

# Linear Algebra: Theory and Applications

Ethan Epperly

December 21, 2016



# Contents

<b>0</b>	<b>Prelude: Proof, Sets, and Linear Systems</b>	<b>1</b>
0.1	Proofs . . . . .	1
<b>1</b>	<b>Vectors</b>	<b>3</b>
1.1	A Physicist's Perspective: Vectors as Directed Line Segments . . . . .	3
1.2	A Computer Scientist's Perspective: Vectors as Ordered Lists . . . . .	6
1.3	A Mathematician's Perspective: Vectors as Objects Satisfying Axioms . .	8
1.4	Subspaces . . . . .	11
1.5	Span, Basis, and Linear Independence . . . . .	14
<b>2</b>	<b>The Second Chapter</b>	<b>17</b>



# Chapter 0

## Prelude: Proof, Sets, and Linear Systems

The purpose of this chapter is to give the briefest possible introduction to the tools we'll need to do linear algebra—namely proofs, sets, and linear systems.

### 0.1 Proofs

A proof is an argument meant to convince a mathematician of the truth of a claim. If we were to try to convince a computer of the truth of a mathematical claim, we would have to be absolutely rigid and precise in our language, using only mathematical symbols. In writing a convincing argument to a human, however, we should make judicious use of English sentences to tie our mathematical ideas together.

The simplest kind of proof is a **direct proof**. Suppose we want to show “If  $A$ , then  $Z$ ”.



# Chapter 1

## Vectors

Vectors are some of the most important objects in mathematics, and the study of vectors is at the core of linear algebra. However, the term "vector" doesn't mean the same thing to everyone. If you were to ask a physicist, they would tell you a vector is a mathematical object that has length and direction.<sup>1</sup> A computer scientist would say a vector is an ordered sequence of numbers. A mathematician would tell you a vector is an object that satisfies the axioms of a vector space.

What makes linear algebra so powerful is that the tools of linear algebra work regardless of which viewpoint you adopt. In fact, there are times when thinking in terms each of these definitions will be the most useful. This chapter will begin by defining vectors in each of these paradigms, and then explore the algebraic properties of vectors,

### 1.1 A Physicist's Perspective: Vectors as Directed Line Segments

Take it as a given that humans live in 3D space. What we mean by three-dimensional space will be treated formally later. For now, we'll take as a working definition of 3D space that we need three coordinates to specify the location of a point in 3D space.<sup>2</sup> Given this definition of 3D make the following definition:

**Definition 1** (Vector in Three Dimensions). A **vector** in three dimensions is a directed line segment in 3D space. By "line segment" we mean a contiguous section of a line with finite length, and by "directed" we mean that one endpoint of the line segment is the **tail** and the other endpoint is the **tip**. We think of a vector as going from its tail to its tip.

We usually denote vectors by lower case letters ( $u$ ,  $v$ , and  $w$  in particular are quite popular) with an arrow above them (i.e.  $\vec{u}$ ,  $\vec{v}$ ,  $\vec{w}$ ). Other texts use boldface in lieu of the arrow and still other texts use no typographic indication to distinguish vectors from numbers.

As we will see (or as you may have already since in physics), vectors are a very important construct. We will often want to use vectors in 1D and 2D physics problem

---

<sup>1</sup>A more erudite physicist may also define a vector as something that transforms like a vector under orthogonal transformations. More on this confusing statement later.

<sup>2</sup>It turns out the "three coordinates to specify a point" in 3D is entirely false. The mathematician David Hilbert showed that is possible to describe a position in 3D space with only 1 coordinate. This gives us a compelling reason to more formally define what it means to be three-dimensional. More on this later.

as well as 3D problems. Further, some models of physics consider higher dimensional spaces.<sup>3</sup> Thus, we want to extend our definition of vectors to any dimensional space.

**Definition 2** (Vector in  $n$  Dimensions). A **vector** in  $n$  dimensions is a directed line segment in  $n$ -dimensional space.

What makes vectors useful is that we can also define operations like addition, subtraction, and multiplication on vectors. Before we define these operations, however, we must define what we mean by equality of vectors.

**Definition 3** (Equality of Vectors). Two vectors are said to be **equal** with they have the same length and same direction. An equivalent way of defining equality is to say two vectors  $\vec{v}$  and  $\vec{w}$  are equal if we can move  $\vec{v}$  on top of  $\vec{w}$  such that the tail of  $\vec{v}$  is on the tail of  $\vec{w}$  and the tip of  $\vec{v}$  is on the tip of  $\vec{w}$ .

What this means is that where a vector is placed in  $n$ -dimensional space doesn't matter—all that matters is its length and its direction. For convenience, we generally draw vectors with their tails at the **origin**, which is just an arbitrary point in space from which we orient ourselves.

**Proposition 1.** *Let an origin  $O$  be given. For every point  $P$  in space, there is a unique vector  $\vec{x}_P$  such that the tip of  $\vec{x}_P$  lies on  $P$  when the tail of  $\vec{x}_P$  is placed on  $O$ .  $\vec{x}_P$  is said to be the **position vector** of  $P$ .*

Thus, once we fix an origin  $O$ , every point has a unique position vector and every vector corresponds to a unique position. We can now go about defining addition of vectors.

**Definition 4** (Addition of Vectors). Let  $\vec{u}$  and  $\vec{v}$  be vectors. The **sum** of  $\vec{u}$  and  $\vec{v}$  is a new vector  $\vec{u} + \vec{v}$  obtained by placing the tail of  $\vec{v}$  on the tip of  $\vec{u}$  and drawing a new vector from the tail of  $\vec{v}$  to the tip of  $\vec{u}$ .

How is this definition useful? Consider the following thought experiment. Suppose I begin at a point  $P$  and walk in a straight line to a point  $Q$ . Let  $\vec{y}$  be the vector from  $P$  to  $Q$ , and  $\vec{x}_P$  be the position vector of  $P$ . Then the position vector of  $Q$  is  $\vec{x}_Q = \vec{x}_P + \vec{y}$ .

**Proposition 2.** *Addition of vectors is commutative. That is, for all vectors  $\vec{u}$  and  $\vec{v}$ ,  $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ .*

To convince yourself of this, draw yourself a picture representing  $\vec{u} + \vec{v}$  and  $\vec{v} + \vec{u}$  in two dimensions. Do you see why addition must be commutative?

There is a special vector, which we will call  $\vec{0}$  such that  $\vec{v} + \vec{0} = \vec{v}$  for all vectors  $\vec{v}$ . This "zero vector" corresponds to the vector for which the tail and the tip are the same point. The zero vector is the only vector with 0 length.

We now go about defining vector subtraction. Think about subtraction of counting numbers using a number line. If we want to subtract 3 from 5, we simply start at the position 5 and move three steps to the left to get 2. Another way of thinking about subtracting 3 from 5 is to introduce negative numbers. Then  $5 - 3$  is just the same as saying  $5 + (-3)$  which is 2. The number  $-3$  means count three to the left, where positive 3 means count three to the right. Notice, if we add 3 and  $-3$  we get 0. Since  $-3$  is the number you add to 3 to get 0, we say  $-3$  is the **additive inverse** of 3. In the same way, every vector  $\vec{v}$  has an additive inverse  $-\vec{v}$ .

---

<sup>3</sup>General relativity, for example, is based on four dimensional Minkowski space.



**Proposition 3.** *For every vector  $\vec{v}$  there is a unique vector  $-\vec{v}$  such that  $\vec{v} + (-\vec{v}) = \vec{0}$ .*

*Proof.* Say  $\vec{v}$  goes from the origin  $O$  to a point  $P$ . Let  $-\vec{v}$  be the vector from  $P$  to  $O$ . Then, clearly by definition of addition,  $\vec{v} + (-\vec{v}) = \vec{0}$ . Uniqueness follows immediately from the definition of vector equality.  $\square$

The geometric interpretation of  $-\vec{v}$  is clear:  $-\vec{v}$  is the vector with the same length of  $\vec{v}$  and the exact opposite direction.

**Proposition 4.** *For all vectors  $\vec{v}$ ,*

$$\vec{v} = -(-\vec{v}).$$

*That is, the additive inverse of the additive inverse of  $\vec{v}$  is  $\vec{v}$ . Additionally,*

$$\vec{0} = -\vec{0}$$

*And  $\vec{0}$  is the only vector equal to its own inverse.*

Having defined the additive inverse, the definition of subtraction is obvious.

**Definition 5** (Subtraction of Vectors). Let  $\vec{u}$  and  $\vec{v}$  be vectors. Then

$$\vec{u} - \vec{v} = \vec{u} + (-\vec{v})$$

Having defined vector-vector addition and subtraction, we naturally move on to multiplication. It turns out that there isn't a simple geometric way to multiply two vectors. We will study two types of vector-vector multiplication later.<sup>4</sup> There is, however, a way of defining number-vector multiplication in a natural way.

**Definition 6** (Scalar Multiplication). Let  $\vec{u}$  be a vector. If  $c \geq 0$ , then  $c\vec{u}$  is a vector with the same direction as  $\vec{u}$  with  $c$  times the length of  $\vec{u}$ . If  $c < 0$ , then  $c\vec{u} = (-c)(-\vec{u})$ .

That is,  $c\vec{v}$  is the vector  $\vec{v}$  scaled to be  $|c|$  times as long, with its direction flipped if  $c < 0$ . Since the real number  $c$  scales the vector  $\vec{v}$  we say  $c$  a **scalar**. In fact, since scaling vectors is something real numbers often do in linear algebra, we say all real numbers are scalars. We call scalar-vector multiplication "scalar multiplication".

**Proposition 5.** *For all vectors  $\vec{v}$ ,*

$$\begin{aligned} 1\vec{v} &= \vec{v} \\ (-1)\vec{v} &= -\vec{v} \\ 0\vec{v} &= \vec{0} \end{aligned}$$

**Proposition 6** (Distributive Laws). *For all real numbers  $c$  and  $d$  and vectors  $\vec{v}$  and  $\vec{u}$ ,*

$$\begin{aligned} c(\vec{u} + \vec{v}) &= c\vec{u} + c\vec{v} \\ (c + d)\vec{u} &= c\vec{u} + d\vec{u} \end{aligned}$$

---

<sup>4</sup>Both of these vector-vector multiplications—the dot and cross products—have unsatisfactory properties. The dot product of two vectors is not a vector at all, but a real number. The cross product of two vectors is another vector, but the cross product is only defined for vectors in 3D and isn't commutative or associative.

**Proposition 7** (Associative Law). *For all real numbers  $c$  and vectors  $\vec{v}$  and  $\vec{u}$ ,*

$$(cd)\vec{v} = c(d\vec{v})$$

In summary, vectors (as we've defined them here) have the following properties.

**Theorem 8.** *Vectors, defined as directed line segments in  $n$ -dimensional space, satisfy the following properties:*

1.  $\vec{u} + \vec{v}$  is a vector in  $n$ -dimensional space
2.  $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
3.  $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
4. There is a vector  $\vec{0}$  such that, for all  $\vec{u}$ ,  $\vec{u} + \vec{0} = \vec{u}$
5. For every vector  $\vec{u}$  there is a vector  $-\vec{u}$  such that  $\vec{u} + (-\vec{u}) = \vec{0}$
6.  $c\vec{u}$  is a vector in  $n$ -dimensional space
7.  $(cd)\vec{u} = c(d\vec{u})$
8.  $(c + d)\vec{u} = c\vec{u} + d\vec{u}$
9.  $c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$
10.  $1\vec{u} = \vec{u}$

## 1.2 A Computer Scientist's Perspective: Vectors as Ordered Lists

In the C++ programming language, a vector is a container object which stores an ordered sequence of other objects. In C++, these objects can be anything the programmer wants—integers, real numbers, strings, student records, and even other vectors. In mathematics, for the moment at least, we will restrict ourselves to real numbers. This leads us to make an entirely different definition of vectors:

**Definition 7** (Vector in  $n$  Dimensions). A **vector** in  $n$  dimensions is an ordered list of  $n$  real numbers.

Mathematicians, always fond of notational shorthand, use  $\mathbb{R}$  to denote the set of real numbers. Since  $n$ -dimensional vectors are just lists of  $n$  real numbers, we use the notation  $\mathbb{R}^n$  to denote the set of all  $n$ -dimensional vectors. Several notation are used for  $n$ -dimensional vector. Let  $\vec{v}$  be an  $n$ -dimensional vector. Then  $v_1$  is the first element,  $v_2$  is the second element, and  $v_n$  is the  $n$ th element. The vector  $\vec{v}$  is notated by any of the following

$$\vec{v} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = (v_1, v_2, \dots, v_n) = \langle v_1, v_2, \dots, v_n \rangle^5$$

---

<sup>5</sup>This last notation is unfortunate because it conflicts with our notation for span

The first of these is called the **row vector** representation and the second the **column vector** representation. The row and column vector notations are often done with parentheses rather than square brackets. Vector addition and subtraction and scalar multiplication couldn't be defined more naturally.

**Definition 8** (Arithmetic on Vectors). Let  $\vec{u}$  and  $\vec{v}$  be vectors and  $c$  a scalar. Then

$$\vec{u} + \vec{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix} \quad \vec{u} - \vec{v} = \begin{bmatrix} u_1 - v_1 \\ u_2 - v_2 \\ \vdots \\ u_n - v_n \end{bmatrix} \quad c\vec{u} = \begin{bmatrix} cu_1 \\ cu_2 \\ \vdots \\ cu_n \end{bmatrix}$$

We now have two working definition of vectors: vectors as directed line segments and vectors as ordered lists. As you might expect, these definitions are equivalent—that is every vector-as-directed-line-segment can be represented as an ordered list and every vector-as-order-list can be represented as a directed line segment.

The way we get between these two representations is by what is called a **basis**, which we will introduce rather informally here and present a more detailed treatment later. For now, we'll say a basis in  $n$ -dimensional space is a set of any  $n$  vectors<sup>6</sup>  $\vec{v}_1, \vec{v}_2, \vec{v}_3, \dots, \vec{v}_n$  such that I can make any  $n$ -dimensional vector  $\vec{u}$  by going some amount  $c_1$  in the  $\vec{v}_1$  direction, some amount  $c_2$  in the  $\vec{v}_2$  direction, etc. In more mathematical terms, a basis is a set of vectors  $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \dots, \vec{v}_n\}$  such that any vector  $\vec{u}$  can be written as

$$\vec{u} = c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_n\vec{v}_n$$

The conversion between vectors as directed line segments to vectors as ordered lists is then just

$$\underbrace{\vec{u} = c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_n\vec{v}_n}_{\text{Vector as direct line segment}} \mapsto \underbrace{(c_1, c_2, c_3, \dots, c_n)}_{\text{Vector as ordered list}}$$

And the conversion from vectors as ordered lists to vectors as directed line segments is just

$$\underbrace{(c_1, c_2, c_3, \dots, c_n)}_{\text{Vector as ordered list}} \mapsto \underbrace{\vec{u} = c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_n\vec{v}_n}_{\text{Vector as direct line segment}}$$

We call the coefficients  $c_1, c_2, \dots, c_n$  the **coordinates** of  $\vec{u}$  with respect to the basis  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ . We normally pick our basis vectors to perpendicular vectors with length 1. In 2 or 3 dimensions, we call the line our first basis vector is on the  $x$  axis, the the line our second basis vector is on the  $y$  axis, and the line our first basis vector is on the  $z$  axis. We use  $\hat{i}$ ,  $\hat{j}$ , and  $\hat{k}$  to denote the  $x$ ,  $y$ , and  $z$  basis vectors respectively. When we're in a higher dimensional space, we often use  $\hat{e}_1, \hat{e}_2, \hat{e}_3, \dots, \hat{e}_n$  for our basis.

Notice that the conversion from vectors-as-directed-line-segments to vectors-as-ordered-lists is not unique: if we pick a different set of basis vectors, the conversion between vectors-as-directed-line-segments to vectors-as-ordered-lists will be different. Thus, when we're converting between these two representations we always need to specify which basis we're using.

---

<sup>6</sup>Here, vectors-as-directed-line-segments

*Remark.* If every vector-as-directed-line-segment can be represented as a vector-as-ordered-list and visa versa, does it mean that these two objects are the same thing? Many would say yes. We shall take the somewhat controversial position that they *aren't* the same thing. We choose to make this distinction since directed line segments have an existence entirely separate from their coordinate representation with respect to a particular set of basis vectors. We will call the set of vectors as directed line segments  $\mathbb{L}^n$  to distinguish it from the set of vectors as ordered lists  $\mathbb{R}^n$ .

### 1.3 A Mathematician's Perspective: Vectors as Objects Satisfying Axioms

Suppose you have a friend who excitedly comes up to you and claims to have invented a new number system. He says, in his number system, that there is a smallest number “orez.” Each number has a unique successor: “orez”'s successor is “eno” and “eno”'s successor is “owt”. Further, you can add numbers together: “orez + owt = owt” and “eno + owt = eerht”. The number “orez” has the unique property that for any number  $n$  “orez +  $n$  =  $n$ ”.

Clearly, your friend has not invented a new system at all but rather just relabeled the counting numbers with new funny names. This admittedly silly situation speaks to an important point, however. When we do mathematics, we don't care what an object “is” in some fundamental sense, but *what its properties are*. We could just as easily do number theory using the numbers “0, 1, 2, 3, ...” as “orez, eno, owt, eerht, ...”.

So we return to the question of “What is a vector?” one more time. To a mathematician, what is a vector? *Something that satisfies the properties of a vector*. Which properties? *The properties outlined in Theorem 8*.

More precisely, let's consider a set of objects which I will call  $V$ . We say the elements of  $V$  are vectors and the entire set  $V$  is a **vector space**. I define an operation  $+$  that takes two vectors from  $V$  and gives me back another vector in  $V$ . We also define an operation  $\cdot$  that takes a scalar from  $\mathbb{R}$  and a vector from  $V$  and returns another vector in  $V$ . Whatever my operations are, I make sure they satisfy all the properties in Theorem 8.

**Definition 9.** A **vector space** is a set  $V$ , whose elements are called vectors, along with operations  $+$  and  $\cdot$  such that:

1. For all  $\vec{u}, \vec{v} \in V$ ,  $\vec{u} + \vec{v} \in V$
2. For all  $\vec{u}, \vec{v} \in V$ ,  $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
3. For all  $\vec{u}, \vec{v}, \vec{w} \in V$ ,  $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
4. There is a vector  $\vec{0}$  such that, for all  $\vec{u} \in V$ ,  $\vec{u} + \vec{0} = \vec{u}$
5. For every vector  $\vec{u} \in V$  there is a vector  $-\vec{u} \in V$  such that  $\vec{u} + (-\vec{u}) = \vec{0}$
6. For all  $\vec{v} \in V$ ,  $c\vec{u} \in V$
7. For all  $c, d \in \mathbb{R}$  and  $\vec{u} \in V$ ,  $(cd)\vec{u} = c(d\vec{u})$
8. For all  $c, d \in \mathbb{R}$  and  $\vec{u} \in V$ ,  $(c + d)\vec{u} = c\vec{u} + d\vec{u}$

9. For all  $c \in \mathbb{R}$  and  $\vec{u}, \vec{v} \in V$ ,  $c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$
10. For all  $\vec{u} \in V$ ,  $1\vec{u} = \vec{u}$

The advantage of doing is are initially unclear. We've taken two perfectly good definitions of vectors that are tangible and concrete and replaced them by a much more abstract definition. The advantage of the mathematical definition is that if we *only* use the 10 properties in the definition, any theorems we prove about one vector space applies to *every* vector space. It turns out that there are many other interesting vector spaces other than just  $\mathbb{R}^n$  and  $\mathbb{L}^n$ . Let's run through some examples of vector spaces.

**Example 1.**  $\mathbb{R}^n$  and  $\mathbb{L}^n$  are both vector spaces. Of special note,  $\mathbb{R}$  is a vector space.

**Example 2.** The set of polynomials of degree less than or equal to  $n$  form a vector space. We will call this space  $\mathbb{P}_n$ . The set of all polynomials also forms a vector space, which we will call  $\mathbb{P}$ .

**Example 3.** Ok this one's a bit strange. We're gonna call this one  $\mathbb{R}^*$ . The elements of  $\mathbb{R}^*$  are just positive real numbers. The “addition” operator (which I will denote by  $\oplus$ ) is *multiplication* of real numbers. So  $a \oplus b = a \times b$ . The “scalar multiplication” operator (which I will denote by  $\odot$ ) is *exponentiation*. So  $c \odot a = a^c$ . Now you don't just have to believe me when I say that, we can prove it.

**Proposition 9.**  $\mathbb{R}^*$  is a vector space.

*Proof.* Let  $x, y, z$  be in  $\mathbb{R}^*$  (these are our “vectors”) and  $c, d$  be scalars.

- (1)  $x \oplus y = xy$  which is another positive real number.
- (2)  $x \oplus y = xy = yx = y \oplus x$
- (3)  $x \oplus (y \oplus z) = x \oplus (yz) = x(yz) = (xy)z = (x \oplus y)z = (x \oplus y) \oplus z$
- (4) This one's really strange. In  $\mathbb{R}^*$  our “0” element (or additive inverse) is actually 1! To see this, simply take  $x \oplus 1 = x \times 1 = x$ .
- (5) The inverse of  $x$  is just  $x^{-1}$ . To verify this, simply take  $x \oplus x^{-1} = xx^{-1} = 1$ . (Remember 1 is our additive inverse.)
- (6)  $c \odot x = x^c$  which is another positive real number.
- (7)  $(cd) \odot x = x^{cd} = (x^c)^d = d \odot x^c = d \odot (c \odot x)$
- (8)  $c \odot (x \oplus y) = c \odot (xy) = (xy)^c = x^c y^c = x^c \oplus y^c = (c \odot x) \oplus (c \odot y)$
- (9)  $(c + d) \odot x = x^{c+d} = x^c x^d = x^c \oplus x^d = (c \odot x) \oplus (d \odot x)$
- (10)  $1 \odot x = x^1 = x$

□

*Remark.* Notice that when we add and multiply scalars to scalars we use the normal arithmetic operations  $+$  and  $\times$ . It's only when we add “vectors” and multiply vectors by scalars that we use our vector operations  $\oplus$  and  $\odot$ .

**Example 4.** The set of functions from any set  $S$  to the real numbers forms a vector space. Call this vector space  $S^{\mathbb{R}}$ .

**Example 5.** The set of continuous functions on an interval  $[a, b]$  is a vector space. This vector space is denoted  $C([a, b])$  or  $C^0([a, b])$ . The set of functions continuous on the entire real line is denoted  $C(-\infty, \infty)$  or  $C^0(-\infty, \infty)$ .

**Example 6.** The set of solutions to the differential equation

$$y''(x) + p(x)y'(x) + q(x)y(x) = 0$$

Is a vector space.

**Example 7.** Consider the system of  $m$  linear equations in  $n$  unknowns

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

Consider the vector  $\vec{x} = (x_1, x_2, \dots, x_n)$  to be a **solution** of the linear equation if the  $x_1, x_2, \dots, x_n$  satisfy the linear system (make all the equations true) when plugged into the system. The set of solutions to this linear system is a vector space.

*Remark.* If any of the right hand sides of the linear system were to be made non-zero, the solutions would no longer form a vector space. A system of linear equations is said to be **homogenous** if all the right hand sides are 0, and **inhomogenous** otherwise.

**Problem 1.** Often, a non-example of a mathematical idea can be just as elucidating as an example of it. The following sets fail to be vector spaces. Which of the axioms do they fail? Which of them do they not fail?

1. The set of integers  $\mathbb{Z}$ . (Hint: Think about scalar multiplication.)
2. Solutions to an *inhomogenous* linear equation under normal vector addition and scalar multiplication.
3. Solutions to the differential equation  $y''(x) + p(x)y'(x) + q(x)y(x) = r(x)$  where  $r(x) \not\equiv 0$ .
4. Polynomials of degree  $n$ . (Hint: How is this space different than  $\mathbb{P}_n$ ?)
5. Let  $\mathfrak{R}$  be the set of all rotations of a point about the origin in two-dimensional space. Addition of rotations is such successive application of rotations. For example, a  $60^\circ$  rotation followed by a  $30^\circ$  rotation is a  $90^\circ$  rotation. Scalar multiplication by  $c$  means apply the rotation  $c$  times. For example, 2.5 times a rotation of  $30^\circ$  is  $75^\circ$ . (Hint: A rotation of  $0^\circ$  equals a rotation of  $360^\circ$  equals a rotation of  $720^\circ$  and so on.)

Invent a non-example of a vector space yourself. What axioms does it fail and not fail?

**Problem 2.** Pick a vector space from the text that we didn't prove was a vector space and prove it is a vector space.

**Problem 3.** A sequence is just a list of numbers. We use the notation  $(a_n)$  to denote the sequence  $a_1, a_2, \dots$ . In this notation  $a_1$  is the first number,  $a_2$  is the second, and so on.

1. Prove the set of sequences forms a vector space under the vector addition  $(a_n) + (b_n) = (a_n + b_n)$  and scalar multiplication  $c(a_n) = (ca_n)$ .
2. A sequence  $(a_n)$  is said to converge to 0 if

$$\lim_{n \rightarrow \infty} a_n = 0$$

Prove the set of sequences converging to 0 forms a vector space.

## 1.4 Subspaces

As we saw in the previous section, the solutions to a homogeneous system of linear equations in  $n$  unknowns is a vector space. We represent these solutions by vectors in  $\mathbb{R}^n$ . But not every vector in  $\mathbb{R}^n$  is a solution to the system of linear equations.<sup>7</sup> So what we have is a smaller vector space contained within a larger vector space. Smaller vector spaces contained within larger vector spaces is a very important concept in linear algebra, and motivates the following definition:

**Definition 10.** Let  $V$  be a vector space. A subset<sup>8</sup>  $W$  of  $V$  is said to be a **(vector) subspace** if  $W$  is also a vector space with the same vector operations  $+$  and  $\cdot$  as  $V$ .

So we're led naturally to a question: given a set  $W \subseteq V$ , how can we check if  $W$  is a subspace? Well we just need to check the 10 properties stated in Definition 9. In fact, we only need to check axioms 1 and 6, the so-called closure axioms, since axioms 2-5 and 7-10 are satisfied since  $W$  uses the same vector operations  $+$  and  $\cdot$  as  $V$ .

**Theorem 10.** A set  $W \subseteq V$  is a subspace of  $V$  if, and only if,

1. For every  $\vec{u}, \vec{v} \in W$ ,  $\vec{u} + \vec{v} \in W$ , and
2. For every  $\vec{u} \in W$  and every scalar  $c \in \mathbb{R}$ ,  $c\vec{u} \in W$ .

*Remark.* Axioms 1 and 6 are often referred to as the **closure** axioms. This is because

**Corollary 11.** If  $W$  is a subspace of  $V$ ,  $\vec{0} \in W$ .

*Proof.* Let  $\vec{w}$  be any vector in  $W$ . Then, by property 6,  $0\vec{w} = \vec{0} \in W$ . □

**Example 8.** Let  $V$  be any vector space. Then the set containing the zero vector,  $\{\vec{0}\}$  is a subspace of  $V$ . This is the so-called **trivial subspace** of  $V$ .  $V$  is also a subspace of itself. A subspace that is neither  $V$  nor the trivial subspace is said to be a proper subspace.

---

<sup>7</sup>Unless all the coefficients  $a_{ij}$  are 0.

<sup>8</sup>Recall that a set  $W$  is a subset of a set  $V$  if every element in  $W$  is in  $V$ . When  $W$  is a subset of  $V$ , we write  $W \subseteq V$ .

Before we look at subspaces in more abstract vector spaces, let's think about what subspaces exist in the space of vectors-as-directed-line-segments. We'll begin with  $\mathbb{L}^2$ . As always, we have the trivial subspace and  $\mathbb{L}^2$  itself. So let's suppose we have a subspace  $W$  with a vector  $\vec{w}$  in it. Since  $W$  is closed under scalar multiplication, all the vectors on the line containing  $\vec{w}$  must be in  $W$ . So the line containing  $\vec{w}$  is the smallest subspace containing  $\vec{w}$ . Suppose now that we have another vector  $\vec{v}$  in  $W$  that is not on the line containing  $\vec{w}$ , then we must have  $W = \mathbb{L}^2$ . Can you convince yourself why?<sup>9</sup> In summary, the only subspaces of  $\mathbb{L}^2$  are: the origin, lines containing the origin, and  $\mathbb{L}^2$  itself.

Very similarly, one can deduce that the subspaces of  $\mathbb{L}^3$  are precisely the origin, the lines containing the origin, the planes containing the origin, and  $\mathbb{L}^3$  itself. What's striking about this result is how "boring" the subspaces of  $\mathbb{L}^3$  are. The only proper subspaces you have are lines and planes. And a plane through the origin is starting to sound a lot like  $\mathbb{L}^2$ , which can be thought of as a plane. So what we have is that all the proper subspaces of  $\mathbb{L}^3$  are either planes that look like  $\mathbb{L}^2$  or lines that look like  $\mathbb{L}^1$ . Ok here are some more "abstract" examples of subspaces.

**Example 9.**  $\mathbb{P}_1$  is a subspace of  $\mathbb{P}_2$  which is a subspace of  $\mathbb{P}_3$  and so on.  $\mathbb{P}_n$  is a subspace of  $\mathbb{P}$  for all  $n$ .

**Example 10.** Let  $\alpha \in [a, b]$  be a real number. Then the set of all continuous functions  $f$  such that  $f(\alpha) = 0$  is a subspace of  $C([a, b])$ .

Suppose we have two subspaces  $U$  and  $W$  of a vector space  $V$ . Notice both  $U$  and  $W$  contain the trivial subspace  $\{\vec{0}\}$  as a subspace. We might ask the question if there is any "larger" subspace of  $V$  that is both a subspace of  $U$  and a subspace of  $W$ . We might be even bolder and ask what the *largest* subspace of both  $W$  and  $U$  is. This question is answered by the next theorem.

**Theorem 12.** *If  $U$  and  $W$  are subspaces of a vector space  $V$ , then  $U \cap W$ <sup>10</sup> is also a subspace of  $V$ . Moreover,  $U \cap W$  is the largest subspace of  $V$  that is a subspace of both  $U$  and  $W$  in the sense that if  $Z$  is a subspace of  $U$  and  $W$  then  $Z \subseteq U \cap W$ .*

*Proof.* As always, to check  $U \cap W$  is a subspace, we must simply check the closure axioms. Let  $\vec{v}_1$  and  $\vec{v}_2$  be vectors in  $U \cap W$ . Then  $\vec{v}_1$  and  $\vec{v}_2$  are in  $W$  (by definition of  $\cap$ ) and so  $\vec{v}_1 + \vec{v}_2$  is in  $W$  (because  $W$  is a subspace.) Similarly,  $\vec{v}_1$  and  $\vec{v}_2$  are in  $U$  so  $\vec{v}_1 + \vec{v}_2$  is in  $U$ . Since  $\vec{v}_1 + \vec{v}_2 \in U$  and  $\vec{v}_1 + \vec{v}_2 \in W$ , we have  $\vec{v}_1 + \vec{v}_2 \in U \cap W$  (by definition of  $\cap$ .) Closure under scalar multiplication is very similar and is left to you to check for yourself.

Let  $Z$  be a subspace of  $U$  and  $W$ . Let  $\vec{z}$  be in  $Z$ . Then since  $Z$  is a subspace of  $U$  and  $W$  we must have  $\vec{z} \in U$  and  $\vec{z} \in W$ . So  $\vec{z} \in U \cap W$ . Thus  $Z \subseteq U \cap W$  (by definition of  $\subseteq$ ).<sup>11</sup> □

If we consider  $\mathbb{L}^3$  as an example, the fact that intersection of subspaces is a subspace is very intuitive. Imagine intersecting two planes that go through the origin. If the planes are coincide with each other, we get another plane, and if the planes don't coincide with

---

<sup>9</sup>Developing a strong geometric intuition for the algebra of  $\mathbb{L}^n$  can be challenging business, but is ultimately deeply rewarding for studying this subject. Take advantage of the fact that you can draw two-dimensional vectors on a sheet of paper.

<sup>10</sup>Recall that  $U \cap W$  is the set of all elements that are members of both  $U$  and  $W$ .

<sup>11</sup>Get used to these "maximality" or "minimality" arguments because they are important for the next section.



each other, we get a line. Since lines and planes going through the origin are subspaces in  $\mathbb{L}^3$ , we have intersection of subspaces is subspaces.

A very natural follow-up question to “What is the largest subspace that is contained within both  $U$  and  $W$ ?” is “What is the smallest subspace that contains both  $U$  and  $W$ ?” To answer this question, we need a new definition.

**Definition 11** (Addition of Subspaces). If  $U$  and  $W$  are subspaces of  $V$ , then  $U + W$  is the set

$$\{\vec{v} \in V : \vec{v} = \vec{u} + \vec{w} \text{ for some } u \in U \text{ and } w \in W\}$$

That is, it is the set of obtained by adding one element from  $U$  and one element from  $W$ .

**Theorem 13.** *If  $U$  and  $W$  are subspaces then  $U + W$  is a subspace of  $V$  and  $U + W$  is the smallest subspace that contains both  $U$  and  $W$  in the sense that if  $Z$  is a subspace containing both  $U$  and  $W$  then  $U + W \subseteq Z$ .*

**Problem 4.** Suppose we have a set of objects  $S$  a relation  $\preceq$ . We think of the relation as “comparing” objects in  $S$  so that either  $s \preceq s'$  or  $s \not\preceq s'$  (read  $s$  precedes  $s'$  or  $s$  does not precedes  $s'$ ).<sup>12</sup> We say  $\preceq$  is a **partial ordering** if three things are true:

1.  $s \preceq s$  for all  $s \in S$  (Reflexive)
2.  $s \preceq s'$  and  $s' \preceq s$  if and only if  $s = s'$  (Antisymmetry)
3.  $s \preceq s'$  and  $s' \preceq s''$  implies  $s \preceq s''$  (Transitivity)

The less-than-or-equal-to relation  $\leq$  is a partial ordering on the set  $\mathbb{R}$  and the subset relation  $\subseteq$  is a partial ordering on any collection of sets. Define  $U \preceq W$  if  $U$  is a subspace of  $W$ . Show that  $\preceq$  is a partial ordering on the subspaces of  $V$  (by showing it satisfies those three properties.)

A partial ordering is said to be **total** if, for every two elements  $s, s' \in S$  we have  $s \preceq s'$  or  $s' \preceq s$ . Find vector spaces  $V$  for when  $\preceq$  is and isn't total.

**Problem 5.** Prove Theorem 13.

**Problem 6.** Let  $V$  be a vector space and let  $U$  and  $W$  be subspaces of  $V$ . Prove the following facts about addition of subspaces:

1.  $U + U = U$
2.  $U + W = W + U$
3.  $U + V = V$

**Problem 7.** Let  $V$  be a vector space and let  $U$  and  $W$  be subspaces of  $V$ . Define  $U - W$  by

$$U - W = \{\vec{v} \in V : \vec{v} = \vec{u} - \vec{w} \text{ for some } u \in U \text{ and } w \in W\}$$

Show  $U - W = U + W$ .

---

<sup>12</sup>Here  $'$  has no relation to the derivative from calculus.  $s$  and  $s'$  are simply two variables.

## 1.5 Span, Basis, and Linear Independence

*Remark.* The study of mathematics is full of so-called “equivalent conditions”. If we say condition 1 and condition 2 are equivalent, we mean that any object satisfying condition 1 must also satisfy condition 2, and any object satisfying condition 2 must also satisfy condition 1.

Vector spaces can be tricky beasts to wrangle because they have an infinite number of vectors in them.<sup>13</sup> What would be helpful is if we could find some way of expressing vectors in some way by only a *finite* number of vectors, because finite sets are much easier to deal with than infinite sets.

Say we have a set  $S$  of vectors. This set may have either a finite or infinite number of vectors in it. Let  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$  be a set of vectors in  $S$  and let  $c_1, c_2, \dots, c_n$  be scalars. Then

$$\vec{u} = c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_n\vec{v}_n$$

is said to be a **linear combination** of the vectors in  $S$ .<sup>14</sup> So if  $S = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ , then  $\vec{v}_1$ ,  $\vec{v}_1 + \vec{v}_2 + \vec{v}_3$ , and  $\pi\vec{v}_1 + e\vec{v}_2$  are all linear combinations of  $S$ . A natural question to ask is: Given a set  $S$ , what are all the possible linear combinations? We have a name for the set of all possible linear combinations; we call it the **span**.

**Definition 12.** Let  $S$  be a set of vectors from a vector space  $V$ . Then the **span** of  $S$ , denoted  $\langle S \rangle$ , is

$$\langle S \rangle = \{\vec{v} \in V : \vec{v} \text{ is a linear combination of the vectors in } S\}$$

If a set  $S'$  is a subset of  $\langle S \rangle$  then we say the set  $S$  **spans**  $S'$ .<sup>15</sup>

**Theorem 14.** The span of  $S$  is a subspace of  $V$ . Moreover,  $\langle S \rangle$  is the smallest subspace containing  $S$  in the sense that if  $W$  is a subspace then  $S \subseteq W$  implies  $\langle S \rangle \subseteq W$ .

*Proof.* As always, to prove  $\langle S \rangle$  is a subspace, we must check the closure axioms. First let  $\vec{u}_1$  and  $\vec{u}_2$  be in  $\langle S \rangle$ . Then, by definition of span, there are vectors  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$  and  $\vec{w}_1, \vec{w}_2, \dots, \vec{w}_m$  in  $S$  and scalars  $c_1, c_2, \dots, c_n$  and  $k_1, k_2, \dots, k_m$  such that

$$\vec{u}_1 = c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_n\vec{v}_n \quad (1.1)$$

$$\vec{u}_2 = k_1\vec{w}_1 + k_2\vec{w}_2 + \dots + k_m\vec{w}_m \quad (1.2)$$

So then

$$\vec{u}_1 + \vec{u}_2 = c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_n\vec{v}_n + k_1\vec{w}_1 + k_2\vec{w}_2 + \dots + k_m\vec{w}_m \quad (1.3)$$

Notice that (1.3) is saying that  $\vec{u}_1 + \vec{u}_2$  is, itself, a linear combination of vectors in  $S$ . Therefore,  $\vec{u}_1 + \vec{u}_2$  is in the span of  $S$ . Now to check closure under scalar multiplication, we consider

$$\begin{aligned} k\vec{u}_1 &= k(c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_n\vec{v}_n) \\ &= (kc_1)\vec{v}_1 + (kc_2)\vec{v}_2 + \dots + (kc_n)\vec{v}_n \end{aligned} \quad (1.4)$$

<sup>13</sup>There is one vector space with a finite number of elements in it. Can you think of what it is?

<sup>14</sup>Though the set  $S$  is allowed to be either finite or infinite, the vectors  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$  must be a finite collection of vectors. What would it even mean to add up an infinite number of vectors?

<sup>15</sup>We also make the quite strange definition that the span of the empty set is the trivial subspace. That is,  $\langle \emptyset \rangle = \{\vec{0}\}$ .

Which is another linear combination of vectors so  $k\vec{u}_1 \in \langle S \rangle$ . Therefore,  $\langle S \rangle$  is a subspace of  $V$ .

Now let  $W$  be a subspace of  $V$  such that  $S \subseteq W$ . Let  $\vec{x}$  be any vector in the span of  $S$ . Then, by definition,  $\vec{x}$  must be of the form

$$\vec{x} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n \quad (1.5)$$

For vectors  $\vec{v}_1, \dots, \vec{v}_n$  in  $S$ . Then  $\vec{x}$  must be in  $W$  because  $W$ , a subspace, is closed under addition and scalar multiplication and contains all the vectors  $\vec{v}_1, \dots, \vec{v}_n$ . Thus,  $\langle S \rangle \subseteq W$ .  $\square$

Before we delve further into the span into the abstract let's take a moment to think about the span geometrically for  $\mathbb{L}^2$  and  $\mathbb{L}^3$ . Remember, the span is the smallest subspace containing a set of vectors. So the span of a single vector  $\vec{v}$  is just the line going through the origin and  $\vec{v}$ . The span of a vector  $\vec{v}$  and another vector  $\vec{w}$  not on the span of  $\vec{v}$  is the plane containing both  $\vec{v}$  and  $\vec{w}$ . Thus, if we are in  $\mathbb{L}^2$ , the set  $\{\vec{v}, \vec{w}\}$  spans  $\mathbb{L}^2$ . If we are in  $\mathbb{L}^3$ , then set  $\{\vec{v}, \vec{w}\}$  is a plane going through the origin,  $\vec{v}$ , and  $\vec{w}$ . If we add another vector  $\vec{x}$  not on  $\langle \{\vec{v}, \vec{w}\} \rangle$ , then  $\{\vec{v}, \vec{w}, \vec{x}\}$  spans  $\mathbb{L}^3$ .

At the beginning of this section, we posed a simple problem: find a way of expressing the infinite vectors in a vector space by some combination of a finite number of vectors. And here we have an answer: just find a finite spanning set for a vector space. For example, the space  $\mathbb{R}^2$  is spanned by  $S = \{(0, 1), (1, 0)\}$ , which means every vector in  $\mathbb{R}^2$  can be expressed as a linear combination of  $(0, 1)$  and  $(1, 0)$ . To see this, try to write  $(5, 2)$  as a linear combination of  $(0, 1)$  and  $(1, 0)$ .

But there is a problem with using mere spanning sets to express all the vectors in our space. To see this, notice  $S_1 = \{(1, 0), (0, 1)\}$  and  $S_2 = \{(3, 0), (1, 2), (5, 7), (10, 12)\}$  both span  $\mathbb{R}^2$ . So really, we don't want to find just a spanning set for our vector space, but a *minimal spanning set*.<sup>16</sup>

**Definition 13.** A **basis** for a vector space  $V$  is a *minimal spanning set* for  $V$ . In other words, a basis is a set  $\beta$  such that  $\langle \beta \rangle = V$  but  $\langle S \rangle \neq V$  for any proper subset  $S \subset \beta$ .

**Example 11.**  $\{1\}$  is a basis for  $\mathbb{R}$ . In fact, if  $a$  is any non-zero real number, then  $\{a\}$  is a basis for  $\mathbb{R}$ .

**Example 12.**  $\{e_1, e_2, \dots, e_n\}$  is a basis for  $\mathbb{R}^n$ , where  $e_i$  is a vector of all zeros except a 1 in position  $i$ . For example, in  $\mathbb{R}^3$ ,  $e_1 = (1, 0, 0)$ ,  $e_2 = (0, 1, 0)$ , and  $e_3 = (0, 0, 1)$ . This basis is called the standard basis of  $\mathbb{R}^n$ .

**Example 13.**  $\{1, x, x^2, \dots, x^n\}$  is a basis for  $\mathbb{P}_n$ , the space of all polynomials of degree less than or equal to  $n$ . While beyond the scope of this discussion, the space  $\mathbb{P}$  of *all* polynomials does not have a *finite* basis, but does have the basis  $\{1, x, x^2, \dots\}$ .

While Definition 13 is an *intuitive* characterization of what a basis is, it's quite an annoying definition to actually *check* if a particular set  $\beta$  is a basis. To find an equivalent, and much more computationally useful way of determining whether a given spanning set is a basis, we introduce the notion of *linear (in)dependence*.

<sup>16</sup>The choice of "minimal" rather than "smallest" is intentional. Saying smallest would imply that there is some "smallest" spanning set, which is simply not true. Both  $\{(1, 0), (0, 1)\}$  and  $\{(5, 0), (1, 2)\}$  are both minimal spanning sets, and one is not smaller (in terms of numbers of elements) than the other.

**Definition 14.** A set  $S$  is linearly independent if no vector in  $S$  can be written as a linear combination of the other vectors in  $S$ .

**Problem 8.** Show that if  $S \subseteq S'$ , then  $\langle S \rangle$  a subspace of  $\langle S' \rangle$ .

## Chapter 2

### The Second Chapter

