consisting of two 'letters', each of which is a real number.

3. What about some really bizarre set that might be infinite; for example, $\mathbb{R}^{\mathbb{R}}$, the set of all functions $\mathbb{R} \to \mathbb{R}$. How can we order this set? In short, I have no idea! However, there are some very deep results from mathematical logic that say that, if we assume a certain axiom of mathematics (the so-called Axiom of Choice), then every set can be ordered in some manner. In fact, it has been shown that the Axiom of Choice of logically equivalent to this ordering property of sets! If you want to learn more then you should consult Wikipedia and take Math 125A in the Fall Semester.[26]

Therefore, no matter how weird or massively infinite a set is, if you are assuming the Axiom of Choice (which we are) then you can put an ordering on that set, even though you will (a priori) have no idea what that ordering is! All that matters is that such an ordering **exists**.

Definition 1.5.3 (Basis; Ordered Basis)**.** Let V be a $\mathbb{K}$-vector space. A nonempty subset $\mathcal{B} \subset V$ is called a $(\mathbb{K})$-*basis of* V if

- $\mathcal{B}$ is linearly independent (over $\mathbb{K}$), and
- if $\mathcal{B} \subset \mathcal{B}'$ and $\mathcal{B}'$ is linearly independent (over $\mathbb{K}$), then $\mathcal{B}' = \mathcal{B}$.

In this case, we say that $\mathcal{B}$ is *maximal linearly independent*.

An *ordered* $(\mathbb{K})$-*basis of* V is a $(\mathbb{K})$-basis of V that is an ordered set.

Remark 1.5.4. 1. You may have seen a basis of a $\mathbb{K}$-vector space V defined as a subset $\mathcal{B} \subset V$ such that $\mathcal{B}$ is linearly independent (over $\mathbb{K}$) and such that $\operatorname{span}_{\mathbb{K}} \mathcal{B} = V$. The definition given above is equivalent to this and it has been used as the definition of a basis to encapsulate the intuition behind a basis: namely, if $\mathbb{K} = \mathbb{R}$, we can think of a basis of an $\mathbb{R}$-vector space as a choice of 'independent directions' that allows us to consider well-defined coordinates. This idea of 'independent directions' is embodied in the fact that a basis must be a linearly independent set; and the assumption of maximal linear independence is what allows us to obtain well-defined coordinates.

However, just to keep our minds at ease our next result will show the equivalence between Definition 1.5.3 and the definition you have probably seen before.

[26] I have to admit that I do not know any mathematical logic but have come across these ideas during my own excursions in mathematics. There are lots of many interesting results that can be obtained if one assumes the Axiom of Choice: one is called the Banach-Tarski Paradox; another, which is directly related to our studies, is the existence of a basis for *any* $\mathbb{K}$-vector space. In fact, the Axiom of Choice is logically equivalent to the existence of a basis for any $\mathbb{K}$-vector space.

2. We will also see in the homework that we can consider a basis to be a *minimal spanning set* (in an appropriate sense to be defined later); this is recorded in Proposition 1.5.9.

3. It is important to remember that a basis is a subset of V and **not** a subspace of V.

4. We will usually not call a basis of a $\mathbb{K}$-vector space a '$\mathbb{K}$-basis', it being implicitly assumed that we are considering only $\mathbb{K}$-bases when we are talking about $\mathbb{K}$-vector spaces. As such, we will only use the terminology 'basis' from now on.

Proposition 1.5.5. *Let V be a $\mathbb{K}$-vector space and $\mathcal{B} \subset V$ a basis of V. Then, $\operatorname{span}_{\mathbb{K}} \mathcal{B} = V$. Conversely, if $\mathcal{B} \subset V$ is a linearly independent spanning set of V, then $\mathcal{B}$ is a basis of V*

Proof: Let us denote $W = \operatorname{span}_{\mathbb{K}} \mathcal{B}$. Then, because $\mathcal{B} \subset V$ we have $W \subset V$. To show that $W = V$ we are going to assume otherwise and obtain a contradiction. So, suppose that $W \neq V$. This means that there exists $v_0 \in V$ such that $v_0 \notin W$. In particular, $v_0 \notin \mathcal{B} \subset W$. Now, consider the subset $\mathcal{B}' = \mathcal{B} \cup \{v_0\} \subset V$.

Then, by Corollary 1.3.5, $\mathcal{B}'$ is linearly independent.

Now, we use the maximal linear independence property of $\mathcal{B}$: since $\mathcal{B} \subset \mathcal{B}'$ and $\mathcal{B}'$ is linearly independent we must have $\mathcal{B}' = \mathcal{B}$, because $\mathcal{B}$ is a basis. Hence, $v_0 \in \mathcal{B}$. But this contradicts that fact that $v_0 \notin \mathcal{B}$. Therefore, our intial assumption, that $W \neq V$, must be false and we must necessarily have $W = V$.

Conversely, suppose that $\mathcal{B}$ is a linearly independent subset of V such that $\operatorname{span}_{\mathbb{K}} \mathcal{B} = V$. We want to show that $\mathcal{B}$ is a basis, so we must show that $\mathcal{B}$ satisfies the maximal linearly independent property of Definition 1.5.3.

Therefore, suppose that $\mathcal{B} \subset \mathcal{B}'$ and that $\mathcal{B}'$ is linearly independent; we must show that $\mathcal{B}' = \mathcal{B}$. Now, since $\mathcal{B} \subset \mathcal{B}'$ we have $V = \operatorname{span}_{\mathbb{K}} \mathcal{B} \subset \operatorname{span}_{\mathbb{K}} \mathcal{B}' \subset V$, using Lemma 1.3.9. Hence, $\operatorname{span}_{\mathbb{K}} \mathcal{B}' = V = \operatorname{span}_{\mathbb{K}} \mathcal{B}$. Assume that $\mathcal{B} \neq \mathcal{B}'$; we aim to provide a contradiction. Then, for each $w \in \mathcal{B}' \setminus \mathcal{B}$ we have $w \in \operatorname{span}_{\mathbb{K}} \mathcal{B}' = \operatorname{span}_{\mathbb{K}} \mathcal{B}$, so that there exists an expression

$$w = \lambda_1 b_1 + ... + \lambda_n b_n,$$

where $b_1, ..., b_n \in \mathcal{B}$. But this means that we have a nontrivial[27] linear relation among vectors in $\mathcal{B}'$ (recall that, as $\mathcal{B} \subset \mathcal{B}'$, we have $b_1, ..., b_n \in \mathcal{B}'$). However, $\mathcal{B}'$ is linearly independent so that no such nontrivial linear relation can exist. Hence, our initial assumption of the existence of $w \in \mathcal{B}' \setminus \mathcal{B}$ is false, so that $\mathcal{B}' = \mathcal{B}$. The result follows. □

Corollary 1.5.6. *Let V be a $\mathbb{K}$-vector space, $\mathcal{B} \subset V$ a basis of V. Then, for every $v \in V$ there exists a unique expression*

$$v = \lambda_1 b_1 + ... + \lambda_n b_n,$$

where $b_1, ..., b_n \in \mathcal{B}$, $\lambda_1, ..., \lambda_n \in \mathbb{K}$, $n \in \mathbb{N}$.

Proof: By Proposition 1.5.5, we have that $\operatorname{span}_{\mathbb{K}} \mathcal{B} = V$ so that, for every $v \in V$, we can write v as a linear combination of vectors in $\mathcal{B}$

$$v = \lambda_1 b_1 + ... + \lambda_n b_n, \quad b_1, ..., b_n \in \mathcal{B},$$

where we can further assume that none of $\lambda_1, ..., \lambda_n$ is equal to zero.

We need to show that this expression is *unique*: so, suppose that we can write v as a different linear combination

$$v = \mu_1 b_1' + ... + \mu_k b_k', \quad b_1', ..., b_k' \in \mathcal{B},$$

again assuming that none of the $\mu_1, ..., \mu_k$ are equal to zero.

Therefore, we have

$$\lambda_1 b_1 + ... + \lambda_n b_n = v = \mu_1 b_1' + ... + \mu_k b_k',$$

giving a linear relation

$$\lambda_1 b_1 + ... + \lambda_n b_n - (\mu_1 b_1' + ... + \mu_k b_k') = 0_V.$$

[27] Why is this linear relation nontrivial?

Thus, since $\mathcal{B}$ is linearly independent this linear relation must be trivial and, furthermore, since we have assumed that none of the λ's or μ's are zero, the only way that this can happen is if $n = k$ and, without loss of generality, $b_i = b'_i$ and $\lambda_i = \mu_i$. Hence, the linear combination given above is unique. $\square$

Corollary 1.5.7. *Let V be a $\mathbb{K}$-vector space, $\mathcal{B} = (b_1, \dots, b_n) \subset V$ an ordered basis containing finitely many vectors. Then, V is isomorphic to $\mathbb{K}^n$.*

Proof: This is just a simple restatement of Corollary 1.5.6: we define a function

$$[-]_{\mathcal{B}} : V \to \mathbb{K}^n \; ; \; v \mapsto [v]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix},$$

where

$$v = \lambda_1 b_1 + \dots + \lambda_n b_n,$$

is the unique expression for v coming from Corollary 1.5.6. Uniqueness shows that $[-]_{\mathcal{B}}$ is indeed a well-defined function.

It will be left to the reader to show that $[-]_{\mathcal{B}}$ is a bijective $\mathbb{K}$-linear morphism, thereby showing that it is an isomorphism. $\square$

Definition 1.5.8. Let V be a $\mathbb{K}$-vector space, $\mathcal{B} = \{b_1, \dots, b_n\} \subset V$ an ordered basis containing finitely many vectors. Then, the linear morphism

$$[-]_{\mathcal{B}} : V \to \mathbb{K}^n \; ; \; v \mapsto [v]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix},$$

introduced above is called the *$\mathcal{B}$-coordinate map* or *$\mathcal{B}$-coordinate morphism*.

The following Proposition provides yet another viewpoint of the idea of a basis: it says that a basis is a spanning set that satisfies a certain minimality condition.

Proposition 1.5.9. *Let V be a $\mathbb{K}$-vector space, $\mathcal{B} \subset V$ a basis of V. Then, $\mathcal{B}$ is a minimal spanning set - namely,*

- $\text{span}_{\mathbb{K}} \mathcal{B} = V$, *and*
- *if $\mathcal{B}' \subset \mathcal{B}$ is such that* $\text{span}_{\mathbb{K}} \mathcal{B}' = V$ *then $\mathcal{B}' = \mathcal{B}$.*

A proof of this Proposition will appear as a homework exercise.

Despite all of these results on bases of vector spaces we have still yet to give the most important fact concerning a basis: that a basis **exists** in an arbitrary $\mathbb{K}$-vector space.

The proof of the general case requires the use of a particularly subtle lemma, called Zorn's Lemma. You can read about Zorn's Lemma on Wikipedia and there you will see that Zorn's Lemma is equivalent to the Axiom of Choice (although the proof of this fact is quite difficult). You will also read on Wikipedia that Zorn's Lemma is logically equivalent to the existence of a basis for an arbitrary $\mathbb{K}$-vector space.

Theorem 1.5.10. *Let V be a $\mathbb{K}$-vector space. Then, there exists a basis $\mathcal{B} \subset V$ of V.*

Proof: Case 1: There exists a finite subset $E \subset V$ such that $\text{span}_{\mathbb{K}} E = V$.

In this case we will use the Elimination Lemma (Lemma 1.3.10) to remove vectors from E until we obtain a linearly independent set. Now, if E is linearly independent then E is a linearly independent spanning set of V and so, by Proposition 1.5.5, E is a basis of V. Therefore, assume that E is linearly dependent. Then, if we write E as an ordered set $E = \{e_1, \dots, e_n\}$, we can use Lemma 1.3.10 to remove a vector from E so that the resulting set is also a spanning set of V; WLOG, we can assume that the vector we remove is e_n. Then, define $E^{(n-1)} = E \setminus \{e_n\}$ so that we have $\text{span}_{\mathbb{K}} E^{(n-1)} = V$. If $E^{(n-1)}$ is linearly independent then it must be a basis (as it is also a spanning set). If $E^{(n-1)}$ is linearly dependent then we can again use Lemma 1.3.10 to remove a vector from $E^{(n-1)}$ so that the resulting set is a spanning set of V; WLOG, we can assume that the vector we remove is e_{n-1}. Then, define $E^{(n-2)} = E \setminus \{e_{n-2}\}$ so that we have $\text{span}_{\mathbb{K}} E^{(n-2)} = V$. Proceeding in a similar fashion as before we will either have that $E^{(n-2)}$ is linearly independent (in which case it is a basis) or it will be linearly dependent and we can proceed as before, removing a vector to obtain a new set $E^{(n-3)}$ etc.

Since E is a finite set this procedure must terminate after finitely many steps. The stage at which it terminates will have a produced a linearly independent spanning set of V, that is, a basis of V (by Proposition 1.5.5).

Case 2: There does not exist a finite spanning set of V.

In this case we must appeal to Zorn's Lemma: basically, the idea is that we will find a basis by considering a maximal linearly independent subset of V. Zorn's Lemma is a technical result that allows us to show that such a subset always exists and therefore, by definition, must be a basis of V. □

Theorem 1.5.11 (Basis Theorem)**.** *Let V be a $\mathbb{K}$-vector space and $\mathcal{B} \subset V$ a basis such that $\mathcal{B}$ has only finitely many vectors. Then, if $\mathcal{B}'$ is another basis of V then $\mathcal{B}'$ has the same number of vectors as $\mathcal{B}$.*

Proof: Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ and $\mathcal{B}' = \{b'_1, \ldots, b'_m\}$ be two distinct bases of V. Then, by Corollary 1.5.7, we have isomorphisms

$$[-]_{\mathcal{B}} : V \to \mathbb{K}^n, \quad \text{and} \quad [-]_{\mathcal{B}'} : V \to \mathcal{B}'.$$

Hence, we obtain an isomorphism (since the composition of two isomorphisms is again an isomorphism, by Lemma 0.2.4)

$$[-]_{\mathcal{B}'}^{-1} \circ [-]_{\mathcal{B}} : \mathbb{K}^n \to \mathbb{K}^m,$$

where $[-]_{\mathcal{B}'}^{-1} : \mathbb{K}^m \to V$ is the inverse morphism of $[-]_{\mathcal{B}'}$. Thus, using Theorem 1.4.11, we must have $n = m$, so that $\mathcal{B}$ and $\mathcal{B}'$ have the same size. □

Theorem 1.5.11 states that if V is a $\mathbb{K}$-vector space admitting a finite basis $\mathcal{B}$, then every other basis of V must have the same size as the set $\mathcal{B}$.

Definition 1.5.12. Let V be a $\mathbb{K}$-vector space, $\mathcal{B} \subset V$ a basis of V containing finitely many vectors. Then, the size of $\mathcal{B}$, $|\mathcal{B}|$, is called the *dimension of V (over $\mathbb{K}$)* and is denoted $\dim_{\mathbb{K}} V$, or simply $\dim V$ when no confusion can arise. In this case we will also say that V is *finite dimensional*. If V is a $\mathbb{K}$-vector space that does not admit a finite basis then we will say that V is *infinite dimensional*.

The Basis Theorem (Theorem 1.5.11) ensures that the dimension of a $\mathbb{K}$-vector space is a well-defined number (ie, it doesn't change when we choose a different basis of V).

Now that we have introduced the notion of dimension of a $\mathbb{K}$-vector space we can give one of the fundamental results of finite dimensional linear algebra.

Theorem 1.5.13. *Let V and W be $\mathbb{K}$-vector spaces such that* $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W < \infty$ *is finite. Then, V is isomorphic to W.*

This result, in essence, classifies all finite dimensional $\mathbb{K}$-vector spaces by their dimension. It tells us that any linear algebra question we can ask in a $\mathbb{K}$-vector space V (for example, a question concerning linear independence or spans) can be translated to another $\mathbb{K}$-vector space W which we know has the same dimension as V. This follows from Proposition 1.4.12.

This principle underlies our entire approach to finite dimensional linear algebra: given a $\mathbb{K}$-vector space V such that $\dim_{\mathbb{K}} V = n$, Theorem 1.5.13 states that V is isomorphic to $\mathbb{K}^n$ and Corollary 1.5.7 states that, once we have a basis $\mathcal{B}$ of V, we can use the $\mathcal{B}$-coordinate morphism as an isomorphism from V to $\mathbb{K}^n$. Of course, we still need to find a basis! We will provide an approach to this problem after we have provided the (simple) proof of Theorem 1.5.13.

Proof: The statement that V and W have the same dimension is just saying that any basis of these vector spaces have the same number of elements. Let $\mathcal{B} \subset V$ be a basis of V, $\mathcal{C} \subset W$ a basis of W. Then, we have the coordinate morphisms

$$[-]_{\mathcal{B}} : V \to \mathbb{K}^n \quad \text{and} \quad [-]_{\mathcal{C}} : W \to \mathbb{K}^n,$$

both of which are isomorphisms. Then, the morphism

$$[-]_{\mathcal{C}}^{-1} \circ [-]_{\mathcal{B}} : V \to W,$$

is an isomorphism between V and W. $\square$

References

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