FOUNDATION OF SCIENTIFIC COMMUNICATION:
SLICE OF PIE

Linear Algebra for Quantum Mechanics

Author: *Remko Osseweijer and Ashhad Shahzad*

Course: *Quantum Mechanics*
Date: *January 9, 2026*

1. Introduction

In this brief paper of lecture notes, we build the necessary mathematical framework to describe the simplest of quantum mechanical systems. This framework is the field of linear algebra. We will begin by formally defining a vector space that inhabits the vectors. We will consider many properties of these spaces, among which subspaces, bases, and dimension. We will then consider linear maps between vector spaces that introduce a sense of moving through them. The linear maps turn out to be equivalent to matrices. We will extract many properties of the matrices, like the transpose and determinant. Finally, we cover the method discovering special vectors and scalars associated to matrices, namely the eigenvectors and eigenvalues.

Beware. This paper is a work in progress and may contain mistakes and misconceptions. Please notify the authors if you find any.

2. Notations

In the following document, we make use of the convention for notation:
Scalars $\lambda, a, b, c, \ldots$
Vector spaces $V, W, \ldots$
Vectors $v, w, \ldots$
Basis vectors e_i
Dual space V^*
Dual vectors $\omega, \xi, \ldots$
Dual basis vectors θ_i
Field $\mathbb{F}$
Matrices $M, N, \ldots$
Adjoint $M^\dagger$
Complex conjugate $\bar{v}$
Linear maps $\mathcal{L}(V, \mathbb{R})$
Hilbert space $\mathcal{H}$

3. Vector spaces

The notion of vectors is one that is familiar in classical physics when more than one dimension is considered. An obvious example of a vector is the Newtonian force on an object, that usually is represented by an arrow-like object. Familiar readers will know that forces can be added and scaled appropriately by adding these arrows tail to end. In this chapter, we seek to formally define vector spaces and extract their multitude of interesting properties.

3.1. Definition

We begin by formally stating the definition of a vector space.

Definition 3.1 (Vector spaces)**.** *Let V be a non-empty mathematical set. Then V is called a* ***vector space over*** $\mathbb{F}$ *(where $\mathbb{F}$ usually stands for $\mathbb{R}$ and $\mathbb{C}$) if the following axioms are satisfied.*

1. *For all $v, w \in V$, we have $v + w = w + v$.* ***(Commutativity)***
2. *For all $v, w, z \in V$, we have $v + (w + z) = (v + w) + z$.* ***(Associativity)***
3. *There exists an element $0 \in V$ such that $v + 0 = v$ for all $v \in V$.* ***(Identity of addition)***
4. *For all $v \in V$, there exists an element $(-v) \in V$ such that $v + (-v) = 0$.* ***(Existence of additive inverse)***
5. *For all $v \in V$ and $\lambda, \mu \in \mathbb{F}$, we have $\lambda(\mu v) = (\lambda\mu)v$.* ***(Compatibility of scalar multiplication)***
6. *There exists an $1 \in \mathbb{F}$ such that $1v = v$ for all $v \in V$.* ***(Identity of multiplication)***
7. *For all $v, w \in V$ and $\lambda \in \mathbb{F}$, we have $\lambda(v + w) = \lambda v + \lambda w$.* ***(Vector distributivity)***
8. *For all $v \in V$ and $\lambda, \mu \in \mathbb{F}$, we have $(\lambda + \mu)v = \lambda v + \mu v$.* ***(Scalar distributivity)***

When $\mathbb{F} = \mathbb{R}$ we call V a **real vector space**, and when $\mathbb{F} = \mathbb{C}$ we call V a **complex vector space**[1]. Elements of V are called **vectors** and elements of $\mathbb{F}$ are usually called **scalars**. These axioms ensure that vectors behave nicely under addition and scalar multiplication. Notice that we now have a definition for the notion of a vector that is significantly different from the physicist's: "a vector is an object that behaves like a vector". Regardless, the reader should be cautious to view vectors as physical objects like arrows in a geometric space. The concept of vectors here is abstract, and sometimes it makes no sense to attach a notion of geometric direction to them.[2]

More often than not, we forgo mentioning the scalar field $\mathbb{F}$ when considering a vector space V and leave it understood to the context. When we say "consider a vector space $V\ldots$", what we mean is "consider a vector space V over a field $\mathbb{F}\ldots$".

It bears fruitful to consider an example in action; arguably, one that is the most important for us.

[1]In fact, one could consider vector spaces over an arbitrary $\mathbb{F}$ as long as $\mathbb{F}$ is a so-called field. Informally, this means that $\mathbb{F}$ behaves nicely under addition, subtraction, multiplication and division. Both $\mathbb{R}$ and $\mathbb{C}$ are examples of fields.

[2]Consider, for example, the RGB vectors MAKiT is fond of showing off. More generally, in general relativity it turns out that we cannot even consider distance or velocity vectors as arrows between two points in spacetime; we must view them as abstract (local) vectors.

Example 3.2 (*Real n-tuples*)**.** Consider the set of the n-tuples of real numbers. Explicitly, what we mean by the n**-tuples of real numbers** is the set

$$\mathbb{R}^n = \left\{ v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \text{ with } v_1, v_2, \ldots, v_n \in \mathbb{R} \right\}.$$

We wish to show that $\mathbb{R}^n$ is a vector space, for which we must demonstrate that all eight axioms are satisfied for $V = \mathbb{R}^n$ and $\mathbb{F} = \mathbb{R}$. In order to proceed, we first need to define addition and scalar multiplication for the n-tuples to be able to check for the axioms. We define addition and scalar multiplication of n-tuples point-wise. For two elements $v, w \in \mathbb{R}^n$, we can write

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \quad \text{and} \quad w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix},$$

for some $v_1, \ldots, v_n, w_1, \ldots, w_n \in \mathbb{R}$. The addition $v + w$ is then defined as the n-tuple,

$$v + w = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix}.$$

Similarly, scalar multiplication is defined for $v \in V$ and $\lambda \in \mathbb{R}$ as

$$\lambda v = \lambda \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \\ \vdots \\ \lambda v_n \end{pmatrix}.$$

Let us now show the proof for the axioms of commutativity and vector distributivity as an example. Let us consider two elements $v, w \in \mathbb{R}^n$,

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \quad \text{and} \quad w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix},$$

for $v_1, \ldots, v_n, w_1, \ldots, w_n \in \mathbb{R}$. For two real numbers $a, b \in \mathbb{R}$, we know that they commute, $a + b = b + a$, thus

$$v + w = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix} = \begin{pmatrix} w_1 + v_1 \\ w_2 + v_2 \\ \vdots \\ w_n + v_n \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = w + v.$$

The commutativity of n-tuples directly follows from the commutativity of $\mathbb{R}$. To show the distributivity of vectors, consider additionally an $\lambda \in \mathbb{R}$. Then,

$$\lambda(v+w) = \lambda\left(\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}\right) = \lambda \begin{pmatrix} v_1+w_1 \\ v_2+w_2 \\ \vdots \\ v_n+w_n \end{pmatrix} = \begin{pmatrix} \lambda(v_1+w_1) \\ \lambda(v_2+w_2) \\ \vdots \\ \lambda(v_n+w_n) \end{pmatrix}$$
$$= \begin{pmatrix} \lambda v_1+\lambda w_1 \\ \lambda v_2+\lambda w_2 \\ \vdots \\ \lambda v_n+\lambda w_n \end{pmatrix} = \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \\ \cdots \\ \lambda v_n \end{pmatrix} + \begin{pmatrix} \lambda w_1 \\ \lambda w_2 \\ \cdots \\ \lambda w_n \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \lambda \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \lambda v + \lambda w.$$

Here we use that for $a, b, \lambda \in \mathbb{R}$, we have $\lambda(a+b) = \lambda a + \lambda b$. Once again, vector distributivity follows from distributivity of elements in $\mathbb{R}$. ∇

Exercise 3.3. Complete the previous example and show that the remaining six vector space axioms are also satisfied for $V = \mathbb{R}^n$ and $\mathbb{F} = \mathbb{R}$, proving that $\mathbb{R}^n$ is a vector space over $\mathbb{R}$. Make sure that every step is justified by using the defined addition and multiplication and using the assumed properties of $\mathbb{R}$. △

A vector in $\mathbb{R}^n$ is thus a n-tuple of length n with real numbers as entries, where the scalars are single real numbers.

This is, of course, the usual notion of a vector we are familiar with. The reader should make sure that this example is clear in the new machinery, as we will refer to this example in almost all further developments on vector spaces.

Having set the stage, there is an immediate generalisation.

Example 3.4. Consider the set of n-tuples of complex numbers,

$$\mathbb{C}^n = \left\{ z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} \text{ with } z_1, z_2, \ldots, z_n \in \mathbb{C} \right\},$$

with addition and scalar multiplication defined piecewise. One can show that $\mathbb{C}^n$ is a vector space over $\mathbb{C}$ in precisely the same way we showed that $\mathbb{R}^n$ is a vector space over $\mathbb{R}$. ∇

Exercise 3.5. If you need more practise with the axioms, show that $\mathbb{C}^n$ with piecewise addition and multiplication is a vector space over $\mathbb{C}$. △

Exercise 3.6. If you need more practise with the axioms but the preceding example was too easy, consider for a nonempty fixed set S the set of functions ,

$$\mathbb{F}^S = \{f : S \to \mathbb{F}\},$$

i.e. an $f \in \mathbb{F}^S$ is a function $f : S \to \mathbb{F}$ that takes elements $x \in S$ to $f(x) \in \mathbb{F}$. For all $f, g \in \mathbb{F}^S$ and $\lambda \in \mathbb{F}$, define the addition $f + g \in \mathbb{F}^S$ and scalar multiplication $\lambda f \in \mathbb{F}^S$ as,

$$(f+g)(x) = f(x) + g(x) \text{ and } (\lambda f)(x) = \lambda(f(x)),$$

for all $x \in S$. Show that $\mathbb{F}^S$ forms a vector space over $\mathbb{F}$. △

From now on, V is always a vector space over a field $\mathbb{F}$ unless specifically mentioned otherwise.

3.2. Subspaces

Often, we are dealing with smaller spaces within vector spaces that we desire to behave nicely. Preferably, we would like the smaller space to act as a vector space. For this, we write the following definition.

Definition 3.7. *Consider a subset $W \subset V$. Then W is a* ***subspace*** *of V is the following conditions are satisfied.*

1. *For all $v, w \in W$, we have $v + w \in W$. (**Closed under addition**)*
2. *For all $v \in W$ and $\lambda \in \mathbb{F}$, we have $\lambda v \in W$. (**Closed under scalar multiplication**)*
3. *We have $0 \in W$. (**Identity of addition**)*

Since all elements of W are inside V, the subspace W induces the axioms from the vector space V. Together with the three conditions for subspaces, this ensures that a subspace W itself is a vector space.

Exercise 3.8. Prove that a subspace is a vector space by accounting for all the vector space axioms.

3.3. Linear independence and bases

The vector addition, natural as it may be, provides us with elegant ways to manipulate vectors. First of all, we can add and scale several vectors to obtain new vectors.

Definition 3.9 (Linear combination)**.** *Suppose we have the vectors $v_1, \ldots, v_m \in V$ and the scalars $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$. Then the vector $w \in V$ defined by*

$$w = \lambda_1 v_1 + \cdots + \lambda_m v_m, \tag{1}$$

is called a ***linear combination*** *of $v_1, \ldots, v_m$.*

The set of all linear combinations of some vectors $v_1, \ldots, v_m \in V$, called the span and denoted by $\text{span}(v_1, \ldots, v_m)$, turns out to be a subspace of V. We can explicitly write the span as,

$$\text{span}(v_1, \ldots, v_m) = \{\lambda_1 v_1 + \cdots + \lambda_m v_m : \lambda_1, \ldots \lambda_m \in \mathbb{F}\}.$$

Exercise 3.10. For some fixed vectors $v_1, \ldots, v_m \in V$, prove that $\text{span}(v_1, \ldots, v_m)$ is a subspace of V. *Hint: Use the definition of a subspace.*

Definition 3.11. *Let $v_1, \ldots, v_m \in V$ be fixed vectors. If $\text{span}(v_1, \ldots, v_m) = V$, then $v_1, \ldots, v_m$ is said to* ***span*** *V.*

Now, consider we have vectors $(v_1, ..., v_n)$, then they cover the subspace $\text{span}(v_1, ..., v_n)$. Let $w = \lambda_1 v_1 + ... + \lambda_n v_n$ be some linear combination of our vectors, then if we consider $\text{span}(v_1, ..., v_n, w)$ it is the exact same as the subspace $\text{span}(v_1, ..., v_n)$. The vector w is then redundant and we can drop it from consideration. In general, if any vector v_i is a linear combination of $v_1, ..., \hat{v}_i, ..., v_n$ then $\text{span}(v_1, ..., v_n)$ is the same as $\text{span}(v_1, ..., \hat{v}_i, ..., v_n)$. This motivates the next definition.

Definition 3.12. *Let $v_1, \ldots, v_m \in V$ be vectors. Then $v_1, \ldots, v_m$ are said to be **linearly dependent** if there exist $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$ with at least one nonzero, such that*

$$\lambda_1 v_1 + \cdots + \lambda_m v_m = 0.$$

*We say that $v_1, \ldots, v_m$ are **linearly independent** if all $\lambda_1, \ldots, \lambda_m$ must be zero to satisfy the above equation.*

Note that if two such scalars are nonzero, say λ_i and λ_j, then we can express v_i in terms of v_j, i.e. v_i depends on v_j. Only when all $\lambda_1, \ldots, \lambda_m$ are zero can we truly not express any vector from $v_1, \ldots, v_m$ in terms of the others.

We have come to the climax of the material.

Definition 3.13 (Bases)**.** *Suppose that $v_1, \ldots, v_m \in V$ that are linearly independent and* $\text{span}(v_1, \ldots, v_m) = V$*. Then we call $\{v_1, \ldots, v_m\}$ a **basis** of V.*

Clearly, this basis would not be unique; we could simply replace v_1 with λv_1 for any nonzero and non-one $\lambda \in \mathbb{F}$ and obtain a different basis for V.

Definition 3.14. *Any vector space V that can be spanned by a finite number of vectors $v_1, \ldots, v_m \in V$ is called a **finite-dimensional vector space**. A vector space that is not finite-dimensional is called an **infinite-dimensional vector space**.*

Immediately , we will for now assume our vector space V is finite-dimensional.

[bla bla]

Lemma 3.15. Two different bases of a vector space have the same number of elements

Definition 3.16. *The number of vectors in a basis for a vector space V is called the **dimension of** V, denoted by* $\dim V$.

Lemma 3.17. Every finite-dimensional vector space has a basis.

Exercise 3.18. Prove the previous lemma. *Hint: Think on how the vectors that span the vector space can be made linearly independent.*

After all those definitions and statements, the time must be ripe to treat an example in greater detail. Naturally, we come back to our vector space $\mathbb{R}^n$ from which we know our vectors as column arrays.

Example 3.19. Basis for $\mathbb{R}^n$.

4. Linear functions

4.1. Linear Maps

Suppose we have two arbitrary vector spaces V and W. We would like to set up a relationship between the elements of V and W however we cannot do this arbitrarily. The reason why vector spaces are so useful is because of their linear structure, and so the maps between vector spaces must preserve that structure. We then axiomatize the following properties of a linear map $T : V \to W$.

$$T(v_1 + v_2) = T(v_1) + T(v_2)$$
$$T(\lambda v_1) = \lambda T(v_1)$$
$$T(0_V) = 0_W$$

The last property asserts that the additive identity in V map to the additive identity in W. It is derivable from the other axioms but we state it here since it is a useful property to consciously have in mind. The set of all linear maps from V to W is denoted $\mathcal{L}(V, W)$. We write the space of linear maps from V to itself as simply $\mathcal{L}(V)$. It turns out we can also endow $\mathcal{L}(V, W)$ with a vector space structure, note we already have

$$\lambda T(v_1) = T(\lambda v_1)$$

If we assert now that for any two maps $T, S \in \mathcal{L}(V, W)$ we have

$$(T + S)(v_1) = T(v_1) + S(v_1)$$

Then it turns out that $\mathcal{L}(V, W)$ is indeed a vector space.

Exercise 4.1. Verify that $\mathcal{L}(V, W)$ satisfies the axioms of a vector space. (Hint: It may be useful to fix an arbitrary vector for the linear maps to act on).

At first glance it not be obvious how such maps are useful, so below are a few examples

Example 4.2. Consider the ring of all polynomials in the variable x over $\mathbb{R}$ denoted by $\mathbb{R}[x]$. Differentiation is then a linear map $\mathcal{L}(\mathbb{R}[x])$ since for any polynomials $p(x), q(x)$ and real numbers a, b

$$\frac{d}{dx}\left(ap(x) + bq(x)\right) = a\frac{d}{dx}p(x) + b\frac{d}{dx}q(x)$$

Exercise 4.3. Prove that definite integration as a map in $\mathcal{L}(\mathbb{R}[x], \mathbb{R})$ defined as

$$p(x) \mapsto \int_a^b p(x)$$

is a linear map. Moreover, prove that multiplication by a polynomial, such as x^2, is a linear map in $\mathcal{L}(\mathbb{R}[x])$.

Now suppose we have a Linear map $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$. The composition of these two maps is then given by their successive application

$$ST(u) = S(T(u))$$

It is important to keep in mind that compositions such as above only make sense when the range of the first operator (here T) is a subset of the domain of the second operator (here S). The key properties of the composition of linear maps are

1. For linear maps T_1, T_2, T_3, whenever the composition

$$T_1T_2T_3(u)$$

is well defined (in the sense of agreement in the range and domain of successive operators) the composition is **associative**, i.e.

$$(T_1T_2)T_3(u) = T_1(T_2T_3(u))$$

2. For a linear operator $\mathcal{L}(V, W)$ we have

$$I_W T(v) = T I_V(v) = T(v)$$

Where I_W and I_V are the identity operators on W and V

3. For linear operators S and T_1, T_2, if the compositions ST_1 and ST_2 are well defined then composition is **distributive**, i.e.

$$S(T_1 + T_2)(v) = ST_1(v) + ST_2(v)$$

It is important to note that composition of linear maps is not in general commutative, i.e $ST \neq TS$

Exercise 4.4. Let differentiation in the variable x and multiplication by x^3 be two linear maps $\mathcal{L}(\mathbb{R}[x])$. Show that these two maps do note commute.

4.2. Linear functionals

A particularly interesting family of linear maps is one that goes from V to the field $\mathbb{F}$, so $\mathcal{L}(V, \mathbb{F})$. These maps are referred to as functionals on V. As we showed earlier, the space $\mathcal{L}(V, \mathbb{F})$ has a vector space structure, what is interesting however, is that the vector space has the same dimensionality as V.

We start by noting that if V has a basis $\{e_i\}$ then a linear functional $T \in \mathcal{L}(V, \mathbb{F})$ acts on elements of V as follows

$$\begin{aligned} T(v) &= T\left(\sum_i \lambda_i e_i\right) \\ &= \sum_i \lambda_i T(e_i) \end{aligned}$$

Now, the linear functional T can be determined uniquely by its action on the basis $\{e_i\}$, since suppose there was a functional S such that $S(e_i) = T(e_i)$ for all i, then the linear functionals would agree on the decomposition of any vector v into the linear combination of the basis, and thus v itself. The evaluation $T(e_i)$ for all i is a scalar of the field $\mathbb{F}$, and so for n basis vectors one would check the action of T on each basis, yielding n scalars that are uniquely associated with T, and so T would be identified with element of $\mathbb{F}^n$.

To rephrase the isomorphism in more concrete terms, it is clear that every functional T has a representation in $\mathbb{F}^n$, when a basis on V is fixed, conversely for any element $f \in \mathbb{F}^n$ one can construct a functional such that $T(e_i) = f_i$. Lastly the zero functional corresponds to exactly the zero vector.

4.3. Matrices

We have covered the general concept of linear maps between vector spaces $\mathcal{L}(V, W)$ as well as a representation of $\mathcal{L}(V, \mathbb{F})$ in terms of vectors of $\mathbb{F}^n$. Since $\mathcal{L}(V, W)$ too is a vector space, a natural question is how we can represent it concretely.

Recall that in the case of linear functionals, the action of T on any basis e_i was a scalar. For a general linear map, the action of T on a basis e_i will result in some vector w_i in the target space. A general linear map too can be determined uniquely by its action on the basis for the same reason as linear functionals can.

Since the action on each basis results in a vector $w_i \in W$, the total action on n basis vectors can be represented as a vector in W^n. If W is an m-dimensional vector space then one can rewrite $W^n = \mathbb{F}^{mn}$. And so we see that the space $\mathcal{L}(V, W)$ has dimensionality $\dim(V) \times \dim(W)$.

Now we move on to write out a representation of the linear map as an element of $\mathbb{F}^{mn}$. Recall the action of T on a basis was given by

$$\begin{aligned} T(e_i) &= w_i \\ &= \sum_j \lambda_{ji} f_j \end{aligned}$$

Where we have expressed the vector w_i in terms of the basis $\{f_j\}$ of the vector space W. Now although T is an element of $\mathbb{F}^{mn}$, it would be cumbersome and non-enlightening to write out a single row or column with mn entries. Instead, a better way is to write the action of T on the ith basis as one column and similarly for the other basis vectors and write everything out as a grid of numbers

$$T = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1n} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{m1} & \lambda_{m2} & \dots & \lambda_{mn} \end{bmatrix}$$

And algebraically, the action of T on an arbitrary vector v can be written as

$$\begin{aligned} T(v) &= T\left(\sum_i \gamma_i e_i\right) \\ &= \sum_i \gamma_i T(e_i) \\ &= \sum_i \gamma_i w_i \\ &= \sum_i \gamma_i \sum_j \lambda_{ji} f_j \\ &= \sum_{i,j} \gamma_i \lambda_{ji} f_j \end{aligned}$$

Exercise 4.5. Let A be an m-by-n matrix and B an n-by-l matrix. Prove that the ik-th entry of AB is given by

$$\sum_{j=1}^{n} A_{ij} B_{jk}$$

4.4. Operations are properties of matrices

In this section we review some properties of a matrix that are useful in various physics as well as computational contexts.

The first operation we'll go over is the transpose. To motivate this operation, consider a map $T : V \to W$ between vector spaces. Then it turns out that there is a canonical map $T^* : W^* \to V^*$ defined via

$$T^*(\phi) := \phi \circ T$$

Where $\phi \in W^*$. What this essentially means is that given a map $T : V \to W$, we can take our vectors from V to W and operate on them using the linear functionals in W^*, and in doing so, we have a well defined notion on how the linear functionals W^* act on vectors in V, and so given the map T we can define T^* as a map between the dual spaces called the pullback of T.

To see where the transpose comes in, consider $\{e_i\}$ and $\{f_i\}$ to be bases of V and W respectively, and $\{e^i\}$, $\{f^i\}$ for the dual spaces. Now applying T^* to f^i and evaluating it on e_j we get

$$\begin{aligned}\left(T^*(f^i)\right)(e_j) &= (f^i \circ T)(e_j) \\ &= f^i(T(e_j)) \\ &= f^i\left(\sum_k \lambda_{kj} f_k\right) \\ &= \sum_k \lambda_{kj}\delta^i_k \\ &= \lambda_{ij}\end{aligned}$$

This was one element, to get the rest of the row corresponding to f^i we operate with all basis vectors e_j and obtain

$$T^*(f^i) = \sum_j \lambda_{ij} e^j \tag{2}$$

But note that the action of T on e_i was given by

$$T(e_i) = \sum_j \lambda_{ji} f_j$$

We see that the indices denoting the rows and columns are swapped, in particular, it turns out the rows and columns of T are interchanged to give T^*, and this is exactly what the operation of transpose corresponds to

$$T^T = T^*$$

It is also important to note that while the pullback T^* is canonical, the transpose T^T is not since it depends on the coordinate representation of V and W.

Next, we turn attention to the determinant of a matrix. Such an operation is only well defined for square matrices, or more accurately, for maps of the form $T : \mathbb{R}^n \to \mathbb{R}^n$. Geometrically, given the hyper-volume of a unit n-cube in the domain of T, the determinant of T describes how the hyper-volume scales during the transformation. A determinant of 2 implies that the unit n-cube has now twice the volume as before, vice versa for a determinant of $\frac{1}{2}$. The determinant also encodes the orientation of the transformation, in particular the negative sign indicates that the orientation has flipped.

The determinant of a 2-by-2 matrix is given by

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

The determinant of a diagonal n-by-n matrix is given by

$$\begin{vmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_m \end{vmatrix} = \prod_i \lambda_i$$

In general, the determinant is calculated using some cofactor BS that currently the author does not have the energy to write out.

5. Eigenvectors and eigenvalues

From now on, we will be exclusively concerned with linear maps between the same vector space V with $\dim V = n$. These maps, as we have seen, can be represented by $n \times n$ dimensional matrices. Given such a matrix $A \in M(n, \mathbb{C})$, we could be interested in a subspace $W \subset V$ of V that is invariant under the matrix. In other words,the vectors $v \in W$ remain in W when they transform under A, so $Av \in W$ for all $v \in W$. We say that W is an **invariant subspace** of V under A. This would mean that the entire subspace is isolated from the rest of the vector space.

It is this intriguing property that motives us to define the following properties of matrices. We can consider a one-dimensional invariant subspace $W = \text{span}(v) \subset V$ for $v \in V$ such that $Av \in W$. But W is the span of v and contains only scalar multiples of v. In other words, the expression $Av \in W$ is equivalent to saying $Av = \lambda v$ for some $\lambda \in \mathbb{F}$. Conversely, if we have some vector $v \in V$ such that $Av = \lambda v$, then we can immediately conclude that the subspace $\text{span}(v)$ is invariant under A.

This relation is so important that we gave it a name.

Definition 5.1. *Let A be an $n \times n$ matrix. Then $\lambda \in \mathbb{F}$ is called an **eigenvalue** if $Av = \lambda v$ for some nonzero $v \in V$. The associated vector v is called an **eigenvector corresponding to** λ.*

Thus, a matrix A has an eigenvalue λ if and only if it has a one-dimensional invariant subspace. Moreover, we see that eigenvectors are only scaled upon transformation; they do not change direction!

We now want to formulate a method to find the eigenvalues of a matrix. We can rewrite the expression into the form

$$(A - \lambda I)v = 0.$$

6. Quantum Mechanics

In the previous chapters, we have exploratively laid down the framework of the theory of linear algebra. We have seen how vector spaces are defined and how linear functions behave as carriers between these vector spaces. We have seen that these linear functions naturally provoke the notion of a matrix, which we have studied extensively. We declared many operations on matrices, such as the determinant and the transpose, and we found many properties, such as invertibility, unitarity and hermicity. We developed a framework to find the so-called eigenvectors and eigenvalues of matrices. At last, we found that we could define a canonical dual space to a vector space that hosts the linear functions acting on the vectors. Coincidentally, these dual spaces have remarkable properties and can even be used to define inner products.

It is now time to apply this grand theory to quantum mechanics.

6.1. The Postulate of Quantum

We begin with a postulate that is the backbone of quantum mechanics.

Postulate 6.1. *Every quantum system has an associated complex vector space $\mathcal{H}$.*

For our purposes, we will assume that this vector space $\mathcal{H}$ is also a Hilbert space, which means that $\mathcal{H}$ has an associated inner product. We will denote the vectors of the Hilbert space as $|\psi\rangle \in \mathcal{H}$, and we often refer to these vectors as **kets**, following Dirac.

The nature of $\mathcal{H}$ differs for different quantum systems. For example, when we consider the spin of a single electron, the crucial vectors are the spin up ket $|\uparrow\rangle$ and the spin down ket $|\downarrow\rangle$ (equivalent to the qubit states $|0\rangle$ and $|1\rangle$). The Hilbert space is thus two-dimensional, with a basis given by $\{|\uparrow\rangle, |\downarrow\rangle\}$ and an arbitrary ket in the Hilbert space is given by $|\psi\rangle = z|\uparrow\rangle + w|\downarrow\rangle$ for $z, w \in \mathbb{C}$.[3]

But if we wish to describe the x position of the electron instead, then we find that a finite-dimensional Hilbert space will not suffice. Any position on the x-axis should be represented by a vector, and the last time the authors checked, the reals do not appear to be discrete. Infinite-dimensional Hilbert spaces require extraordinary care and precision, lest one drown in the subtleties and nuances. We will try to avoid them as such.

We must be a bit careful when we discuss the Hilbert space vectors. Often, we want to think of vectors like arrays, just as we do for vectors in $\mathbb{R}^n$. For the spin of the electron, we can, for example, represent the kets as,

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

From this, we see that we could describe the Hilbert space $\mathcal{H}$ with $\mathbb{C}^2$.

But more often than not, we cannot represent a ket using an array. The example immediately fails when we consider a system with more than one electron, say two, where the ket is now described as a tensor product between the spin directions of either electron. If both electrons were to be in an up spin, the ket would be given by $|\uparrow\rangle \otimes |\uparrow\rangle$,

[3] To be able to talk about physical states, we require one more condition. Namely that two vectors different only by a scalar multiplication represent the same state. Thus, a ket/vector is an element in the Hilbert space, whereas a physical state is a line in the Hilbert space.

whereas if the first electron has spin up and the second electron has spin down, then the ket is given by $|\uparrow\rangle \otimes |\downarrow\rangle$. In other examples, the states are instead represented by matrices or even spinors. This is why it is important for us to consider the vectors as abstract elements of the vector space and not as arrays of numbers.[4]

6.2. Quantum Mathematics

In the previous section, we have given a brief introduction based on a defining postulate. Here we recap and expand on the mathematical set-up of quantum mechanics.

We begin with a complex Hilbert space $\mathcal{H}$ associated to some physical system. The vectors in this Hilbert space are written as kets $|\psi\rangle \in \mathcal{H}$. Physical states are rays in the Hilbert space, i.e. both $|\psi\rangle, |\psi\rangle' \in \mathcal{H}$ with

$$|\psi\rangle = z\,|\psi'\rangle$$

for some $z \in \mathbb{C}$, represent the same physical state.[5] This gives us the freedom to choose a vector on the ray to represent the physical state. We often choose a normalised vector, because this will make the manipulations a lot simpler. Normalisation can be defined using the inner product $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{F}$ of the Hilbert space $\mathcal{H}$. For a state $|\psi\rangle \in \mathcal{H}$, the inner product $\langle |\psi\rangle, |\psi\rangle \rangle$ is a real number representing the squared length of the vector $|\psi\rangle$. Thus we define the **normalised** vector as

$$|\psi'\rangle = \frac{|\psi\rangle}{\sqrt{\langle |\psi\rangle, |\psi\rangle \rangle}}.$$

It follows that $\langle |\psi'\rangle, |\psi'\rangle \rangle = 1$, which is precisely what we want.

We can consider the linear functionals of the Hilbert space $\mathcal{H}$, which is, of course, the dual space $\mathcal{H}^*$. The elements of the dual space, i.e., functions that take vectors to $\mathbb{C}$, are denoted in the Dirac notation as bras $\langle \phi| \in \mathcal{H}^*$. The functional $\langle \phi|$ acting on a vector $|\psi\rangle$ is then denoted as $\langle \phi| (|\psi\rangle) = \langle \phi|\psi\rangle \in \mathbb{C}$. This is a different way to write the inner product.[6]

We can take a vector $|\psi\rangle \in \mathcal{H}$ to a dual vector $\langle \psi| \in \mathcal{H}^*$ using the adjoint, $(|\psi\rangle)^\dagger = \langle \psi|$, and inversely, $(\langle \psi|)^\dagger = |\psi\rangle$. We can thus write the normalisation condition of a vector $\psi \in \mathcal{H}$ as $\langle \psi|\psi\rangle = 1$.

Linear functionals of kets are bras, but it turns out that linear functions $A : \mathcal{H} \to \mathcal{H}$ also serve an important role in quantum mechanics. These linear functions are called **operators**.[7]

Though the quantities $|\phi\rangle\,|\psi\rangle$ and $\langle \phi|\,\langle \psi|$ do not make sense (a vector-vector multiplication is not defined, nor is a function-function operation), the quantity $|\phi\rangle\,\langle \psi|$ does make sense! This is an object that, when acted upon by a ket $|\varphi\rangle$, returns another vector $|\phi\rangle\,\langle \psi|\rangle\, \varphi = (\langle \psi|\varphi\rangle)\,|\phi\rangle$. Thus $|\phi\rangle\,\langle \psi| : \mathcal{H} \to \mathcal{H}$ is an example of an operator A[8].

As vectors go, we can expand any vector into basis vectors.[9]

[4] In order to write arbitrary vectors in the Hilbert space as arrays or matrices, we formally need to seek aid in representation theory.

[5] This is, of course, an equivalence relation where the equivalence classes are one-dimensional subspaces of the Hilbert space $\mathcal{H}$. This is akin to considering the projective Hilbert space $\mathbb{P}\mathcal{H}$.

[6] This is actually a definition: $\langle |\phi\rangle, |\psi\rangle \rangle \equiv (|\phi\rangle)^\dagger\,|\psi\rangle = \langle \phi|\psi\rangle$.

[7] There also exist antilinear functions, which act like $A(cv) = \bar{c}A(v)$. These are not relevant to our discussion, but their existence should not go unnoticed. The time reversal operator is such an antilinear function.

[8] This is not a double dagger, but a footnote.

[9] This footnote is merely because the author is curious on how the footnote symbol looks like.

6.3. Hamiltonian

6.4. Position and momentum

Patch notes

[Rem] January 1st at 2pm: Defined vector spaces and vectors in the formal (boring) mathematical way with some fun examples and challenging exercises.

[Rem] January 2nd at 1am: Gave the vector spaces a basis; they should be happy now.

[Rem] January 2nd at 4pm: We have unlocked dimensions now!

[Ash] January 2nd at 11:59pm: Wrote about maps between linear spaces, functionals, matrices and their representations

[Rem] January 2nd at 11pm: Quantum mechanics is a fact now. We would have been the first to discover it if no one else was before us.