Quantum Mechanics - Linear Algebra Review

Kevin Langhoff, (With edits by Irian D'Andrea)

September 5, 2023

Abstract

These are notes reviewing relevant topics of linear algebra for Physics 137A. Throughout, I have tried to stick to "bra-ket" notation to acclimate students to this notation. I have also excluded many proofs when I believe the time and effort involved in discussing them isn't worth the lessons they teach. The proofs I do include are almost always worth going through to understand how to use bra-ket notation effectively. Some of the proofs, for example those discussing hermitian operators, are proofs all quantum mechanics students should know by heart.

Contents

1 Introduction

Fluency in linear algebra is essential for an understanding of quantum mechanics. To see why this is, we can take a look at the postulates of quantum mechanics

1. **States:** The state of a quantum system is given by a vector in a Hilbert space. We denote such states using "bra-ket" notation as $|\psi\rangle$.

2. **Observables:** Every observable quantity associated with a quantum state is related to a hermitian operator.

3. **Measurements:** The result of measurement of an observable $\hat{\Omega}$ is one of the eigenvalues of $\hat{\Omega}$. The probability of measuring an eigenvalue ω is related to the inner product of the state being measured, $|\psi\rangle$, and the eigenvector, $|\omega\rangle$, with eigenvalue ω. $P(\omega) = |\langle\omega|\psi\rangle|^2$.

4. **Collapse:** After measurement, the wave function collapses to the eigenvector associated with the eigenvalue measured.

5. **Evolution:** The state of a quantum system evolves by application of a unitary operator $U(t_f, t_i)$ as $|\psi(t_f)\rangle = \hat{U}(t_f, t_i)\,|\psi(t_i)\rangle$. Infinitesimally this can be written as

$$i\hbar\frac{d}{dt}\,|\psi(t)\rangle = \hat{H}(t)\,|\psi(t)\rangle$$

As you can see, there is a boatload of linear algebra terminology that we have to become familiar with. There is one further difficulty. Often we will be interested in the position of a particle is space. The state of the system will then be described by a wave-function $\psi(x)$. At first, this may not seem to be a *vector*, but in fact it is... just in an infinite dimensional vector space. Since this doesn't correspond to our familiar notion of a vector being an object with magnitude and direction, we first have to generalize our understanding of what a vector really is.

2 Vector Spaces

2.1 Basic Definitions

Quantum mechanics is our motivation for studying linear algebra, so we will use the standard notation of quantum mechanics for linear agebraic concepts. The standard quantum mechanical notation for a vector in a vector space is the following:

$$|\psi\rangle \,.$$

ψ is a label for the vector (any label is valid, although we prefer to use simple labels that give contextual clues to the meaning of the vector). The notation $|\cdot\rangle$ is used to indicate that the object is a vector, so it fills the same notational role that $\vec{\cdot}$ plays for ordinary 3D vectors.

Here is a pretentious sounding definition of what a vector space is, according to the people in the math department:

Definition (Vector Space)**.** A vector space is a set V, whose elements are called *vectors*, which is equipped with the operations of vector addition and scalar multiplication which satisfy the following axioms:

Let $|v\rangle\,, |u\rangle\,, |w\rangle$ be vectors in V and α, β be numbers ("scalars"). Then:

1. Commutive (+) : $|v\rangle + |u\rangle = |u\rangle + |v\rangle$
2. Associative (+) : $(|v\rangle + |u\rangle) + |w\rangle = |v\rangle + (|u\rangle + |w\rangle)$
3. Identity (+) : $\exists\ |0\rangle \in V$ such that $|0\rangle + |v\rangle = |v\rangle$ for all $v \in V$
4. Negation (+) : $\forall\, |v\rangle \in V\ \exists\ |-v\rangle \in V$ such that $|v\rangle + |-v\rangle = |0\rangle$
5. Associative (·) : $(\alpha\beta)\, |v\rangle = \alpha(\beta\, |v\rangle)$
6. Distributive (·) : $(\alpha + \beta)\, |v\rangle = \alpha\, |v\rangle + \beta\, |v\rangle$
7. Distributive (·) : $\alpha(|v\rangle + |u\rangle) = \alpha\, |v\rangle + \alpha\, |u\rangle$
8. Unity (·) : $\exists 1 \in \mathbb{F}$ such that $1\, |v\rangle = |v\rangle$ for all $|v\rangle \in V$

If the scalars α, β are real numbers, V is called a real vector space. If the numbers α, β are complex, then V is called a complex vector space. It is not customary to call vectors real or complex, only the scalars which might multiply them. In quantum mechanics we are interested in complex vector spaces, so we will assume that is the case from here on out.

There is no need to memorize these. The main thing we would like to take from this is that there are many examples of vector spaces which may not fit our predisposed notion of "arrow like things".

One notational issue I would like to point out is that usually we will use $|0\rangle$ to mean the ground state of a quantum system, which is NOT the zero vector. In fact, I can't think of a single situation in which people do not just use the "number" 0 to mean the zero vector. You can tell whether it's a number or a vector at a glance from context:

$$\begin{aligned} 6 + 3 \cdot (-2) &= 0 \quad \text{(Clearly a number)} \\ |\psi\rangle + |\phi\rangle + |\chi\rangle &= 0 \quad \text{(Clearly the zero vector)} \end{aligned}$$

Examples.

- The ordinary vectors $\vec{v}$ in 3D space form a vector space. It is a real vector space.
- All 2×2 matrices with elements in $\mathbb{R}$ form a real vector space with the operations

$$\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} + \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & d_1 + d_2 \end{pmatrix} \quad \text{and} \quad r \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ra & rb \\ rc & rd \end{pmatrix}$$

- The set of functions $f : \mathbb{R} \to \mathbb{R}$ is a real vector space where addition and scalar multiplication are defined point-wise as

$$\begin{aligned} (f_1 + f_2)(x) &= f_1(x) + f_2(x) \\ (c \cdot f)(x) &= cf(x) \end{aligned}$$

- The set of functions $f : [0, 1] \to \mathbb{R}$ with $f(0) = f(1) = 0$ is a vector space where addition and scalar multiplication are defined point-wise as above.
- The set of functions $f : [0, 1] \to \mathbb{R}$ with $f(0) = 0$ and $f(1) = 1$ is NOT a vector space. Why?
- The set of polynomials of degree 0 through 2 form a vector space.
- Does the set of 2×2 matrices of the form

$$|z\rangle \equiv \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix}$$

  for a complex number z, with vector addition given by matrix multiplication

$$|z\rangle + |z'\rangle \equiv \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & z' \\ 0 & 1 \end{pmatrix}$$

  and scalar multiplication given by

$$\alpha\, |z\rangle \equiv \begin{pmatrix} 1 & \alpha z \\ 0 & 1 \end{pmatrix}$$

  form a vector space? (The above equations probably look odd, but remember we're free to pick the rules for addition and multiplication however we want, so long as they work).

The vector spaces you should keep in the back of your mind throughout the following discussion are:

- The set of n-tuples of complex numbers $(z_1, ..., z_n)^T$ arranged as column vectors. The rule for addition of two vectors is:

$$\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} + \begin{pmatrix} z_1' \\ z_2' \\ \vdots \\ z_n' \end{pmatrix} = \begin{pmatrix} z_1 + z_1' \\ z_2 + z_2' \\ \vdots \\ z_n + z_n' \end{pmatrix}$$

  and the rule for multiplication by a complex number z is

$$z \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} zz_1 \\ zz_2 \\ \vdots \\ zz_n \end{pmatrix}$$

  This is a vector space usually referred to as $\mathbb{C}^n$.

- The set of complex-valued continuous functions $f : \mathbb{R} \to \mathbb{C}$ with addition and scalar multiplication defined pointwise as above.

Sometimes we are interested in subsets of vector spaces which are also vector spaces. This motivates the following definition:

Definition (Vector Subspace)**.** A *vector subspace* S of a vector space V is a subset of V which is a vector space itself.

To show a set S is a vector subspace it is enough to show that S is closed under vector addition and scalar multiplication, i.e.

$$\begin{gathered} |s_1\rangle + |s_2\rangle \in S \text{ for all } |s_1\rangle, |s_2\rangle \in S \\ c \cdot |s\rangle \in S \text{ for all } |s\rangle \in S \text{ and } c \in \mathbb{F} \end{gathered}$$

2.2 Basis

For computational purposes, it is often convenient to decompose a vector using *basis vectors.* The example $\mathbf{v} = v_x\hat{\boldsymbol{x}} + v_y\hat{\boldsymbol{y}} + v_z\hat{\boldsymbol{z}}$ should be familiar. We now formalize this:

Definition (Linear Independence)**.** A set of vectors $\{|v_1\rangle, .., |v_n\rangle\}$ is said to be linearly independent if

$$\sum_{i=1}^{n} c_i |v_i\rangle = 0 \implies c_i = 0 \ \forall i$$

In other words, none of the vectors can be written as a linear combination of the others. If a set of vectors are not linearly independant, they are said to be *linearly dependant*.

Examples. Show that $(1, -1)$, $(1, 2)$, and $(2, 1)$ are linearly dependant.

Definition (Span)**.** The *span* or *spanning set* of a collection of vectors $\{|v_1\rangle, \ldots, |v_n\rangle\}$ is the set of all linear combinations of those vectors. Explicitly, for the case of two vectors, their span is:

$$\text{Span}\{|v\rangle, |u\rangle\} = \{\alpha |v\rangle + \beta |u\rangle : \alpha, \beta \in \mathbb{C}\}$$

Exercise. Show that the spanning set of any collection of vectors is a vector space.

Definition (Dimension)**.** The *dimension* of a vector space is the maximum number of vectors which can be linearly independent. The dimension of vector space V is denoted $\text{Dim}(V)$.

Alternatively, $\text{Dim}(V)$ is the *minimum* number of vectors needed to span the space; in other words, for the span of those vectors to equal V itself.

If you find $\text{Dim}(V) = n$ linearly independent vectors in V, then any vector in V can be written as a linear combination of those vectors. To see that this is true, let the n linearly independent vectors be called $\{|b_1\rangle, \ldots, |b_n\rangle\}$. Now let $|w\rangle$ be some vector. If we assume that the number of $|b_i\rangle$'s is equal to the dimension n of the vector space, then any larger set of vectors, say $\{|b_1\rangle, \ldots, |b_n\rangle, |w\rangle\}$ is no longer linearly independent. Therefore, there exist nonzero complex numbers $c_1, \ldots, c_n, z$ such that $c_1 |b_1\rangle + \ldots + c_n |b_n\rangle + z |w\rangle = 0$. Moving $|w\rangle$ to the other side and dividing by z, we explicitly see that $|w\rangle$ can be written as a linear combination of the $\{|b_i\rangle\}$. This important fact is summarized in the following definition:

Definition (Basis)**.** Any set of $\text{Dim}(V) = n$ linearly independent vectors in V is called a *basis* of V.

Given the basis of V $\{|b_1\rangle, ..., |b_n\rangle\}$, then any $v \in V$ can be written as

$$|v\rangle = \sum_{i=1}^{n} v_i |b_i\rangle$$

where v_i are called *components of* $|v\rangle$. These components are defined uniquely with respect to a given basis. Further, we can perform vector addition component wise as

$$|v\rangle + |u\rangle = \sum_{i=1}^{n} (v_i + u_i) |b_i\rangle$$

and scalar multiplication as

$$c \cdot |v\rangle = \sum_{i=1}^{n} (cv_i) |b_i\rangle$$

Because all the properties of vectors can be encoded in their components, it is common to associate the vector $|v\rangle$ in V to the column vector $(v_1, \ldots, v_n)^T$ living in the vector space of n-tuples of complex numbers, $\mathbb{C}^n$

$$\sum_{i=1}^{n} v_i |b_i\rangle \xrightarrow{\{b\}} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \tag{1}$$

The act of picking a basis $\{|b_1\rangle, \ldots, |b_n\rangle\}$ and then working directly with the column vector of components is referred to as using the *matrix representation* with respect to the basis $\{|b_i\rangle\}$. In this way all finite dimensional vector spaces are in some sense "the same" as $\mathbb{C}^n$, where n is the dimension of the vector space. (The official term is *isomorphic*.)

2.3 Inner Product Spaces

We would like our vector space to have the notion of distance and direction. We therefore equip our vector space with an *inner product*. In this section we focus on vector spaces over the complex numbers.

An *inner product* is a function which takes as input two vectors $|v\rangle$, $|w\rangle$ and produces a complex number as output. For the time being, it will be convenient to write the inner product of $|v\rangle$ and $|w\rangle$ as $(|v\rangle, |w\rangle)$; for notational clarity, $(\cdot, \cdot)$ will be used to start off, and then occasionally thereafter. The usual notation in quantum mechanics for the inner product $(|v\rangle, |w\rangle)$ is $\langle v|w\rangle$, where $|v\rangle$ and $|w\rangle$ are vectors in the vector space, and the notation $\langle v|$ is used to indicate the *dual vector* to the vector $|v\rangle$, which can be thought of a row vector which acts on ("eats") a ket ("column vector") and outputs the complex number $\langle v| (|w\rangle) \equiv \langle v|w\rangle \equiv (|v\rangle, |w\rangle)$.

Definition (Inner Product)**.** Let $|v\rangle, |u\rangle, |w\rangle$ be vectors in V. An *inner product* is a map $(\cdot, \cdot) : V \times V \to \mathbb{C}$ satisfying:

1. $(\cdot, \cdot)$ is linear in the second slot: $(|u\rangle, \alpha |v\rangle + \beta |w\rangle) = \alpha(|u\rangle, |v\rangle) + \beta(|u\rangle, |w\rangle)$
2. Skew symmetry: $(|u\rangle, |w\rangle) = (|w\rangle, |u\rangle)^*$
3. $(|v\rangle, |v\rangle) \geq 0$ (a real number) with equality if and only if $|v\rangle = 0$.

If a vector space is equipped with an inner product it is called an *inner product space*.

Examples.

- The vector space of n-tuples of complex numbers has an inner product given by

$$((y_1, \dots, y_n), (z_1, \dots, z_n)) \equiv \sum_{i=1}^{n} y_i^* z_i = \begin{pmatrix} y_1^* & \dots & y_n^* \end{pmatrix} \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \tag{2}$$

- Let $\psi(x)$ and $\phi(x)$ be complex-valued continuous functions that are normalizable. Then the map defined by

$$(\psi, \phi) \equiv \int_{-\infty}^{\infty} \psi^*(x)\phi(x)dx$$

  turns the vector space into an inner product space.

Exercise.

- Verify that both the maps given above are inner products.
- Why is the map $((y_1, y_2), (z_1, z_2)) = y_1 z_1 + y_2 z_2$ *not* and inner product for the vector space of pairs of complex numbers (z_1, z_2)?
- Show that the inner product is anti-linear in its first slot: $(\alpha |u\rangle + \beta |v\rangle, |w\rangle) = \alpha^*(|u\rangle, |w\rangle) + \beta^*(|v\rangle, |w\rangle)$

Definition (Orthogonal Vectors)**.** Vectors $|v\rangle$ and $|u\rangle$ are *orthogonal* if $\langle v|u\rangle = 0$.

Definition (Inner Product Space Norm)**.** Define the *norm* of a vector $|v\rangle$ by

$$\|v\| \equiv \sqrt{|\langle v|v\rangle|}$$

Definition (Unit vector)**.** A vector $|v\rangle$ is said to be a *unit vector* if $|| \, |v\rangle \, || = 1$. We also say that "$|v\rangle$ *is normalized*". It is conventinet to speak of "normalizing" a vector by dividing it by its norm. We call $|v\rangle / || \, |v\rangle \, ||$ the *normalized* form of $|v\rangle$, for any non-zero vector $|v\rangle$.

Definition (Orthonormal Set)**.** A set $\{|e_1\rangle, ..., |e_n\rangle\}$ of vectors is said to be *orthonormal* if

$$\langle e_i|e_j\rangle = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

If the set $\{|e_1\rangle, ..., |e_n\rangle\}$ is a basis, then we say it is an *orthonormal basis.*

With these conventions, the inner product on a finite-dimensional vector space can be given a convenient matrix representation. Let $|v\rangle = \sum_{i=1}^n v_i |e_i\rangle$ and $|u\rangle = \sum_{j=1}^n u_j |e_j\rangle$ be the decompositions of the vectors $|u\rangle$ and $|v\rangle$ with respect to an orthonormal basis $\{|e_1\rangle, \ldots, |e_n\rangle\}$. Then, since $\langle e_i|e_j\rangle = \delta_{ij}$

$$\begin{aligned}\langle v|u\rangle = \sum_{i=1}^n \sum_{j=1}^n v_i^* u_j \langle e_i|e_j\rangle &= \sum_{i=1}^n v_i^* u_i \\ &= \begin{pmatrix} v_1^* & \ldots & v_n^* \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}\end{aligned}$$

That is, the inner product of two vectors is equal to the vector inner product between two matrix representations of those vectors, provided the representation are written with respect to the same orthonormal basis. We also see that the dual vector $\langle v|$ has a nice interpretation as the row vector whose components are complex conjugates of the corresponding components of the column vector representation of $|v\rangle$.

As a small side-note, inner-product spaces also have two useful inequalities.

1. **Schwartz Inequality:** $|\langle v|u\rangle| \leq \|v\| \cdot \|u\|$
2. **Triangle Inequality:** $\|v + u\| \leq \|v\| + \|u\|$

With all the above definitions, we can now define a *Hilbert space*:

Definition (Hilbert Space)**.** A *Hilbert space*, $\mathcal{H}$, is a complex or real inner product space which is complete.

The completeness requirement is a mathematical aspect which we will never need to consider in this course. It just says that all Cauchy sequences converges to points in $\mathcal{H}$. Note that a Hilbert space does not need to be infinite dimensional, which is a common misconception by new students of quantum mechanics. However, we will consider many cases in which it is infinite dimensional.

2.4 Concluding example

Examples are usually more useful than general descriptions, so let's put some of these ideas to use. We'll consider the vector space of real polynomials with degree 0 through 2. This is a three dimensional vector space, since the minimum collection of polynomials needed to span this space is three: for instance $\{1, x, x^2\}$. Let's equip this space with an inner product:

$$(|P\rangle, |Q\rangle) = \int_{-1/2}^{1/2} dx P(x)Q(x) \tag{3}$$

where $P(x)$ and $Q(x)$ are polynomials. The most obvious basis to work with is the basis $\{|0\rangle, |1\rangle, |2\rangle\} = \{1, x, x^2\}$. However, this is neither normalized nor orthogonal, so is not the most convenient choice. We can build an orthogonal basis by *projecting out* the parts of $\{|0\rangle, |1\rangle, |2\rangle\}$ which are not orthogonal. We have

$$
\begin{aligned}
\langle 0|0\rangle &= \int_{-1/2}^{1/2} dx 1 = 1 \\
\langle 0|1\rangle &= \int_{-1/2}^{1/2} dx x = 0 \\
\langle 1|2\rangle &= \int_{-1/2}^{1/2} dx x^3 = 0 \\
\langle 1|1\rangle &= \int_{-1/2}^{1/2} dx x^2 = \frac{1}{12} \\
\langle 0|2\rangle &= \int_{-1/2}^{1/2} dx x^2 = \frac{1}{12}
\end{aligned}
$$

We see that all we have to do is normalize $|1\rangle$ by diving by $\sqrt{1/12}$, then modify the third basis vector to be orthogonal to $|0\rangle$. From 3d vector analysis, we know that the dot product of two vectors gives the extent to which they point in the same direction, $\vec{A}\cdot\vec{B} = A_{||}|B|$, where $A_{||}$ is the part of $\vec{A}$ that is parallel to $\vec{B}$. If one were subtract off the part of $\vec{A}$ parallel to $\vec{B}$, what remains should be orthogonal. Explicitly, if $\vec{B}$ is normalized, then $\vec{A}' = \vec{A} - (\vec{A}\cdot\vec{B})\vec{B}$ is orthogonal to $\vec{B}$, since

$$\vec{B}\cdot\vec{A}' = \vec{B}\cdot\vec{A} - (\vec{A}\cdot\vec{B})||\vec{B}|| = 0.$$

We can use this to construct the vector $|2'\rangle = |2\rangle - \frac{1}{12}|0\rangle$, which is orthogonal to all the other basis vectors. We can then normalize it by multiplying by $6\sqrt{5}$. This procedure is called the *Grahm-Schmidt* procedure, and can be used to turn any basis into an orthonormal basis. Our final orthonormal basis is

$$
\begin{aligned}
|e_0\rangle &= 1 \\
|e_1\rangle &= 2\sqrt{3}x \\
|e_2\rangle &= 6\sqrt{5}\left(x^2 - \frac{1}{12}\right)
\end{aligned}
$$

If can we decompose a polynomial $|P\rangle$ into a linear combination $|P\rangle = c_0|e_0\rangle + c_1|e_1\rangle + c_2|e_2\rangle$ of these, we can start doing inner products easily, without integrals. This is clearly a big advantage if you need to compute inner products over and over.

Next time we will discuss linear operators and how they can be expressed nicely in Bra-ket notation.

3 Linear Operators

The next thing we need to talk about are linear operators. The example you should keep in you mind is an $n \times n$ matrix which acts on an n-dimensional column vector and transforms it somehow. Another example of an operator, which isn't a matrix, is a derivative acting on the vector space of functions. We now formalize this concept.

3.1 Basic Definitions

Definition (Linear Operator)**.** A *linear operator* between vector spaces V and W is defined to be any function $A : V \to W$ which takes a vector in V as an input and produces a vector in W as an output, and which is linear in its inputs. This means it satisfies:

$$\hat{A}(\alpha\,|v\rangle + \beta\,|u\rangle) = \alpha\hat{A}(|v\rangle) + \beta\hat{A}(|u\rangle)$$

where $|v\rangle, |u\rangle$ are any two vectors in V, and α, β are any (complex) numbers. The vector $\hat{A}(|v\rangle)$ lives in the space W, and is the output when the function $\hat{A}$ is evaluated on the input $|v\rangle$. It is conventional to drop the parentheses and write simply $\hat{A}\,|v\rangle$.

When one says that a linear operator $\hat{A}$ is defined *on* a vector space, V, they usually mean that it is a linear operator from V to V.

Examples.

- An $n \times n$ matrix which acts on n-dimensional column vectors is a linear operator.
- A derivative acting on the vector space of functions is a linear operator.
- The operator acting on functions as $\hat{S}[f(x)] = f(x)^2$ is <u>not</u> linear. (Why not?)
- Dual vectors are linear operators. For a fixed vector $|v_0\rangle$ in inner product space V, we can define a linear operator, $\langle v_0| : V \to \mathbb{C}$, which takes in a vector and spits out a complex number. The linear operator is defined by its action on an arbitrary vector $|w\rangle$, via:

  $$\langle v_0|\,(|w\rangle) \equiv (|v_0\rangle\,, |w\rangle)$$

  This is a linear operator because the inner product is linear in its second slot, and because $\mathbb{C}$ is itself a vector space. We call the linear operator $\langle v_0|$ *the dual vector to* $|v_0\rangle$. For each $|v\rangle$ in V, a unique linear operator $\langle v|$ exists. It is through this construction that $\langle v|$ is defined, and we usually drop the parenthesis to write $\langle v|\,(|w\rangle) = \langle v|w\rangle$.

We will often be interested in operators which map a vector space to itself, i.e. $\hat{A} : V \to V$. If we have two such operators, we can get another operator by composing them:

$$(AB)\left|v\right\rangle = A(B\left|v\right\rangle)$$

Another important object is called the *commutator* which is given by

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$$

If $[\hat{A}, \hat{B}] = 0$ than we say that $\hat{A}$ and $\hat{B}$ *commute*[1]. This will be of great importance in quantum mechanics as only those observables which commute with each other are compatible[2].

3.2 Describing Operators

Usually the most convenient way to understand linear operators is in terms of their *matrix representations* with respect to a basis. This idea hinges on the observation that *a linear operator is fully specified by its action on the basis vectors*. To see this, let $\{\left|b_1\right\rangle, \ldots, \left|b_n\right\rangle\}$ be a basis for an n-dimensional vector space V, and let $\hat{A}$ be some linear operator on V. When I say that the action of $\hat{A}$ on the basis is known, I mean that we know what $\hat{A}\left|b_j\right\rangle$ is for each $j = 1, \ldots, n$; let's call it $\left|b'_j\right\rangle$. Then since $\left|b'_j\right\rangle$ is just some vector in V, we can expand it as a linear combination of the basis vectors:

$$\hat{A}\left|b_j\right\rangle = \sum_{i=1}^{n} A_{ij}\left|b_i\right\rangle$$

So the collection of n^2 numbers A_{ij} tell us the action of $\hat{A}$ on the basis vectors.

Now let $\left|v\right\rangle$ be some other vector in the space with decomposition $\left|v\right\rangle = \sum_j v_j \left|b_j\right\rangle$ with respect to the basis. Acting on this with the operator $\hat{A}$ we have:

$$\hat{A}(\left|v\right\rangle) = \hat{A}\left(\sum_{j=1}^{n} v_j \left|b_j\right\rangle\right)$$

Now we can use the fact that the operator is linear:

$$= \sum_{j=1}^{n} v_j \hat{A}\left(\left|b_j\right\rangle\right)$$

[1]In this case, $AB = BA$. That's what I would call commuting!

[2]*Compatible* means a measurement of one does not disrupt or affect a measurement of the other. Position and momentum are not compatible observables.

and the known action of $\hat{A}$ on basis vectors:

$$= \sum_{j=1}^{n} v_j \sum_{i=1}^{n} A_{ij} |b_i\rangle$$
$$= \sum_{i,j=1}^{n} A_{ij} v_j |b_i\rangle$$

So, we see that we can compute $\hat{A}(|v\rangle)$, for any vector $|v\rangle$, provided we know the numbers A_{ij}. But if you squint even harder, you will see something very nice. The matrix representation of the vector $\hat{A}|v\rangle$ with respect to basis $\{|b_i\rangle\}$ is the same as the matrix representation of $\hat{A}$ multiplying the matrix representation of $|v\rangle$:

$$\hat{A}(|v\rangle) = \sum_{i,j=1}^{n} A_{ij} v_j |b_i\rangle \xrightarrow{\{b\}} \begin{pmatrix} A_{11} & A_{12} & \dots \\ \vdots & \ddots & \\ A_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

We thus call the numbers A_{ij} the matrix representation of $\hat{A}$ with respect to the basis $\{|b_i\rangle\}$. An example is in order.

Examples.

- Suppose V is a 2d vector space with basis vectors $|0\rangle$ and $|1\rangle$. The linear operator $\hat{A}$ is known to have action $\hat{A}|0\rangle = |1\rangle$ and $\hat{A}|1\rangle = 2|0\rangle$. The matrix representation of $\hat{A}$ with respect to this basis is then:

  $$\hat{A} \longmapsto \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$$

  You can read this as follows. The first column tells you where the first basis vector $|0\rangle$ lands, and the second column tells you where the second basis vector $|1\rangle$ lands. This is a nice rule to remember.

- Go to `https://shad.io/MatVis/` and visualize the above transformation by moving the green dot to $(0, 1)$ and the red dot to $(2, 0)$. Play around and get a feel for what kinds of linear transformations are possible.

It will be convenient to be able to build up linear operators out of basic building blocks. The *outer product* of two vectors $|u\rangle$ and $|v\rangle$, denoted $|u\rangle\langle v|$, is such a building block. It is defined to be the linear operator which acts on inputs $|w\rangle$ via

$$(|u\rangle\langle v|)(|w\rangle) \equiv |u\rangle\langle v|w\rangle = |u\rangle(|v\rangle, |w\rangle).$$

That is, it projects the vector $|w\rangle$ onto the vector $|v\rangle$, and then rotates the result into the direction of $|u\rangle$.

The outer product of two basis vectors drawn from an orthonormal basis is especially useful. Let $\{|e_1\rangle, \ldots, |e_n\rangle\}$ be an orthonormal basis for inner product space V. Let's consider the object:

$$|e_i\rangle \langle e_j|$$

This is a linear operator from V to itself. What is its matrix representation? Allow it to act on a basis vector $|e_k\rangle$:

$$|e_i\rangle \langle e_j| (|e_k\rangle) = \delta_{jk} |e_i\rangle$$

Thus the action of $|e_i\rangle \langle e_j|$ on a basis vector $|e_k\rangle$ is to send it to the zero vector, unless $k = j$, in which case it gets mapped to $|e_i\rangle$. As an example, in a 3d vector space, we have matrix representations:

$$|e_1\rangle \langle e_1| \xrightarrow{\{e\}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad |e_2\rangle \langle e_2| \xrightarrow{\{e\}} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad |e_3\rangle \langle e_3| \xrightarrow{\{e\}} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

as well as so-called off-diagonal elements:

$$|e_1\rangle \langle e_2| \xrightarrow{\{e\}} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad |e_2\rangle \langle e_1| \xrightarrow{\{e\}} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad |e_1\rangle \langle e_3| \xrightarrow{\{e\}} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

etc.. Hopefully you see what I mean by these outer products being building blocks. By taking linear combinations of these building blocks one can create any $n \times n$ matrix they like, for example:

$$\begin{pmatrix} a & b & 0 \\ 0 & 0 & 0 \\ 0 & 0 & c \end{pmatrix} = a \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Similarly, any linear operator $\hat{A}$ on vector space V can be written as

$$\hat{A} = \sum_{i,j} A_{ij} |e_i\rangle \langle e_j|$$

where A_{ij} is its matrix representation with respect to basis $\{|e\rangle\}$.

The outer product can also be calculated directly via matrix multiplication, for example:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

etc.

The most useful formula involving the outer product is described below.

3.3 Completeness Relation

Before continuing, let me introduce you to what will become your favorite formula in linear algebra. It is called the *completeness relation* which says that the identity operator can be decomposed as:

$$\boxed{\mathbb{1} = \sum_{i \in \mathcal{I}} |e_i\rangle \langle e_i|}. \tag{4}$$

where $\mathcal{I}$ is some indexing set which may or may not be infinite (for example $\mathcal{I} = \{1, 2, 3\}$ or $\mathcal{I} = \mathbb{R}$ are possible), and $\{e_i\}_{i \in \mathcal{I}}$ is an orthonormal basis.

Proof. To verify that this is indeed the identity operator, we need to show that $\mathbb{1} |v\rangle = |v\rangle$ for any vector $|v\rangle$ in the space V. To this end, let $|v\rangle$ by a vector in V. It has components $v_i = \langle e_i|v\rangle$ with respect to orthonormal basis $\{|e_i\rangle\}$. We have:

$$\mathbb{1} |v\rangle = \sum_{i \in \mathcal{I}} |e_i\rangle \langle e_i|v\rangle = \sum_{i \in \mathcal{I}} |e_i\rangle v_i = |v\rangle$$

$\square$

The interpretation of this equation is very simple: it simply restates the fact that every vector can be decomposed into a linear combination of basis vectors. If you squint at the matrix representations of $|e_1\rangle \langle e_1|$, $|e_2\rangle \langle e_2|$, and $|e_3\rangle \langle e_3|$, you can already see that the equation is true there.

It is difficult to overstate how useful this equation is.

As an example application, is often used to change the basis under which a vector is decomposed. Assume $\{|e_1\rangle, ..., |e_n\rangle\}$ and $\{|e'_1\rangle, ..., |e'_n\rangle\}$ are two different orthonormal bases (resorting to n-dimensional case). We can the write

$$v = \sum_i v_i |e_i\rangle = \sum_i v_i \Big(\sum_{j=1}^n |e'_j\rangle \langle e'_j| \Big) |e_i\rangle = \sum_{j=1}^n v'_j |e'_j\rangle$$

where the new components v'_j are given by

$$v'_j = \sum_{i=1}^n \langle e'_j|e_i\rangle v_i$$

which is the "change of basis" formula. Let's continue, making use of the completeness relation liberally.

3.4 Working in Orthonormal Bases

In quantum mechanics we are lucky enough to (almost) always have naturally occurring orthonormal bases available to us, so it's worthwhile to discuss this special case separately. As previously mentioned, a linear operator $\hat{A}$ is fully specified by its action on any basis. If that basis happens to be orthonormal, we can find a clean formula for the matrix representation A_{ij}.

Let $\{|e_i\rangle\}_{i\in\mathcal{I}}$ be an orthonormal basis for inner product space V, and let $\hat{A}$ be a linear operator on V. Then, for any vector $|v\rangle$ in V, if we recall the decomposition

$$|v\rangle = \sum_{k\in\mathcal{I}} v_k \, |e_k\rangle \, , \quad v_k = \langle e_k|v\rangle$$

we can write

$$\begin{aligned}\hat{A}\,|v\rangle &= \Big(\sum_{j\in\mathcal{I}} |e_j\rangle\,\langle e_j|\Big)\hat{A}\Big(\sum_{i\in\mathcal{I}} |e_i\rangle\,\langle e_i|\Big)\,|v\rangle \\ &= \sum_{i,j\in\mathcal{I}} A_{ji} v_i \,|e_j\rangle \, , \quad A_{ji} = \langle e_j|\,\hat{A}\,|e_i\rangle\end{aligned}$$

We find that the matrix representation is given by $A_{ij} = \langle e_i|\,\hat{A}\,|e_j\rangle$. The numbers $\langle e_i|\,\hat{A}\,|e_j\rangle$ are called the *matrix elements* of the operator $\hat{A}$. The numbers themselves depend on the basis you're working in, while of course $\hat{A}$ does not.

4 Eigenvalues and Eigenvectors

An *eigenvector* of a linear operator $\hat{A} : V \to V$ is a non-zero vector $|v\rangle$ such that $\hat{A}\,|v\rangle = \lambda\,|v\rangle$, where λ is a complex number known as the *eigenvalue* of $\hat{A}$ corresponding to $|v\rangle$. It will often be convenient to use the same label for both the eigenvalue and the eigenvector, ie $\hat{A}\,|\lambda\rangle = \lambda\,|\lambda\rangle$.

The *eigenspace* corrosponding to eigenvalue λ is the set of all vectors which have eigenvalue λ. It is a vector subspace of the vector space on which $\hat{A}$ acts.

Examples. Consider the linear operator $\hat{A}$ which acts on a two-dimensional vector space with basis vectors $|e_1\rangle$ and $|e_2\rangle$ via $\hat{A}\,|e_1\rangle = |e_2\rangle$, and $\hat{A}\,|e_2\rangle = |e_1\rangle$. The matrix representation of $\hat{A}$ in this basis is

$$\hat{A} \xrightarrow{\{e\}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Consider the vectors $|a\rangle = |e_1\rangle + |e_2\rangle$, and $|b\rangle = |e_1\rangle - |e_2\rangle$. They have matrix representations:

$$|a\rangle \xrightarrow{\{e\}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad , \quad |b\rangle \xrightarrow{\{e\}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

The action of $\hat{A}$ on these vectors is

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = +1 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$
$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = -1 \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

so we see that $|a\rangle$ is an eigenvector of $\hat{A}$ with eigenvalue $+1$, and $|b\rangle$ is an eigenvector of $\hat{A}$ with eigenvalue -1.

There is an algorithm for finding all the eigenvalues of a matrix that is widely taught in linear algebra classes. You write out the formula $\mathrm{Det}(A - \lambda I) = 0$ explicitly as a polynomial in λ, and then solve for the roots λ^* of the equation. The roots are the eigenvalues. I will not describe it more here, since it is not of that much interest to us, but one useful takeaway from this formula is that there is always at least one eigenvalue of every operator, since every polynomial has at least one complex root.

When an eigenspace is more than one dimensional, we say that it is *degenerate*. For example, the matrix A given by

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

has a two-dimensional eigenspace corrosponding to the eigenvalue 2. The eigenvectors $(1, 0, 0)$ and $(0, 1, 0)$ are said to be *degenerate* because they are linearly independant eigenvectors of A with the same eigenvalue.

5 Hermitian and Unitary Operators

Here we would like to discuss two special types of operators which are without question the most important in the study of quantum mechanics.

Definition (Adjoint)**.** The *adjoint* (or *Hermitian conjugate*) of a matrix is defined in a basis as $(A^\dagger)_{ij} = A^*_{ji}$. This could also be written $(A)^\dagger = (A^*)^T = (A^T)^*$, where T denotes the transpose, and $*$ denotes complex conjugation of all components. For example,

$$\begin{pmatrix} 1 & 2i \\ 1+i & -3i \end{pmatrix}^\dagger = \begin{pmatrix} 1 & 1-i \\ -2i & 3i \end{pmatrix}$$

The symbol † is usually pronounced "dagger".

We can also define it independently of basis. The adjoint of an operator $\hat{A}$ on inner product space V is defined to be the operator $\hat{A}^\dagger$ for which

$$(|u\rangle, \hat{A}|v\rangle) = (\hat{A}^\dagger |u\rangle, |v\rangle) \tag{5}$$

for any vectors $|u\rangle, |v\rangle$ in V. This definition is the reason I have adopted the $(\cdot, \cdot)$ notation[3]. Once this is comfortable, you should start using $\langle \cdot | \cdot \rangle$ without issue.

The adjoint of an operator is *unique*: suppose $\hat{B}$ and $\hat{C}$ are two would-be adjoints of $\hat{A}$, that is, operators satisfying Eq. 5 in place of $\hat{A}^\dagger$. Then consider the quantity $((\hat{B} - \hat{C})|u\rangle, |v\rangle)$. By the (anti) linearity of the inner product, this is equal to $(\hat{B}|u\rangle, |v\rangle) - (\hat{C}|u\rangle, |v\rangle)$. Each term is equal to $(|u\rangle, A|v\rangle)$ by assumption, so the expression vanishes for all $|u\rangle, |v\rangle$. Therefore $(\hat{B} - \hat{C})$ is the zero operator, and so in fact $\hat{B} = \hat{C}$.

If we use the notation $\hat{A}|v\rangle = |\hat{A}v\rangle$, then the adjoint is defined by $\langle u|\hat{A}v\rangle = \langle \hat{A}^\dagger u|v\rangle$. Note that this means $\langle Au| = \langle u| A^\dagger$ but this makes some sense since $\langle cu| = \langle u| c^*$.

Exercise.

- If $|w\rangle$ and $|v\rangle$ are any two vectors, show that $(|w\rangle \langle v|)^\dagger = |v\rangle \langle w|$.
- Show that the adjoint is anti-linear:

$$\left(\sum_i a_i \hat{A}_i \right)^\dagger = \sum_i a_i^* A_i^\dagger$$

- Show that $(\hat{A}^\dagger)^\dagger = \hat{A}$.
- Show that $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger \hat{A}^\dagger$.

[3] Otherwise the formula needs to be written awkwardly as $\langle u|Av\rangle = \langle A^\dagger u|v\rangle$, since the notation $\langle u|A|v\rangle$ is ambiguous.

5.1 Hermitian Operators

Definition (Hermitian Operator)**.** A *hermitian operator* is an operator $\hat{H}$ satisfying $\hat{H}^\dagger = \hat{H}$. It is said to be *self-adjoint*.

Hermitian operators have very nice eigenvalues and eigenvectors as we now see. This is very important.

Theorem. *Hermitian operators have real eigenvalues.*

Proof. Let $|\lambda\rangle$ be an eigenvector of a hermitian operator $\hat{H}$ such that $\hat{H}\,|\lambda\rangle = \lambda\,|\lambda\rangle$. Then without loss of generality assume $|\lambda\rangle$ is normalized to one. Then we have

$$\lambda = \langle\lambda|\hat{H}\lambda\rangle = \langle\lambda|\hat{H}|\lambda\rangle = \langle\lambda|\,\hat{H}^\dagger\,|\lambda\rangle = \langle\hat{H}\lambda|\lambda\rangle = \langle\lambda|\,\lambda^*\,|\lambda\rangle = \lambda^*$$

Then the only way we can have $\lambda = \lambda^*$ is if λ is real. □

Theorem. *If $|\lambda_1\rangle$ and $|\lambda_2\rangle$ are two eigenvectors of the hermitian operator $\hat{H}$ with eigenvalues $\lambda_1 \neq \lambda_2$ then $\langle\lambda_1|\lambda_2\rangle = 0$.*

Proof.

$$0 = \langle\lambda_1|(H - H^\dagger)|\lambda_2\rangle = \langle\lambda_1|H\lambda_2\rangle - \langle H\lambda_1|\lambda_2\rangle = (\lambda_2 - \lambda_1)\,\langle\lambda_1|\lambda_2\rangle$$

Since $(\lambda_2 - \lambda_1) \neq 0$ we must have $\langle\lambda_1|\lambda_2\rangle = 0$. □

Exercise.

- Show that if $\hat{A}$ and $\hat{B}$ are Hermitian, then $\hat{A} + \hat{B}$ is also Hermitian.
- Show that if $\hat{A}$ and $\hat{B}$ are Hermitian, then $\hat{A}\hat{B}$ is Hermitian provided that $\hat{A}$ and $\hat{B}$ commute with each other.
- Show that for any operator $\hat{A}$, the operator $\hat{A}^\dagger\hat{A}$ is Hermitian, and that all its eigenvalues are *positive* real numbers.

We will have more to say about Hermitian operators in a moment, but we need to introduce their close friends, unitary operators.

5.2 Unitary Operators

Definition (Unitary Operator)**.** A *unitary operator* is an operator $\hat{U}$ satisfying $\hat{U}^\dagger\hat{U} = \hat{U}\hat{U}^\dagger = I$. In other words, its adjoint is its inverse.

The most important property of unitary operators for quantum mechanics is that they preserve the inner product:

$$(U|u\rangle, U|v\rangle) = (|u\rangle, U^\dagger U|v\rangle) = (|u\rangle, |v\rangle)$$

In braket notation:

$$\langle Uv|Uw\rangle = \langle v| U^\dagger U |w\rangle = \langle v|w\rangle$$

In quantum mechanics, if you have a quantum state $|\psi(t)\rangle$ and measure an operator $\hat{X}$ with eigenvectors $|x_i\rangle$, the probability of measuring the eigenvalue x_i is given by $|\langle x_i|\psi(t)\rangle|^2$. We require that the sum of all these probabilities add to 1 (in other words we must measure some value). This gives the requirement

$$1 = \sum_{x_i} |\langle x_i|\psi(t)\rangle|^2 = \sum_{x_i} \langle\psi(t)|x_i\rangle\langle x_i|\psi(t)\rangle = \langle\psi(t)| \Big(\sum_{x_i} |x_i\rangle\langle x_i|\Big) |\psi(t)\rangle = \langle\psi(t)|\psi(t)\rangle$$

If we assume the state is normalized at time $t = 0$ and evolution is linear then $|\psi(t)\rangle = \hat{E}_t|\psi(0)\rangle$ which gives

$$1 = \langle\psi(t)|\psi(t)\rangle = \langle\psi(0)| \hat{E}_t^\dagger \hat{E}_t |\psi(0)\rangle$$

The only way this is true for all states is if $\hat{E}_t^\dagger \hat{E}_t = I \implies$ evolution is unitary.

Although we wont often be interested in calculating eigenvalues or eigenvectors of unitary matrices in this class we can say a few things:

- The eigenvalues of $\hat{U}$ are of the form $e^{i\phi}$, $\phi \in \mathbb{R}$.

  Proof. Let $|\lambda\rangle$ be an eigenvectors with $\hat{U}|\lambda\rangle = \lambda|\lambda\rangle$. Then we have

  $$1 = \langle\lambda|\lambda\rangle = \langle\lambda|U^\dagger U|\lambda\rangle = |\lambda|^2\langle\lambda|\lambda\rangle = |\lambda|^2 \implies \lambda = e^{i\phi}$$

  □

- The eigenvectors of $\hat{U}$ with different eigenvalues are orthonogonal.

  Proof. Let $|\lambda\rangle$ and $|\mu\rangle$ be eigenvectors of $\hat{U}$ with $\lambda \neq \mu$. Then we have

  $$\langle\lambda|\mu\rangle = \langle\lambda|U^\dagger U|\mu\rangle = \lambda^*\mu\langle\lambda|\mu\rangle$$

  Since λ is of the form $e^{i\phi}$ we have $\lambda^* = \lambda^{-1}$. Then since $\mu \neq \lambda$ we have $\lambda^{-1}\mu \neq 1 \implies \langle\lambda|\mu\rangle = 0$. We get similar results for the diagonalization of unitary matrices as hermitian matrices. In fact both unitary and hermitian matrices are examples of something called *normal matrices* which is why they have several similar nice properties. □

Unitary operators have a nice representation in terms of outer products. Let $|v_i\rangle$ be any orthonormal basis. Define $|w_i\rangle \equiv \hat{U}|v_i\rangle$. Because unitary operators preserve inner products, the set $\{|w_i\rangle\}$ is also an orthonormal basis. The unitary operator $\hat{U}$ can then be written as:

$$\hat{U} = \sum_i |w_i\rangle \langle v_i|$$

To check this, we need only confirm its action on some basis. It clearly has the correct action on the basis $\{|v_i\rangle\}$, so it must be the correct operator.

Conversely, if $\{|v_i\rangle\}$ and $\{|w_i\rangle\}$ are two orthonormal bases, then the operator $\hat{U}$ defined by $\hat{U} \equiv \sum_i |w_i\rangle \langle v_i|$ is always a unitary operator.

Exercise. What are the matrix elements of operator $\hat{U} \equiv \sum_i |w_i\rangle \langle v_i|$ with respect to orthonormal basis $|v_i\rangle$? With respect to the orthonormal basis $|w_i\rangle$?

Let's wrap up with a few more important theorems.

Theorem. *If $\hat{H}$ is a Hermitian matrix, then there exists a unitary matrix $\hat{U}$, with columns given by the eigenvectors of $\hat{H}$, which diagonalizes $\hat{H}$ to the diagonal matrix $\hat{\Lambda}$ with entries given by the eigenvalues of $\hat{H}$.*

Proof. *We prove this for the case where there are no degeneracies. If there are degeneracies, you just have to use Gram-Schmidt to pick a orthonromal eigenvectors in each degenerate subspace.*

Let $\{|\lambda_i\rangle\}_{i\in 1,\ldots,\dim(\hat{H})}$ and $\{\lambda_i\}_{i\in 1,\ldots,\dim(\hat{H})}$ be the eigenvectors and eigenvalues of $\hat{H}$. Without loss of generality we can pick the eigenvectors to be orthonormal. Assume we have $\hat{H}$ written in the basis $\{|e_i\rangle\}_{i\in 1,\ldots,\dim(\hat{H})}$, then we have

$$\begin{aligned}
H_{ij} &= \langle e_i| \hat{H} |e_i\rangle \\
&= \sum_{j,k=1}^{\dim(\hat{H})} \langle e_i|\lambda_j\rangle \langle \lambda_j| \hat{H} |\lambda_k\rangle \langle \lambda_k|e_i\rangle \\
&= \sum_{j,k=1}^{\dim(\hat{H})} \langle e_i|\lambda_j\rangle \lambda_k \delta_{jk} \langle \lambda_k|e_i\rangle \\
&= (\hat{U}\hat{\Lambda}\hat{U}^\dagger)_{ij}
\end{aligned}$$

where $\Lambda_{jk} = \lambda_k \delta_{jk}$ and $U_{ij} = \langle e_i|\lambda_j\rangle$. $\square$

In this proof I actually assumed there even exists $\dim(\hat{H})$ eigenvectors. This is always true for hermitian operators, but not always so for any old operator you may find on the street.

We will often be interested in diagonalizing several hermitian operators at the same time.

Theorem. *Let $\hat{A}$ and $\hat{B}$ both be hermitian operators. If $[\hat{A}, \hat{B}] = 0$ than $\hat{A}$ and $\hat{B}$ can both be simultaneously diagonalized in the same basis.*

Proof. Let $\{|a_i\rangle\}$ be an orthonormal basis which diagonalizes $\hat{A}$ with eigenvalues a_i. Then we have

$$\hat{A}(\hat{B}\,|a_i\rangle) = \hat{B}\hat{A}\,|a_i\rangle = a_i(\hat{B}\,|a_i\rangle)$$

The only way this can be true is if $\hat{B}\,|a_i\rangle = b(a_i)\,|a_i\rangle$ for some $b(a_i)$. Therefore $|a_i\rangle$ are eigenvectors for $\hat{B}$ as well and therefore diagonalize B as well. This is true regardless of the dimesnion of $\hat{A}$ and $\hat{B}$. We often label states which are simultaeously eigenvectors of $\hat{A}$ and $\hat{B}$ by $|a, b\rangle$. ☐

5.3 Expectation Values

As a small aside, lets use what we learned about hermitian operators to find the expectation value. From the postulates of quantum mechanics, for any given observable represented by the hermitian operator $\hat{\Omega}$, a measurement of the state $|\psi\rangle$ will result in the value ω with probability $P(\omega) = |\langle\omega|\psi\rangle|^2$, where ω and $|\omega\rangle$ are the eigenvalue and corresponding eigenvector of $\hat{\Omega}$. Let us say we would like to calculate the *expectation value* (or mean) of $\hat{\Omega}$. This is defined as

$$\begin{aligned}
\left\langle \hat{\Omega} \right\rangle_\psi &= \sum_\omega \omega P(\omega) = \sum_\omega \omega \langle\psi|\omega\rangle \langle\omega|\psi\rangle = \sum_\omega \langle\psi|\hat{\Omega}|\omega\rangle \langle\omega|\psi\rangle \\
&= \langle\psi|\,\hat{\Omega}\Big[\sum_\omega |\omega\rangle\langle\omega|\Big]\,|\psi\rangle = \langle\psi|\,\hat{\Omega}I\,|\psi\rangle \\
&= \langle\psi|\,\hat{\Omega}\,|\psi\rangle
\end{aligned}$$

which is a nice expression to know.

Another extremely important property of unitary matrices which will be very important for the study of quantum mechanics is the relationship between hermitian and unitary matrices. It turns out that there is an bijection between the set of hermitian matrices and the set of unitary matrices of the form

$$\hat{U} = e^{i\hat{H}} \longleftrightarrow \hat{H} = \log\Big(-i\hat{U}\Big)$$

where the exponential of a matrix (or any function of a matrix) is defined by its Taylor series

$$f(\hat{A}) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)\hat{A}^n}{n!}$$

Two checks you can do to convince yourself that this is logical is to count the degrees of freedom and also to consider a unitary matrix very close to the identity

$$\begin{aligned} \hat{U} = I + i\varepsilon\hat{A} + \mathcal{O}(\varepsilon^2) \implies I &= \hat{U}^\dagger\hat{U} = I + i\varepsilon(\hat{A} - \hat{A}^\dagger) + \mathcal{O}(\varepsilon^2) \\ &\implies \hat{A}^\dagger = \hat{A} \implies \hat{A} \text{ is hermitian} \end{aligned}$$

This will be VERY important as every symmetry of the quantum system will correspond to a unitary matrix and will be related to a hermitian operator (a physical observable) as above. You may have seen this before as *Noether's Theorem* and the relations

$$\begin{aligned} \text{Time translation symmetry} &\longleftrightarrow \text{Hamiltonian} \\ \text{Space translation symmetry} &\longleftrightarrow \text{Momentum} \\ \text{Rotation symmetry} &\longleftrightarrow \text{Angular Momentum} \end{aligned}$$

We will make this more concrete as we progress in the course.