Linear Algebra for Physicists

Flip Tanedo

flip.tanedo@ucr.edu ⬨

Department of Physics & Astronomy, University of California, Riverside, CA 92521

Abstract

Lecture notes for Physics 17, a course on linear algebra in preparation for upper-division undergraduate physics coursework at UC Riverside.

Contents

1 Logistics

At any time in this course, you should feel comfortable asking either of the following questions:

1. Is it obvious that...?
2. Why is this significant?

The first question is the way to ask for on-the-spot clarification—I will either appreciate that I did not properly explain a subtle point, *or* I will explain the intuition[1] for why something should be obvious. The second question is a way to remind me that I may have *lost sight of the forest for the trees*: I want this course to *mathematically connect big ideas in physics*. Asking this question is a reminder that making those connections justifies the hard mathematical work we will put into the course.

Finally, I want to comment on the word *obvious*. I write this often. It is somewhat dangerous because it can come off as being arrogant: *this is so obvious to me, if you do not understand you must be deficient.* This is never the reason why I use that word. Instead, the word *obvious* serves a very practical purpose. The goal of this class is not just to be able to "do stuff" (e.g. diagonalize a symmetric matrix), but to also build that intuition that comes from a deeper understanding how the mathematics works. In this sense, every time I write the word *obvious* it is a flag: I am saying something that—with the proper perspective—should be self-evident. If it is not self-evident, then you should stop to interrogate why it is not self-evident. Most likely there is something where a change in perspective may (1) make it obvious, and (2) in so doing deepen your understanding of the subject. So when you see the word 'obvious,' I want you to do a quick check to confirm whether or not the statement is indeed obvious. If it is not, then welcome the opportunity to learn.[2]

2 Motivation

Here are three incredibly significant equations in physics:

$$\mathbf{F} = m\mathbf{a} \qquad R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G_\mathrm{N}}{c^4}T_{\mu\nu} \qquad \hat{H}|\Psi\rangle = E|\Psi\rangle \ . \qquad (2.1)$$

These are Newton's force law, Einstein's field equations, and the Schrödinger equation. They govern classical physics, general relativity, and quantum theory.

Each equation looks rather unique: they seem to each be speaking their own mathematical language. Newton's law is written with boldfaced vectorial quantities $\mathbf{F} = (F_x, F_y, F_z)^T$ that should look very familiar to any physics undergraduate. Einstein's equation has these funny μ and ν indices on every term—have you seen these before? Do they look intimidating? If you ever want to make your equations look "technical" and "physicsy," you should dress them up with indices. The Schrödinger equation has no indices, but instead has these funny angle-brackety things... and that $\hat{H}$ looks suspicious. Where did H get a hat, and what is the content of this equation other than $\hat{H} = E$?

Each of these equations turns out to be a "vectorial" equation. Each one is actually shorthand for a number of equations. Newton's equation is shorthand for three equations, one for each component. Einstein's equation is shorthand for 16 equations, one for each combination of the indices μ and ν that run over four values[3]. The Schrödinger equation is

---

[1] Your sense of mathematical and physical intuition is incredibly valuable. This is one of the key traits that makes a physics training unique.

[2] There is, of course, the possibility that what I have written is *not* obvious. For example, if I have made a typo... in which case, please let me know.

[3] The four values are the three directions of space and one direction of time.

shorthand for an *infinite* number of equations, one for each allowed energy of a quantum system.

The mathematical formalism that unifies these different ideas (and notations) of 'vector' is called linear algebra. It may sound humble: after all, "linear" systems are *easy*, aren't they? Did we not just spend years of our lives learning fancy things like *calculus* and *differential equations* to deal with functions that are more complicated than *lines*? In some sense, yes: linear algebra is about lines and planes in different numbers of dimensions.[4] However, linear algebra turns out to be far more richer than what you may be used to from high school.

In this course we will see how the three equations in (2.1) are connected by the mathematics of linear algebra. We will dive into the different notation and shamelessly pass between $\mathbf{v}$, v^i, and $|v\rangle$ to describe the same abstract vector. We will connect to the mathematical description of *symmetry* and see how it is an underlying theme in our descriptions of nature. And we will do all of this in a way that will make the instructors of the linear-algebra-for-mathematicians course and linear-algebra-for-engineers course vomit a little in disgust. Consider that one of privileges of being a physicist.

3 Basics

3.1 Pre-conceptions

If this were a mathematics course, then we would start by very carefully defining words like *vector* and *matrix*. As a physics student, you already have a working definition of these words. It is probably something like this:

> A vector has a magnitude and a direction. We write a vector as an array of three numbers arranged in a column. A matrix is an array of nine numbers arranged in a 3×3 block. There is a rule for how to apply (multiply) the matrix to the vector to produce a new vector.

The problem is that you already know too much to learn linear algebra as a mathematics student. You have already seen the tip of the iceberg and so have preconceptions about what vectors are and how they work. You may remember from freshman mechanics that forces are vectors. So are momenta and velocities. You may also recall the idea of a force field—like the electric field—which is actually a whole bunch of vectors: one for each point in space. Examples of matrices are a little more subtle: you may recall that you can represent rotations as matrices. Speaking of rotations, there was another thing that showed up called the moment of inertia *tensor*. It looked like a matrix, but we never called it the "moment of inertia matrix." What the heck is a tensor, anyway?

And so, you see that starting this course like a mathematics course could cause trouble. The mathematics professor would start by defining a vector. That definition will say nothing about magnitudes or directions, and will not even say anything about arrays of numbers. That definition will clash with the hard-earned intuition that you built from your physics

---

[4]On the other hand: a good chunk of the calculus that we do is also implicitly linear. Physicists often Taylor expand and keep only the $\mathcal{O}(\varepsilon)$ term. Integration boils down to summing trapezoids whose angley-bits are given by the first derivative of a function... the linear component.

education thus far. It will be perplexing, and may make you feel rather unhappy. What do these mathematicians know, anyway? Or maybe its the physics that is wrong, or have we just completely misunderstood everything and we are just now noticing that we are hopelessly lost? We begin to spiral into a black hole of confusion.

> Fortunately, *this is not a mathematics course.*

As a consequence, we will not give a rigorous definition of a vector. We start with a familiar definition of vectors and lay out which qualities are general, and which properties are specific. Then we will come to appreciate the approximation that *"everything is a vector."* So let us start with something comfortably familiar, even though it constitutes only the simplest example of a vector.

3.2 Real Three-Vectors

Let us write $\mathbf{v}$ to be a vector. This is a standard convention for writing a vector. In this course we will use a few different notations for vectors according to convenience. Notation is neither physics nor mathematics, it is simply a shorthand for a physical or mathematical idea.

In fact, let us focus on a particular type of vector: **real three-vectors**. These are the familiar vectors that we can write as a column of three numbers that effectively represent the coordinates in three-dimensional space:

$$\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} , \tag{3.1}$$

where x, y, and z are real numbers. These numbers are called the **components** of the vector $\mathbf{v}$.

Exercise 3.1. *There is something very perverse about this "vector." The variable names x, y, and z imply that $\mathbf{v}$ is something that physicists like to call a "position vector." If you say this to a mathematician they will vomit. By the end of this course, you should appreciate why the notion of a position vector makes no sense.* Hint: *You may have some intuition for this already: a velocity vector tells you about the instantaneous motion of a particle relative to its present position. Try to write the analogous statement for a "position vector.*[5]*"*

This three-dimensional space is called [three-dimensional] **real space** and we write it as $\mathbb{R}^3$. This is because a vector is an element of three-dimensional real space specified by *three* real numbers.

Three-dimensional real space is an example of a **vector space**, which is just a stupidly formal way of saying that it is where vectors live. Vectors are *elements* of a vector space. A vector space is the set of all possible allowed vectors of a given type. For $\mathbb{R}^3$, the vector space is composed of all possible triplets of real numbers.

---

[5]I am not a mathematician, but you see that even I have to write "position vector" in condescending quotation marks. In lecture I use even more condescending air quotes.

Example 3.1. *It should be no surprise that we can imagine real two-dimensional space, $\mathbb{R}^2$. This is a vector space where each vector may be written as two real numbers. You can also imagine writing real four-dimensional space, $\mathbb{R}^2$, or complex two dimensional space, $\mathbb{C}^2$.*

From the above example, you should have some intuition for what the **dimension** of a vector space means: the dimension counts how many numbers you need to specify a vector. For real vector spaces, $\mathbb{R}^d$, the dimension is the number d. We will always assume that d is a positive integer.[6]

3.3   Mathematical properties

We now make some general statements about vector spaces. These apply to all vector spaces, not just $\mathbb{R}^3$, but you can keep $\mathbb{R}^3$ in mind as we go over them.

First we should make clear that there are now two different kinds of objects: *vectors* and *numbers*. We will have all sorts of notation for vectors, but let us write them with a boldfaced Roman letter for now, e.g. **v**. We typically write numbers as lowercase italicized Roman letters like a or sometimes Greek letters like α.

You already know everything there is to know about numbers.[7] Most relevant is that you can multiply numbers with each other (including division, the inverse of multiplication) and you can add them together (including subtraction). For the first part of this course, we will focus on real numbers, $\mathbb{R}$. Later we will also allow for complex numbers, $\mathbb{C}$.

What about vectors? Vectors have the following properties:

1. You can add vectors. This assumes that they are in the same vector space. The sum of two vectors is also a vector in the same vector space.

2. Vector addition is associative. This means that in the sum $\mathbf{v} + \mathbf{w} + \mathbf{u}$, it does not matter if you add $(\mathbf{v} + \mathbf{w})$ first and then add $\mathbf{u}$, or if you take $\mathbf{v}$ and then add it to $(\mathbf{w} + \mathbf{u})$. This is the kind of 'obvious' property that we tend to take for granted.

3. Vector addition is commutative. $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$. This is also kind of obvious. But recall that matrix multiplication is not commutative.

4. There is a zero vector, $\mathbf{0}$, that does leaves any other vector unchanged under addition. $\mathbf{v} + \mathbf{0} = \mathbf{v}$. This should be totally obvious. I bet you even know what the components of $\mathbf{0}$ are.

5. There is an additive inverse (negative vectors). If $\mathbf{v}$ is a vector, then $-\mathbf{v}$ is a vector and satisfies $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.

6. You can multiply vectors by numbers. This is called rescaling or scalar multiplication. All the usual properties of multiplication by numbers holds: associativity, commutivity, distributive law.

---

[6]The notion of a non-integer-dimensional space does show up occasionally. These do not even have to be particularly exotic: you can look up the dimension of a fractal.

[7]Formally, what I mean by 'number' is what mathematicians call a **field**. This simply means some objects where one can add, subtract, multiply, and divide as you would expect. This term is a little tedious for us because physicists usually mean something else when they say 'field.' Usually we mean something like the electric field or the field associated with the Higgs boson.

Example 3.2. *The first property implies that once you have identified one vector in a vector space, $\mathbf{v}$, then you can immediately have an infinite number of vectors. This is because $2\mathbf{v} = \mathbf{v} + \mathbf{v}$ must also be a vector. Then $3\mathbf{v} = 2\mathbf{v} + \mathbf{v}$ must also be a vector. And so forth.*

Let us go over a few of these properties.

3.4 Notation: Indices

One theme in this course is that we will repeatedly refine our notation to suit our needs. Let us introduce an *index* notation where we write the components of vectors $\mathbf{v}$ and $\mathbf{w}$ as follows:

$$\mathbf{v} = \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} \qquad\qquad \mathbf{w} = \begin{pmatrix} w^1 \\ w^2 \\ w^3 \end{pmatrix} . \tag{3.2}$$

We see that a boldfaced Roman letter, u, corresponds to a vector. The *components* of the vector are u^1, u^2, u^3. The "x-component" of $\mathbf{u}$ is called u^1: we use the same letter as the vector, but italicized rather than boldfaced. The upper index is *not* some kind of power, it simply means "the first component."

Example 3.3. *If you see $\mathbf{s}$, this is understood to be a vector that has multiple components. If it is a three-vector, it has three components. If you see s^2, then this means that this is the second component of the vector $\mathbf{s}$. The component of a vector is a number.*

You may worry that this notation introduces ambiguity. If we see q^2, is this the square of some number q, or is it the second component of some vector $\mathbf{q}$? The answer depends on context. You should avoid choosing variable names where there is ever the potential for ambiguity. If you have a vector that you call $\mathbf{q}$, then do not use the letter q for anything else.

3.5 Arithmetic

All vector spaces allow addition and subtraction. This is defined component-wise. The sum of $\mathbf{v}$ and $\mathbf{w}$ is

$$\mathbf{v} + \mathbf{w} = \begin{pmatrix} v^1 + w^1 \\ v^2 + w^2 \\ v^3 + w^3 \end{pmatrix} . \tag{3.3}$$

What this means is that the *sum* of two vectors is also a vector. That means that if $\mathbf{v}$ and $\mathbf{w}$ are vectors in $\mathbb{R}^3$, then $(\mathbf{v} + \mathbf{w})$ is a vector in $\mathbb{R}^3$. The components of the vector $(\mathbf{v} + \mathbf{w})$ are simply the sum of the components of $\mathbf{v}$ and $\mathbf{w}$. A few formal properties that generalize to all vector spaces:

- Vector addition is associative. This means that in the sum $\mathbf{v} + \mathbf{w} + \mathbf{u}$, it does not matter if you add $(\mathbf{v} + \mathbf{w})$ first and then add $\mathbf{u}$, or if you take $\mathbf{v}$ and then add it to $(\mathbf{w} + \mathbf{u})$. This is the kind of 'obvious' property that we tend to take for granted.
- Vector addition is commutative. $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$. This is also kind of obvious. But recall that matrix multiplication is not commutative.

- There is a zero vector, **0**, that does leaves any other vector unchanged under addition. **v** + **0** = **v**. This should be totally obvious. The components of **0** are obviously all zero.

- There is an additive inverse (negative vectors). If **v** is a vector, then −**v** is a vector and satisfies **v** + (−**v**) = **0**.[9]

3.6   Notation: Indices again

We can express this using index notation. Let us call this sum **u** so that **u** ≡ **v** + **w**. Then we can succinctly write the components of **u** in one line:

$$u^i = v^i + w^i \ . \tag{3.4}$$

The variable i is called an **index**. What values does the index take? In this example, it is clear that (3.4) holds for $i = 1, 2, 3$. That is, i takes values from 1 to the dimension of the space. The typical convention is that we do not have to state the range of index values because it should be understood from the space itself.

With that in mind, it should be clear that if **q** is the difference of two vectors, then the components of **q** may be succinctly written:

$$\mathbf{q} = \mathbf{v} - \mathbf{w} \qquad\qquad \Leftrightarrow \qquad\qquad q^i = v^i - w^i \ . \tag{3.5}$$

In fact, as physicists we typically use the two statements above interchangeably. If you know the components of a vector, then you know the vector.

3.7   Rescaling: multiplication by a number

Another operation that exists in a vector space is rescaling: we multiply a vector by a number. Let α be a number. If you want to nitpick, let us restrict α to be a real number. If we have a vector **v** with components v^i, then $\alpha\mathbf{v}$ is also a vector.[8] The components of $\alpha\mathbf{v}$ are

$$(\alpha v)^i = \alpha v^i \ , \tag{3.6}$$

by which we mean

$$(\alpha\mathbf{v}) = \begin{pmatrix} \alpha v^1 \\ \alpha v^2 \\ \alpha v^3 \end{pmatrix} \ . \tag{3.7}$$

The parenthesis on the left-hand side is sloppy notation to mean "the vector that is the vector **v** rescaled by the number α." Another way of saying this is that there is a vector **w** ≡ $\alpha\mathbf{v}$ whose components are $w^i = \alpha v^i$.

Example 3.4. *Let us do one explicit example with numbers. Suppose the vectors* **v** *and* **w** *have components*

$$\mathbf{v} = \begin{pmatrix} 4.2 \\ -2.6 \\ 7.0 \end{pmatrix} \qquad\qquad\qquad \mathbf{w} = \begin{pmatrix} 5.3 \\ 2.1 \\ -2.5 \end{pmatrix} \ . \tag{3.8}$$

---

[8]"Also a vector" means that it is also an element of the vector space; so $(\alpha\mathbf{v})$ is an element of $\mathbb{R}^3$ is **v** is an element of $\mathbb{R}^3$.

7

I can rescale each vector by different numbers: $\alpha = 10$, $\beta = 2$. We can consider the vector that comes from adding these rescaled vectors:

$$\mathbf{u} \equiv \alpha\mathbf{v} + \beta\mathbf{w} \ . \tag{3.9}$$

The second component of $\mathbf{u}$ is $u^2 = -26 + 4.2 = -21.8$.

At this point it is useful to define some jargon. A **scalar** is a number. This is in contrast to vectors (and matrices and tensors) which we can think of as arrays of numbers. In fact, every time you see the word scalar, you should just think "number." Another name for 'rescaling a vector by a number' is *scalar multiplication.*

3.8 Operations that are not (yet) allowed

In these definitions, we make a big deal about how the sum of two vectors *is also a vector.* Or how the rescaling of a vector by a number *is also a vector.* This is in contrast to operations that are either not allowed or that do not produce vectors. An example of an operation that is not allowed is adding together vectors from two different vector spaces. The following proposed sum of a vector in $\mathbb{R}^3$ with a vector in $\mathbb{R}^2$ does not make sense:

$$\begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} + \begin{pmatrix} w^1 \\ w^2 \end{pmatrix} = ? \tag{3.10}$$

If you find yourself adding vectors from two different vector spaces, then you have made a mistake.

Another operation that requires care is rescaling a real vector by a complex number. If $\mathbf{v}$ is a vector in $\mathbb{R}^3$ and we try to multiply it by a complex number, $\alpha = 2 + 3i$, then the resulting "vector" is not a vector in $\mathbb{R}^3$:

$$(\alpha\mathbf{v})^i = (2 + 3i)v^i \notin \mathbb{R} \ , \tag{3.11}$$

that is: the components of $\alpha\mathbf{v}$ are not real numbers, and so this cannot be an element of aa vector space that is *defined* to have real components. Later on we will generalize to the case of *complex vector spaces*, but we will treat that with some care.[9]

Thus far, we have introduced the *nouns* of this course: vectors. We have identified a few *verbs* that let us do things with these vectors:

1. Addition takes two vectors in a vector space and returns a vector in the same vector space.

2. Rescaling takes a vector and a number and returns a vector in the same vector space.

We can rewrite this in the language of *mappings* (or *functions*) as follows. Let V be a vector space, say $V = \mathbb{R}^3$. Let us write $\mathbb{R}$ mean [real] numbers. Then the above statements tell us that addition and rescaling can be thought of as maps:

---

[9]If you want to be fancy, you can replace 'number' with the mathematical notion of a field. Both the real numbers and the complex numbers are examples of fields. In my mind a field is just a class of number, though mathematicians have fancier definitions.

1. Vector addition: $V \times V \to V$

2. Rescaling: $V \times \mathbb{R} \to V$.

Do not be intimidated by the $\times$ symbol here. This "mapping" notation means nothing more and nothing less than the statements above.

We now know everything there is to know about the vector space $\mathbb{R}^3$. We want to learn more about functions (maps) that involve this vector space. How can we combine vectors and numbers to produce other vectors and numbers? What about more complicated objects like matrices and tensors?

3.9 Euclidean three-space

You may object: *wait! I know there are more things you can do with three-vectors!* You remember that there are two types of vector multiplication that we use in physics. The **dot product** and the **cross product**.

In $\mathbb{R}^3$, the **dot product** is a map $V \times V \to \mathbb{R}$. That means it takes two vectors and returns a number. The particular number that it returns is typically *defined* to be

$$\mathbf{v} \cdot \mathbf{w} = \sum_i v^i w^i = v^1 w^1 + v^2 w^2 + v^3 w^3 . \tag{3.12}$$

The dot product generalizes in linear algebra. It is often called an **inner product** or a **metric** and has a few different notations that we will meet. What is important is that this dot/inner product is an *additional* mathematical function that we attach to a vector space.

Three-dimensional real space combined with the dot product/inner product/metric (3.12) is called Euclidean three-space. In general, a vector space combined with a 'dot product' is called a **metric space**. The word metric should invoke some etymological notion of measurement of distance. Indeed, the dot product is a tool that tells us how 'close' two vectors are to one another—though it is not yet obvious how.

Example 3.5. *Let $\mathbf{r} = (x, y, z)$ be a "position vector" of a point relative to the origin.*[10] *Then the distance of the point from the origin is*

$$d = \sqrt{\mathbf{r} \cdot \mathbf{r}} = \sqrt{x^2 + y^2 + z^2} . \tag{3.13}$$

This gives a notion of how the dot product is related to measuring distances, but it turns out to be a bit of a red herring! The real sense in which the dot product measures the 'closeness' of two vectors is the sense in which it defines an angle between those vectors. (See below.)

The **cross product** is a different story. You may remember the cross product from such hits as[11] angular momentum, $\mathbf{r} \times \mathbf{p}$. It looks like a map that takes two vectors and spits out another vector, $V \times V \to V$. Indeed, this is the case in Euclidean three-space. However, it had some funny properties compared to the dot product. For example, there was something weird with the order of the two vectors: $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$. It is also a bit funny that the direction of

---

[10]It is dangerous to use the phrase "position vector," see Exercise 3.1.

[11]https://tvtropes.org/pmwiki/pmwiki.php/Main/YouMightRememberMeFrom

the output vector is completely different[12] from the directions of the input vectors. It will turn out that this product does not generalize as simply as the dot product, though there is a generalization called the **wedge product** which is outside the scope of this course.[13]

Exercise 3.2. *Define the generalization of the Euclidean three-space metric to Euclidean space in d dimensions. (Easy.)*

Exercise 3.3. *Try to define a generalization of the cross product in two-dimensional Euclidean space. Reflect on why this is much less natural than the generalization of the dot product.*

3.10   Length in Euclidean three-space

Euclidean three-space is real space combined with the Euclidean dot product, (3.12). The [Euclidean] **magnitude** (length) of a three vector $\mathbf{v}$ as $|\mathbf{v}|$ in Euclidean three-space. The magnitude is defined to be

$$|\mathbf{v}| = \sqrt{\mathbf{v} \cdot \mathbf{v}} \ . \tag{3.14}$$

This definition generalizes to Euclidean d-dimensional space with the appropriate generalization of the dot product.

Example 3.6. *Consider the vector*

$$\mathbf{v} = \begin{pmatrix} 3 \\ -4 \\ 0 \end{pmatrix} \tag{3.15}$$

in Euclidean three-space. The magnitude of $\mathbf{v}$ is $|\mathbf{v}| = 5$.

Some references prefer to use the double bar notation, $||\mathbf{v}||$ for the length of a vector. This is to distinguish it from the absolute value of a number, $|-3| = 3$. We will be even more perverse: sometimes we will write v to mean the magnitude of $\mathbf{v}$ when there is no ambiguity.

Example 3.7. *Consider the vector*

$$\mathbf{v} = \begin{pmatrix} -1 \\ 3 \\ 2 \end{pmatrix} \ . \tag{3.16}$$

Then the magnitude *of $\mathbf{v}$ is $|\mathbf{v}| = \sqrt{14}$. We could also write this as $v = \sqrt{14}$, but we should be careful when we write things like v^2 which could either mean the second component of $\mathbf{v}$—which is 3—or the square of the magnitude—which is 14.*

---

[12]The technical meaning of 'completely different' is *orthogonal*, which we define below with the help of the metric.

[13]That is not to say that the wedge product is not relevant in phsyics. The wedge product features prominently in a mathematical field called **differential geometry**, which is in turn the framework for general relativity. The wedge product is related to defining volumes and integration measures.

We see that the dot product (metric) allows us to define length. Because the length of a vector is a number, we can divide the vector $\mathbf{v}$ by its its length $|\mathbf{v}|$ to obtain a **unit vector**, $\hat{\mathbf{v}}$:

$$\hat{\mathbf{v}} = \frac{1}{|\mathbf{v}|}\mathbf{v} \ . \tag{3.17}$$

The right-hand side is simply scalar multiplication by $|\mathbf{v}|^{-1}$. Unit vectors are useful for identifying directions.

Example 3.8. *In grade school one may have learned that a vector is an arrow that has a magnitude and a direction. Unit vectors encode the 'direction' of a vector.*

Example 3.9. *Let $\mathbf{v}$ be defined as in* (3.17). *The unit vector associated with $\mathbf{v}$ is*

$$\hat{v} = \begin{pmatrix} 3/5 \\ -4/5 \\ 0 \end{pmatrix} \ . \tag{3.18}$$

3.11 Angles in Euclidean three-space

Let $\mathbf{v}$ and $\mathbf{w}$ be two vectors in Euclidean three-space. The [Euclidean] angle between these two vectors, θ, is

$$\cos\theta \equiv \hat{\mathbf{v}} \cdot \hat{\mathbf{w}} = \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{v}||\mathbf{w}|} \ . \tag{3.19}$$

Exercise 3.4. *Confirm that this matches the definition of the angle between two vectors that you learned in your youth.*

The above definition of the angle between two vectors is general for any metric space—that is, a vector space equipped with a dot product.

Example 3.10. *The angle between two vectors defines the sense in which two vectors are close to one another. This is the sense in which the dot product (metric) lets you measure the "distance" between two vectors. Note that this is completely different from the notion of distance between two points in space,* (3.13).

4 Index Notation and Summation Convention

There's something that physicists do that tend to drive mathematicians crazy: we write a generic *component of a vector* and refer to it as if it were the vector itself. It is a fairly harmless peccadillo:[14] if I say "the vector v^i," then it is not hard to guess that I mean "the vector $\mathbf{v}$ which has components that I label v^i."

---

[14]There are times when you can get into trouble if you drink your own Kool Aid, so to speak. The reason is that the *component v^i* is simply a number, whereas $\mathbf{v}$ is a vector. Some manipulations are only allowed for numbers and not vectors, and you should be clear that you mean 'the component v^i' if you are treating it like a number, and not 'the *vector* whose components are v^i.' See Example 4.3.

The reason why we have this culture is that this index notation ends up being so damn convenient. In addition to vectors, we will have other objects that have indices: dual vectors, matrices, and tensors. When we write everything in with indices, we can "see" properties of these objects that are not obvious without the indices. Specifically, we can see *how an object transforms under symmetries*. In this course, we will focus on *rotations* of vectors and their generalizations.

There's a second reason why indices are convenient: they allow us to use **Einstein summation convention**. This is a notational shortcut that introduces upper and lower indices to convey sums. Consider, for example, the "matrix multiplication" of a row vector $\widetilde{\mathbf{w}}$ on a column vector $\mathbf{v}$. Nevermind the formal definition of "row vector" as opposed to "column vector." Let's just write it out in components where it is obvious for $\mathbb{R}^3$

$$\widetilde{\mathbf{w}} = \begin{pmatrix} w_1 & w_2 & w_3 \end{pmatrix} \qquad \mathbf{v} = \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} \qquad \widetilde{\mathbf{w}}\mathbf{v} = w_1 v^1 + w_2 v^2 + w_3 v^3 \ . \qquad (4.1)$$

The final expression is familiar, right? It follows the usual rules of matrix multiplication for a matrix that happens to be one row and three columns. We notice that we chose to write the components of $\widetilde{\mathbf{w}}$ with lower indices—this is the convention. Row vectors (which have many names) have indices written as subscripts while column vectors have indices written as superscripts. There is no mathematics here, just a choice of notation. The result of the multiplication is simply a number, which we can write as a sum:

$$\widetilde{\mathbf{w}}\mathbf{v} = \sum_{i=1}^{3} w_i v^i \equiv w_i v^i \ . \qquad (4.2)$$

On the right-hand side we have *defined* the summation convention: *whenever there is exactly one upper index and exactly one lower index with the same letter, we should understand that there is a sum over that index over all of its allowed values.* We call pairs of repeated indices where one is upper and one is lower **contracted indices**.

The value $w_i v^i$ is simply a number. It is not a vector. It does not have any "vectorial" (tensorial) structure. It is not an element of the vector space $\mathbb{R}^3$. It does not transform under rotations. It is *just a number*. In other words, $w_i v^i$ behaves like an object with *no indices*. Contracted indices "cancel each other out."

This is significant because we will see that indices tell us how objects transform. Evidently, column vectors and row vectors transform differently since one has an upper index and one has a lower index. Further, when we contract the two indices, we end up with something with no indices: a number that does not transform at all. This may seem like notational overkill—trust me, it is worth building this notation now. We will use it over and over.

Example 4.1. *Matrices M have the following index structure: $M^i{}_j$. There is a first index and a second index—the order matters. The first index is upper, and the second index is lower. Matrix multiplication boils down to a contraction of indices:*

$$(M\mathbf{v})^i = M^i{}_j v^j \ . \qquad (4.3)$$

Let us read this equation carefully. First, $M\mathbf{v}$ is a vector. The ith component of this vector is $(M\mathbf{v})^i$. How is this related to the components of M and $\mathbf{v}$? The right-hand side tells us that we simply take the sum:

$$M^i{}_j v^j = M^i{}_1 v^1 + M^i{}_2 v^2 + M^i{}_3 v^3 \ . \tag{4.4}$$

Example 4.2. *From the above example, you can then excuse the glib statement: "the* vector $M^i{}_j v^j$." *As we explained above,* $M^i{}_j v^j$ *is not a vector, but a component of a vector. However, the point is that even though there are three indices, two of them are contracted so the object effectively only has one upper index. This is the index structure of a vector.*

Exercise 4.1. *Consider the following vector, row vector, and matrix:*

$$\mathbf{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \qquad \widetilde{\mathbf{w}} = \begin{pmatrix} 4 & 5 & 6 \end{pmatrix} \qquad M = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \ . \tag{4.5}$$

These have index structure v^i, w_i, *and* $M^i{}_j$ *respectively. Note that the first index of a matrix is the row and the second is the column, thus* $M^1{}_2 = 2$ *while* $M^2{}_1 = 4$. *Calculate the following:* $(wM)_2$, $(Mv)^1$, $(MM)^1{}_2$. *Here* MM *is understood to be the square of the matrix* M, $(M^2)^i{}_j = M^i{}_k M^k{}_j$.

Example 4.3. *It should be clear that*

$$w_i M^i{}_j = w_1 M^1{}_j + w_2 M^2{}_j + w_3 M^3{}_j = M^1{}_j w_1 + M^2{}_j w_2 + M^2{}_j w_2 = M^i{}_j w_i \ . \tag{4.6}$$

After all, each of the components w_i *and* $M^i{}_j$ *are simply numbers. However: even though* $w_i M^i{}_j = M^i{}_j w_i$, *it is completely incorrect to say* $\widetilde{\mathbf{w}} M = M \widetilde{\mathbf{w}}$. *This is because* $\widetilde{\mathbf{w}}$ *and* M *are tensorial (vector-y) objects. The order of their 'multiplication' matters. You can see this from the matrix notation.*

$$\widetilde{\mathbf{w}} M = \begin{pmatrix} 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \qquad M \widetilde{\mathbf{w}} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 4 & 5 & 6 \end{pmatrix} \ . \tag{4.7}$$

The first multiplication gives a row vector, as you expect since $(wM)_j$ *has one lower index. The second multiplication does not even make sense. What we see is that expressions like* $w_i M^i{}_j = M^i{}_j w_i$ *are valid as long as you are only talking about the components. The glib "physicist slang" of replacing a component by its vector/matrix/tensor can get you into trouble if you have moved components around in a way that is only allowed for numbers, but not vectory-things.*

Since the language is now becoming cumbersome, let us define the word **tensorial** to mean an object with indices. This will replace the phrase "vectory" in our notes.

5 Ket Notation

Quantum mechanics has a completely different notation for vectors. Rather than $\mathbf{v}$, we write $|v\rangle$. This is called a *ket*. The reason for this is that we write row vectors as $\widetilde{\mathbf{w}} = \langle w|$. We call this object a *bra*. Then the "matrix multiplication" $\widetilde{\mathbf{w}}\mathbf{v}$ from (4.2) is succinctly written $\langle w|v\rangle$. The pun is that this is a 'bra–ket' or a *bracket*. At the moment we do not have much use for bra and ket notation, but it is useful to establish this notation early on. It will soon be *very* convenient. To summarize:

$$\text{column vector:} \quad \mathbf{v} = |v\rangle \qquad\qquad \text{row vector:} \quad \widetilde{\mathbf{w}} = \langle w| \tag{5.1}$$

There is no special notation for matrices, M. If you really want to be fancy, though, you can give matrices a hat and write them $\hat{M}$. This is really only important if you need to distinguish between a matrix and a number that happens to have a similar variable name.

Example 5.1. *As an example where the 'hat' notation comes in handy, recall the Schrödinger equation in (2.1):*

$$\hat{H}\,|\Psi\rangle = E\,|\Psi\rangle\ . \tag{5.2}$$

The $\hat{H}$ has a hat because it is a matrix. The E does not have a hat because it is a number. Does this information give the equation more significance? Obviously $\hat{H} \neq E$ since these are two completely different classes of objects. The equation is an equality between two vectors, $|\Psi\rangle$. Evidently, when the matrix $\hat{H}$ hits $|\Psi\rangle$, you end up with a new vector that is simply a rescaling of $|\Psi\rangle$. We say that $|\Psi\rangle$ is an eigenvector *of $\hat{H}$.*

6 Matrices and Linear Transformations

Vectors are the 'nouns' in linear algebra. The word 'linear' refers to the the *verbs*. That is: we would like to act on vectors.

6.1 Jargon

Let us introduce some jargon here. Rather than formal definitions, we give practical "physicist's" definitions.[15] You can look up proper definitions in your favorite mathematics textbook. The following words are closely related: function, map, transformation. I often use them interchangeably, though this is rather sloppy.

A **function** is a mathematical machine that takes some inputs and produces some output. The inputs can be numbers, vectors, or more sophisticated objects. The outputs may also be numbers, vectors, or more sophisticated objects. The outputs do not have to be the same type of object as the inputs—in general they are not. A function that takes one input and returns one output is called a **map**. A map that takes one type of object and returns the same type of object is called a **transformation**. Mathy-folks like to draw diagrams like Fig. 1.

---

[15]If we do something less formally, we say that we are physicists, not mathematicians. If we choose to be more highbrow, then we say that we are physicists, not engineers.
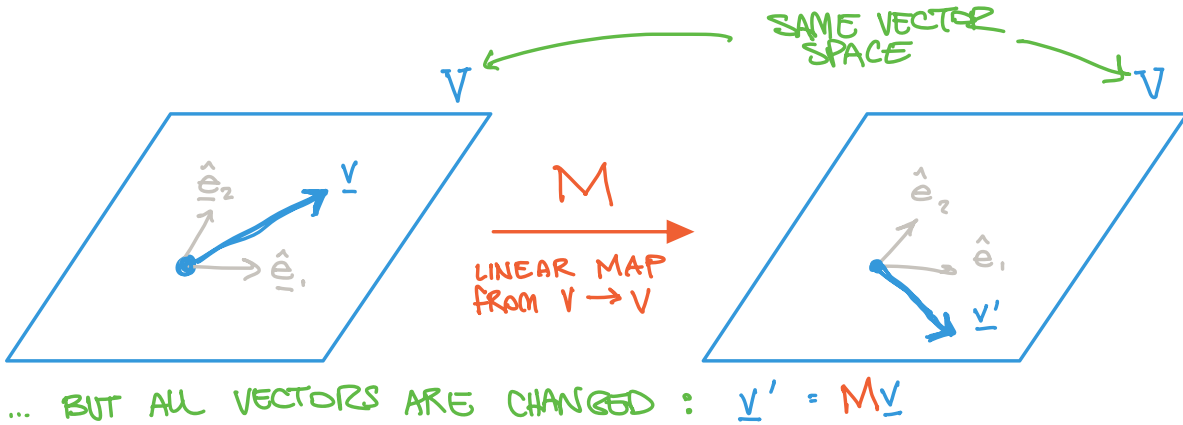
Figure 1: Example of a linear transformation. The transformation M (for matrix) takes a vector $\mathbf{v}$ and turns it into a different vector that we call $\mathbf{v}'$. This new vector is related to $\mathbf{v}$ by the matrix $\mathbf{v}' = M\mathbf{v}$. *Every* vector is transformed under the map M. This is an *active transformation* where all vectors transform, but the basis ($\hat{\mathbf{e}}_i$) stays fixed.

Example 6.1. *The dot product is a **function** that takes in two vectors and outputs a number.*

Example 6.2. *The magnitude is a **map** that takes a vector and returns a number.*

Example 6.3. *The map that takes a vector and returns its unit vector is a **transformation**: it is a function that takes a vector and returns a vector.*

6.2 More Jargon

Pre-image, image, kernel,

6.3 Linear transformations

A function f is **linear** if it satisfies:

$$f(\alpha x) = \alpha f(x) \tag{6.1}$$
$$f(x + y) = f(x) + f(y) \ . \tag{6.2}$$

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) \tag{6.3}$$

Acknowledgments