Linear Algebra

William F. Barnes

October 20, 2020

Contents

1 Symbols . . . 2
2 Axioms . . . 3
3 Inner Product . . . 5
4 Linear Combinations . . . 7
5 Orthonormal Basis . . . 8
6 Normed Vector Space . . . 12
7 Countably Finite Systems . . . 15
8 Countably Infinite Systems . . . 17
9 Operators . . . 19
10 Eigenvectors . . . 22
11 Matrix Operators . . . 24
12 Hermitian Matrix . . . 26
13 Matrix in Hilbert Subspace . . . 28
14 Functions of Operators . . . 30
15 Unitary Operators . . . 31

1 Symbols

Ket Vectors

We begin by replacing the familiar vector notation $\vec{A}$ with the *ket* notation

$$\vec{A} \qquad\qquad \rightarrow \qquad\qquad |A\rangle$$

popularized by Paul Dirac. The object $|\,\rangle$ is equivalent to a vector, and contains all of the information a vector would carry. For instance, a two-component vector may occur as

$$\vec{A} = (A_x, A_y) \qquad\qquad \rightarrow \qquad\qquad |A_x, A_y\rangle\ .$$

Scalars

Any vector component is most generally a complex number z, which always has a complex conjugate z^* such that

$$z = (A, B) = r\,e^{i\theta} \qquad\qquad z^* = (A, -B) = r\,e^{-i\theta}\ ,$$

where A, B, r, θ are real numbers, and $i = \sqrt{-1}$. For two complex numbers $z_1 = (A_1, B_1)$ and $z_2 = (A_2, B_2)$, the addition, multiplication, and division rules are contained in

$$\begin{aligned} z_1 + z_2 &= (A_1 + B_1, A_2 + B_2) \\ z_1 \cdot z_2 &= (A_1 A_2 - B_1 B_2, A_1 B_2 + B_1 A_2) \\ \frac{z_1}{z_2} &= \frac{z_1 \cdot z_2^*}{|z_2|^2}\ , \end{aligned}$$

embedding the notions of commutativity, associativity, and other algebraic properties.

Sets

A *set* is generally defined as a collection of distinct, well-defined objects. Perhaps the most common set is the real numbers, denoted $\mathcal{R}$. Distinguishing the set of integers $\mathcal{Z}$ from the irrational numbers $\mathcal{Q}'$, we can relate each set using the *union* operator:

$$\mathcal{R} = \mathcal{Z} \cup \mathcal{Q}'$$

An individual member of a set is called an *element*. For example, the set $\mathcal{C}$ of complex numbers is comprised of all elements z. This is formally denoted using the *in* symbol $\in$ as

$$z \in \mathcal{C}\ .$$

Mathematical statements can be shortened further by introducing the *for all* symbol $\forall$, along with the *there exists* symbol $\exists$. For instance, the idea that 'for all complex numbers z there exits a complex conjugate z^*' can be written as:

$$\forall z \in \mathcal{C}\ ,\ \exists\, z^* \in \mathcal{C}$$

Spaces

A *space* is a set with some kind of ordered structure. For instance, the space of all ordered pairs of real numbers, i.e., all two-dimensional vectors with real components, is denoted $\mathcal{R}_2$. For a less trivial example, we may define a space $\mathcal{L}_2\left[a, b\right]$ of all functions $\{f\left(x\right)\}$ obeying

$$\int_a^b \left|f\left(x\right)\right|^2 \, dx < \infty \, .$$

A *vector space*, for a given vector $\left|A\right\rangle$, contains the set of all allowed vectors that $\left|A\right\rangle$ could have been.

2 Axioms

A vector space $\mathcal{V}$, comprised of complex elements $\{\left|a\right\rangle\}$, satisfies the four axioms that follow.

(1) Summation

Vectors add in the 'usual' way, obeying commutativity and associativity, respectively:

$$\left|a\right\rangle + \left|b\right\rangle = \left|b\right\rangle + \left|a\right\rangle$$
$$\left|a\right\rangle + \left(\left|b\right\rangle + \left|c\right\rangle\right) = \left(\left|a\right\rangle + \left|b\right\rangle\right) + \left|c\right\rangle$$

(2) Zero Vector

There exists a zero vector that does not contribute to summation. More precisely, we write

$$\exists \left|0\right\rangle \in \mathcal{V} \, : \, \forall \left|a\right\rangle \in \mathcal{V} \, , \qquad \left|a\right\rangle + \left|0\right\rangle = \left|a\right\rangle \, ,$$

where the colon symbol : literally translates to *such that*.

(3) Additive Inverse

Every vector has a 'negative' version of itself called the *additive inverse*. That is,

$$\forall \left|a\right\rangle \in \mathcal{V} \, : \, \exists \left|-a\right\rangle \in \mathcal{V} \, , \qquad \left|a\right\rangle + \left|-a\right\rangle = \left|0\right\rangle \, .$$

(4) Scalar Multiplication

Multiplying a vector $\left|a\right\rangle \in \mathcal{V}$ by a scalar $\alpha \in \mathcal{C}$ results in a new vector in $\mathcal{V}$ that is parallel to $\left|a\right\rangle$, meaning

$$\forall \left|a\right\rangle \in \mathcal{V} \, , \forall \alpha \in \mathcal{C} \, : \, \exists \, \alpha \left|a\right\rangle \in \mathcal{V} \, .$$

Scalar multiplication obeys the following commutativity and associativity rules (for two scalars α and β):

$$\alpha\left(\beta \left|a\right\rangle\right) = \left(\alpha\beta\right)\left|a\right\rangle$$
$$\alpha\left(\left|a\right\rangle + \left|b\right\rangle\right) = \alpha\left|a\right\rangle + \alpha\left|b\right\rangle$$
$$\left(\alpha + \beta\right)\left|a\right\rangle = \alpha\left|a\right\rangle + \beta\left|a\right\rangle$$

Uniqueness of Zero Vector

The first semi-nontrivial issue to address is whether there exist multiple zero vectors in a given vector space. To capture this concern, take two vectors $|a\rangle$, $|b\rangle$ and add a unique zero vector to each:

$$|a\rangle + |0\rangle = |a\rangle$$
$$|b\rangle + |0\rangle' = |b\rangle$$

Using the shorthand $|a\rangle + |b\rangle = |c\rangle$, add the two equations to get

$$|c\rangle + |0\rangle + |0\rangle' = |c\rangle \ .$$

As it appears, the combination $|0\rangle + |0\rangle'$ can only be the zero vector itself:

$$|0\rangle + |0\rangle' = |0\rangle$$

Evidently, $|0\rangle'$ plays an indistinguishable role from $|0\rangle$. In conclusion, we find there exists exactly one (abbreviated $\exists!$) zero vector per vector space:

$$\exists!\,|0\rangle \in \mathcal{V}$$

Uniqueness of Additive Inverse

In a similar spirit, we can show that the additive inverse of a vector is unique. Take two copies of a vector $|a\rangle$ and add a unique additive inverse vector to each:

$$|a\rangle + |-a\rangle = |0\rangle$$
$$|a\rangle + |-a\rangle' = |0\rangle$$

Adding the two equations, we have

$$2\,|a\rangle + |-a\rangle + |-a\rangle' = |0\rangle \ ,$$

which is only true if

$$|-a\rangle + |-a\rangle' = 2\,|-a\rangle$$
$$|-a\rangle' = |-a\rangle \ ,$$

telling us the additive inverse is unique. Our declaration of the additive inverse becomes more specific:

$$\forall\,|a\rangle \in \mathcal{V} \ : \ \exists!\,|-a\rangle \in \mathcal{V}$$

Subtraction Operator

The notion of subtraction can be formally introduced after establishing uniqueness of the additive inverse:

$$|a\rangle - |b\rangle = |a\rangle + |-b\rangle$$

Identities

Directly following from the above axioms, we can write the trivial 'identity' statement

$$1 \cdot |a\rangle = |a\rangle \ .$$

Another identity emerges from asking which α is required to satisfy $\alpha\,|a\rangle = |0\rangle$. By letting $\alpha \to -\alpha$, we gain a second equation $-\alpha\,|a\rangle = |0\rangle$. Since both equations are still equal, it must follow that $\alpha = -\alpha = 0$, telling us

$$0 \cdot |a\rangle = |0\rangle \ .$$

As a mere notational convenience, we may condense the sum of two vectors as

$$|a\rangle + |b\rangle = |ab\rangle \ ,$$

and multiplication by a scalar as

$$\alpha\,|a\rangle = |\alpha a\rangle \ .$$

Bra Vectors

Finally, we introduce the *bra* notation

$$\langle A| \ ,$$

which denotes the *dual vector* of $|A\rangle$, also called a *linear functional*. The vector space occupied by $\langle A|$ is called the *dual space* to that occupied by $|A\rangle$.

3 Inner Product

As an operator, a bra vector can 'act on' a ket vector to produce a scalar:

$$\langle b|a\rangle = \alpha$$

The quantity $\langle b|a\rangle$ is called the *inner product* of vectors $|b\rangle$ and $|a\rangle$, assumed elements of the same vector space. As an axiom, let us 'use up' some available freedom to require that swapping $a \leftrightarrow b$ results in the complex conjugate of the original product:

$$\langle b|a\rangle = (\langle a|b\rangle)^* = \overline{\langle a|b\rangle}$$

Norm

By calculating $\langle a|a\rangle = \overline{\langle a|a\rangle}$, we readily find $\alpha = \alpha^* = \overline{\alpha}$, telling us the self-inner product always yields a real number:

$$\langle a|a\rangle \in \mathcal{R}$$

The square root the self-inner product is called the *norm* of the vector, axiomatically assumed to be positive:

$$\|a\| = \sqrt{\langle a|a\rangle} > 0$$

It immediately follows that the norm of any non-zero vector cannot itself be zero, with the only excepting being the zero vector:

$$\langle a|a\rangle = 0 \qquad \Longleftrightarrow \qquad |a\rangle = |0\rangle$$

Linearity

Finally, we can also require the linearity relation

$$\langle a|\alpha u + \beta v\rangle = \alpha \langle a|u\rangle + \beta \langle a|v\rangle \ .$$

Using what we've established, it follows that

$$\begin{aligned}\langle \alpha u + \beta v|a\rangle &= \langle \alpha u|a\rangle + \langle \beta v|a\rangle \\ &= \overline{\langle a|\alpha u\rangle} + \overline{\langle a|\beta v\rangle} \\ &= \alpha^* \langle u|a\rangle + \beta^* \langle v|a\rangle \ .\end{aligned}$$

Negative Product

It is possible for the inner product to yield a negative result. Indeed, pursuing the calculation $\langle a|-a\rangle$ for $|a\rangle \neq |0\rangle$, we find

$$\begin{aligned}\langle a|-a\rangle &= \langle a| \left(|0\rangle - |a\rangle\right) \\ &= \langle a|0\rangle - \langle a|a\rangle \\ &= -\|a\|^2 \ ,\end{aligned}$$

which makes use of the identity

$$\langle a|0\rangle = 0 \ .$$

Examples

Complex Vector Space

Consider the vector space $\mathcal{C}_n$ whose elements are vectors containing n individual complex numbers, i.e.

$$\exists \, |x\rangle \in \mathcal{C}_n \ : \ |x\rangle = |x_1, x_2, ..., x_n\rangle$$

For two vectors $|a\rangle$ and $|b\rangle$ in $\mathcal{C}_n$, the inner product can be defined as

$$\langle a|b\rangle = \sum_{j=1}^{n} a_j^* b_j \ ,$$

or more generally, the definition can include weighting coefficients

$$\langle a|b\rangle = \sum_{j=1}^{n} a_j^* b_j w_j$$

for $w_j > 0 \in \mathcal{R}$.

Complex Function Space

By analogy to the inner product for vectors, a similar equation can be written for two complex functions $f(z)$, $g(z)$ defined in the interval $z \in [a, b]$ as

$$\langle f|g\rangle = \int_a^b f^*(z)\, g(z)\, dz\ .$$

Of course, the above can be generalized with a weighting function $w(z) > 0 \in \mathcal{R}$ such that

$$\langle f|g\rangle = \int_a^b f^*(z)\, g(z)\, w(z)\, dz\ .$$

4 Linear Combinations

Consider a vector space $\mathcal{V}$ admitting a set of n vectors $\{|\phi_j\rangle\}$ with $j = 1, 2, \ldots, n$. Introducing a set of n complex coefficients $\{c_j\}$, we construct a *linear combination*:

$$|a\rangle = \sum_{j=1}^{n} c_j\, |\phi_j\rangle$$

Span

The linear combination vector $|a\rangle$, along with all other linear combinations of $\{|\phi_j\rangle\}$, occupy a sub-space $\mathcal{V}' \in \mathcal{V}$. In tighter terms, we say the vectors $\{|\phi_j\rangle\}$ *span* the vector space $\mathcal{V}'$.

Basis

If it turns out that $\mathcal{V}' = \mathcal{V}$, any vector allowed in $\mathcal{V}$ can be expressed as some linear combination of its elements. In this case, vectors $\{|\phi_j\rangle\}$ are called a *basis*, and the number n is a positive non-infinite integer called the *dimension* of the space.

Linear Independence

While the notion of ‘span’ makes sure there are ‘not too few’ basis vectors, we introduce *linear independence* to assure there aren’t too many. That is, any basis vector $|\phi_k\rangle$ that can be expressed as a linear combination is *not* really a basis vector, and the dimension of the space may shrink by one.

Equivalently, we may argue that a set of linearly independent basis vectors only satisfies

$$\sum_{j=1}^{n} c_j\, |\phi_j\rangle = |0\rangle$$

when all coefficients $c_j = 0$. To show this we choose any two nonzero c_k and $c_{k'}$ (with the rest zero), reducing the above to

$$c_k\, |\phi_k\rangle = -c_{k'}\, |\phi_{k'}\rangle\ .$$

Clearly, the vector $|\phi_{k'}\rangle$ is not independent from $|\phi_k\rangle$ and either can be excluded from the basis.

Uniqueness of Coefficients

We can show that the coefficients c_j are unique for a given linear combination. Supposing we have a resultant vector $|a\rangle$ that that is 'arrived at' by two different sets of coefficients

$$|a\rangle = \sum_{j=1}^{n} c_j |\phi_j\rangle \qquad\qquad |a\rangle = \sum_{j=1}^{n} c'_j |\phi_j\rangle \ .$$

Adding each equation and dividing by 2, we quickly find

$$|a\rangle = \sum_{j=1}^{n} \left(\frac{c_j + c'_j}{2} \right) |\phi_j\rangle \ ,$$

which only holds if every c_j is equal to c'_j.

5 Orthonormal Basis

Orthogonality

Two vectors $|\phi_j\rangle$, $|\phi_k\rangle$ are *orthogonal vectors* if their inner product is zero:

$$\langle \phi_j | \phi_k \rangle = 0$$

If *all* basis vectors $\{|\phi_j\rangle\}$ are mutually orthogonal, they constitute an *orthogonal basis.*

Normalized Basis

A basis vector is *normalized* if its self-inner product resolves to one:

$$\langle \phi_j | \phi_j \rangle = 1 \ ,$$

in which case the change of notation

$$|\phi_j\rangle \qquad \rightarrow \qquad |e_j\rangle$$

is made. Of course, one can always normalize each vector in an orthogonal basis by dividing out the norm:

$$|e_j\rangle = \frac{1}{\sqrt{\langle \phi_j | \phi_j \rangle}} |\phi_j\rangle$$

If all basis vectors are mutually orthogonal and have a norm of one, the set $\{|e_j\rangle\}$ is called an *orthonormal basis.* We summarize this by writing

$$\langle e_j | e_k \rangle = \delta_{jk} \ ,$$

where δ_{jk} is the Kronecker delta symbol.

Gram-Schmidt Procedure

It turns out that an arbitrary basis $\{|\phi_j\rangle\}$ can always be transformed into an orthonormal basis by the *Gram-Schmidt procedure*. Denote $\{|e_j\rangle\}$ as the desired set of unit-normalized orthogonal vectors, and $\{|e'_j\rangle\}$ as a non-normalized version (a notational convenience). Starting with the $j = 1$ vector, we write the easy result

$$|e'_1\rangle = |\phi_1\rangle \qquad\qquad |e_1\rangle = |e'_1\rangle / \sqrt{\langle e'_1|e'_1\rangle} \,.$$

Next, we need a new vector $|e'_2\rangle$ that involves $|\phi_2\rangle$ and is orthogonal to $|e_1\rangle$. This is achieved by writing

$$|e'_2\rangle = |\phi_2\rangle - \langle e_1|\phi_2\rangle |e_1\rangle \qquad\qquad |e_2\rangle = |e'_2\rangle / \sqrt{\langle e'_2|e'_2\rangle} \,.$$

Continuing for $j = 3$, we need a vector $|e'_3\rangle$ that involves $|\phi_3\rangle$ and is orthogonal to $|e_1\rangle$, $|e_2\rangle$, satisfied by

$$|e'_3\rangle = |\phi_3\rangle - \langle e_1|\phi_3\rangle |e_1\rangle - \langle e_2|\phi_3\rangle |e_2\rangle \ ,$$

subject to the same normalization rule. In the general $j = n$ case, we must have

$$|e'_n\rangle = |\phi_n\rangle - \langle e_1|\phi_n\rangle |e_1\rangle - \langle e_2|\phi_n\rangle |e_2\rangle - \cdots - \langle e_{n-1}|\phi_n\rangle |e_{n-1}\rangle \ ,$$

normalized by

$$|e_n\rangle = \frac{1}{\sqrt{\langle e'_n|e'_n\rangle}} |e'_n\rangle \ .$$

Components

Equipped with the notion of the orthonormal basis, let us consider a linear combination

$$|a\rangle = \sum_{j=1}^{n} a_j |e_j\rangle \ ,$$

and solve for the coefficients a_j. Using what's sometimes called Fourier's trick, notice that multiplying through by any bra vector $\langle e_k|$ will trigger one inner product on the left, and n inner products on the right. However $n - 1$ of these will be *zero*:

$$\langle e_k|a\rangle = \sum_{j=1}^{n} a_j \langle e_k|e_j\rangle = 0 + \cdots + a_j\,\delta_{jk} + \cdots + 0 = a_k$$

Evidently, any coefficient c_j can be reverse-engineered from a linear combination by the relation

$$a_j = \langle e_j|a\rangle \ .$$

The coefficients a_j are synonymous with the *components* of a vector. In pure bra-ket notation, a linear combination reads

$$|a\rangle = \sum_{j=1}^{n} \langle e_j|a\rangle |e_j\rangle \ ,$$

reminding us that the components of a vector are strictly related to the choice of basis.

Isomorphism

Consider a vector space $\mathcal{V}$ admitting a basis $\{|e_j\rangle\}$. A linear combination vector $|a\rangle$, in component form, can be written

$$|a\rangle = |a_1, a_2, \ldots, , a_n\rangle \ .$$

On the right side, we see that the (complex) components form an n-dimensional space of their own, namely $\mathcal{C}_n$. To capture this observation of 'one-to-oneness' between the original vector space and that occupied by its components, we say the n-dimensional inner product space is *isomorphic* with $\mathcal{C}_n$, or

$$\mathcal{V}_{(n)} \cong \mathcal{C}_n \ .$$

Examples

Arbitrary Basis

Use the Gram-Schmidt procedure to produce an orthonormal basis from:

$$|\phi_1\rangle = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \qquad |\phi_2\rangle = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \qquad |\phi_3\rangle = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Starting with $|e_1\rangle$, we have

$$|e_1\rangle = \frac{1}{\sqrt{\langle\phi_1|\phi_1\rangle}} |\phi_1\rangle = \frac{1}{\sqrt{2}} |\phi_1\rangle \ ,$$

telling us $|e_2'\rangle$ is given by

$$|e_2'\rangle = |\phi_2\rangle - \langle e_1|\phi_2\rangle |e_1\rangle = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix} \qquad |e_2\rangle = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} \ .$$

Finally, for $|e_3'\rangle$, we have

$$|e_3'\rangle = |\phi_3\rangle - \langle e_1|\phi_3\rangle |e_1\rangle - \langle e_2|\phi_3\rangle |e_2\rangle = \begin{bmatrix} 2/3 \\ -2/3 \\ 2/3 \end{bmatrix} \ ,$$

normalizing to

$$|e_3\rangle = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \ .$$

Legendre Polynomials

Consider the set of real-valued polynomial functions of order no greater than four

$$P_4(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + c_4 x^4 \,,$$

known as *Legendre polynomials*. Confining the x-domain to the window $-1 \le x \le 1$, we may introduce the inner product of two such polynomials $f(x)$, $g(x)$ as

$$\langle f|g\rangle = \int_{-1}^{1} f(x)\, g(x)\, dx \,.$$

The vectors $|\phi_j\rangle = x^j$ with $j = 0, 1, 2, 3, 4$ form the basis of a five-dimensional vector space.

By the Gram-Schmidt procedure, we can normalize the basis $\{|\phi_j\rangle\}$ starting with

$$|e_0\rangle = \frac{1}{\sqrt{\int_{-1}^{1} dx}} |\phi_0\rangle = \frac{1}{\sqrt{2}} |\phi_0\rangle = \frac{1}{\sqrt{2}} \,.$$

Proceeding for $|e_1\rangle$, we have

$$|e_1'\rangle = |\phi_1\rangle - \langle e_0|\phi_1\rangle |e_0\rangle \qquad\qquad |e_1\rangle = \frac{1}{\sqrt{\langle e_1'|e_1'\rangle}} |e_1'\rangle \,,$$

reducing to

$$|e_1'\rangle = |\phi_1\rangle - \cancel{\langle e_0|\phi_1\rangle |e_0\rangle} \qquad\qquad |e_1\rangle = \sqrt{\frac{3}{2}} |\phi_1\rangle = \sqrt{\frac{3}{2}}\, x \,.$$

Continuing for $|e_2\rangle$, begin with

$$|e_2'\rangle = |\phi_2\rangle - \langle e_0|\phi_2\rangle |e_0\rangle - \cancel{\langle e_1|\phi_2\rangle |e_1\rangle} = |\phi_2\rangle - \frac{\sqrt{2}}{3} |e_0\rangle \,,$$

where normalization requires calculating

$$\langle e_2'|e_2'\rangle = \int_{-1}^{1} \left(x^2 - \frac{1}{3}\right)^2 dx = \frac{8}{45} \,,$$

landing us at

$$|e_2\rangle = \sqrt{\frac{5}{8}} \left(3x^2 - 1\right) \,.$$

Turning the same crank, it's straightforwardly shown that the remaining normalized basis vectors resolve to

$$|e_3\rangle = \sqrt{\frac{7}{8}} \left(5x^3 - 3x\right)$$
$$|e_4\rangle = \frac{3}{8\sqrt{2}} \left(35x^4 - 30x^2 + 3\right) \,.$$

With an orthonormal basis on hand, we can expand an arbitrary function, such as $h(x) = x^4$, in terms of the basis as a linear combination:

$$|h\rangle = \sum_{j=0}^{4} h_j |e_j\rangle \ ,$$

where the components h_j are given by

$$h_j = \langle e_j|h\rangle \ .$$

By symmetry of $h(x) = x^4$, all odd h_j are zero, leaving three calculations to perform:

$$\begin{aligned}
h_0 &= \frac{1}{\sqrt{2}} \int_{-1}^{1} x^4 \, dx = \frac{\sqrt{2}}{5} \\
h_2 &= \sqrt{\frac{5}{8}} \int_{-1}^{1} \left(3x^6 - x^4\right) dx = \sqrt{\frac{8}{5}} \frac{2}{7} \\
h_4 &= \frac{3}{8\sqrt{2}} \int_{-1}^{1} \left(35x^8 - 30x^6 + 3x^4\right) dx = \frac{1}{35} \frac{8\sqrt{2}}{3}
\end{aligned}$$

As a reality check, we can readily verify that $|h\rangle$ still corresponds to x^4, as all other x^n terms cancel out:

$$|h\rangle = \sum_{j=0}^{4} h_j |e_j\rangle = h_0 |e_0\rangle + h_2 |e_2\rangle + h_4 |e_0 4\rangle = |\phi_4\rangle$$

6 Normed Vector Space

Normed Vector Space

The self-inner product of a vector $|a\rangle$, namely

$$\|a\| = \sqrt{\langle a|a\rangle} \in \mathcal{R} \geq 0 \qquad\qquad \|a\| = 0 \iff |a\rangle = 0 \,,$$

was identified as the 'norm' of the vector, a notion that extends to whole vector spaces. A vector space $\mathcal{V}$ is said to be *normed* if two of its elements $|a\rangle$, $|b\rangle$, obey the triangle inequality

$$\|a + b\| \leq \|a\| + \|b\| \ ,$$

easily proven by brute force:

$$\|a + b\| = \sqrt{\langle a + b|a + b\rangle} = \sqrt{\|a\|^2 + \|b\|^2 + \langle a|b\rangle + \langle b|a\rangle}$$

The inner product terms under the square root simplify to

$$\langle a|b\rangle + \langle b|a\rangle = \langle a|b\rangle + \overline{\langle a|b\rangle} = 2\,\mathrm{Re}\left(\langle a|b\rangle\right) \ ,$$

which ranges between 0 and $2\|a\|\|b\|$, proving the triangle inequality. Meanwhile, a normed vector space must also contain the linearity relation

$$\|\alpha a\| = \sqrt{\langle \alpha a|\alpha a\rangle} = |\alpha| \, \|a\|$$

for a complex scalar α.

Examples

Max as Norm

Consider the vector space $\mathcal{R}_2$, i.e. pairs of real numbers (x, y). Let us show that the 'maximum' function

$$\|(x,y)\|_m = \max\{|x|, |y|\}$$

is a norm on $\mathcal{R}_2$.

Taking two vectors

$$|a\rangle = (x_1, y_1) \qquad\qquad |b\rangle = (x_2, y_2) \ ,$$

the 'max' function tells us

$$\|a+b\|_m = \max(|x_1+x_2|, |y_1+y_2|) \ .$$

Note from the triangle inequality that the arguments within the 'max' function obey

$$|x_1+x_2| \le |x_1| + |x_2| \qquad\qquad |y_1+y_2| \le |y_1| + |y_2| \ .$$

Next observe that $|a\rangle$, $|b\rangle$ are subject to

$$|x_1| \le \|a\|_m \qquad\qquad |y_1| \le \|a\|_m$$

$$|x_2| \le \|b\|_m \qquad\qquad |y_2| \le \|b\|_m \ .$$

Summing the x-equations and the y-equations, we find

$$|x_1| + |x_2| \le \|a\|_m + \|b\|_m \qquad\qquad |y_1| + |y_2| \le \|a\|_m + \|b\|_m \ .$$

Tracing back the inequality symbols, we may finally write

$$\begin{aligned}
\|a+b\|_m &= \max(|x_1+x_2|, |y_1+y_2|) \\
&\le \max(|x_1|+|x_2|, |y_1|+|y_2|) \\
&\le \max(\|a\|_m + \|b\|_m, \|a\|_m + \|b\|_m) \\
&\le \|a\|_m + \|b\|_m \ ,
\end{aligned}$$

satisfying a requirement of a norm. To complete the job we must also show a linearity relation:

$$\|\alpha a\|_m = \max(|\alpha x_1|, |\alpha y_1|) = |\alpha| \max(|x_1|, |y_1|) = |\alpha| \, \|a\|_m$$

Sum as Norm

Consider the (same) vector space $\mathcal{R}_2$, i.e. pairs of real numbers (x, y). Let us show that the 'sum' function

$$\|(x,y)\|_s = |x| + |y|$$

is a norm on $\mathcal{R}_2$.

Taking (the same) two vectors

$$|a\rangle = (x_1, y_1) \qquad\qquad |b\rangle = (x_2, y_2) \ ,$$

the 'sum' function gives, same identities used above,

$$\begin{aligned}
\|a+b\|_s &= |a| + |b| \\
&= |x_1 + y_1| + |x_2 + y_2| \\
&\le |x_1| + |y_1| + |x_2| + |y_2| \\
&\le \|a\|_m + \|b\|_m \ .
\end{aligned}$$

To check for linearity, we write

$$\|\alpha a\|_s = |\alpha x_1| + |\alpha y_1| = |\alpha| \left(|x_1| + |y_1|\right) = |\alpha| \, \|a\|_s \ .$$

Unit Ball

In any vector space $\mathcal{V}$, the unit 'ball' $\mathcal{B}_1$ is defined as

$$\mathcal{B} = \{|a\rangle \in \mathcal{V} \ : \ \|a\| \le 1\} \ .$$

Plotting the the 'max' function in the xy-plane, the unit ball resolves to a square frame of side 1, as $\max(a) = 1$ in the unit ball. In terms of the 'sum' function, the ball resolves to a filled diamond with points at $(0, \pm 1)$ and $(\pm 1, 0)$, generated by $|x| + |y| \le 1$.

Identities

Pythagorean Theorem

In the special case $\langle y|x\rangle = \langle x|y\rangle = 0$, the equation of the norm becomes the Pythagorean theorem:

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2$$

Cauchy-Bunyakovsky-Schwarz Inequality

Consider the vector sum $|x\rangle + \lambda |y\rangle = |x + \lambda y\rangle$, where λ is a dimensionless parameter. The norm of such a vector is

$$\begin{aligned}
\|x + \lambda y\|^2 &= \langle x + \lambda y | x + \lambda y\rangle \\
&= \|x\|^2 + \lambda \langle x|y\rangle + \lambda^* \langle y|x\rangle + |\lambda|^2 \|y\|^2 \\
&= \|x\|^2 + 2\,\mathrm{Re}\left(\lambda \langle x|y\rangle\right) + |\lambda|^2 \|y\|^2 \ .
\end{aligned}$$

The middle term has a minimum at 0, and a maximum at $2\,|\lambda|\,\|x\|\|y\|$. In general then, we find

$$2\,\mathrm{Re}\left(\lambda \langle x|y\rangle\right) \le 2\,|\lambda|\,\|x\|\|y\| \ ,$$

leading to the *Cauchy-Bunyakovsky-Schwarz Inequality*:

$$|\langle x|y\rangle| \le \|x\|\|y\|$$

Parallelogram Law

Consider two vectors $|x\rangle$, $|y\rangle$, and two linear combinations $|x+y\rangle$, $|x-y\rangle$. The *parallelogram law* concerns vector addition:

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$$

To prove this, write out the norm of $|x+y\rangle$ and $|x-y\rangle$ and add the resulting equations:

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + \langle x|y\rangle + \langle y|y\rangle$$
$$\|x-y\|^2 = \|x\|^2 + \|y\|^2 - \langle x|y\rangle - \langle y|y\rangle$$

An complementary equation results by subtracting the equations:

$$\|x+y\|^2 - \|x-y\|^2 = 4\,\mathrm{Re}\left(\langle y|x\rangle\right)$$

Polar Identity

We may also write a *polar identity*, given by

$$\langle y|x\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right) ,$$

most easily derived starting from the right:

$$\|x+iy\|^2 = \|x\|^2 - \|y\|^2 + \langle x|iy\rangle + \langle iy|x\rangle$$
$$\|x-iy\|^2 = \|x\|^2 - \|y\|^2 - \langle x|iy\rangle - \langle iy|x\rangle$$

Evidently then,

$$\begin{aligned} i\|x+iy\|^2 - i\|x-iy\|^2 &= i\left(\langle x|iy\rangle + \langle iy|x\rangle + \langle x|iy\rangle + \langle iy|x\rangle\right) \\ &= i\left(i\,\langle x|y\rangle - i\,\langle y|x\rangle + i\,\langle x|y\rangle - i\,\langle y|x\rangle\right) \\ &= 2\left(-\overline{\langle y|x\rangle} + \langle y|x\rangle\right) \\ &= 4\,\mathrm{Im}\left(\langle y|x\rangle\right) \end{aligned}$$

Finally, we verify

$$\begin{aligned} \|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2 &= 4\,\mathrm{Re}\left(\langle y|x\rangle\right) + 4\,\mathrm{Im}\left(\langle y|x\rangle\right) \\ &= 4\left(\mathrm{Re}\left(\langle y|x\rangle\right) + \mathrm{Im}\left(\langle y|x\rangle\right)\right) \\ &= 4\,\langle y|x\rangle\ . \end{aligned}$$

7 Countably Finite Systems

Convergent Sequence

Suppose $\mathcal{V}$ is a normed vector space. The sequence of vectors

$$\{|x_j\rangle \in \mathcal{V}\} \qquad\qquad j = (1, 2, 3, \dots)$$

is said to be *convergent* to the vector $|x\rangle$ if $\forall\ k_\epsilon\ :\ \exists\ \epsilon > 0$, such that if $k > k_\epsilon$, then $\|x - x_k\| < \epsilon$. This statement inspires a definition of a convergent vector:

$$|x\rangle = \lim_{k\to\infty} |x_k\rangle$$

Cauchy Sequence

The (same) sequence of vectors $\{|x_j\rangle \in \mathcal{V}\}$ qualifies as a *Cauchy sequence* if $\forall\, k_\epsilon \,:\, \exists\, \epsilon > 0$, such that if $m, n > k_\epsilon$, then $\|x_m - x_n\| < \epsilon$.

It's easy to show that any convergent sequence qualifies as a Cauchy sequence. For two vectors $|x_m\rangle$, $|x_n\rangle$, we know

$$\|x - x_m\| < \alpha\epsilon \qquad\qquad \|x - x_n\| < \beta\epsilon$$

for two parameters $\alpha, \beta > 0 \in \mathcal{R}$. Adding each, we have

$$\|x - x_m\| + \|x_n - x\| < (\alpha + \beta)\,\epsilon\,,$$

which becomes, using the triangle inequality,

$$\|x - x_m + x_n - x\| < \epsilon\,,$$

reducing to an identity for the Cauchy sequence:

$$\|x_m - x_n\| < \epsilon$$

Complete Spaces

Normed vector spaces in which all Cauchy sequences converge are called *complete spaces*, also known as a Banach spaces. In terms of an orthonormal basis, an arbitrary vector is given by

$$|a\rangle = \sum_{j=1}^{n} a_j \,|e_j\rangle \;\;,$$

where clearly

$$\|a\| = \sqrt{|a|_1^2 + |a|_2^2 + \cdots + |a|_n^2}\,.$$

We may further consider a vector $|b\rangle$ that is itself a Cauchy sequence of vectors $\left|a^{(k\leq m)}\right\rangle$ such that

$$|b\rangle = \sum_{k=1}^{m} \left|a^{(k)}\right\rangle = \sum_{j=1}^{n}\sum_{k=1}^{m} a_j^{(k)}\,|e_j\rangle = \sum_{j=1}^{n} b_j\,|e_j\rangle \;\;,$$

where each coefficient b_j is itself a Cauchy sequence of the complex coefficients a_j:

$$b_j = \sum_{k=1}^{m} a_j^{(k)}$$

Supremum

Consider the complete infinite-dimensional space $\mathcal{C}_{ab}$ of complex-valued functions $f\,(x) \,:\, x \in [a, b]$. Here we define the 'supremum' function

$$\|f\|_{sup} = \max\left\{|f\,(x)| \;:\; x \in [a, b]\right\}\;.$$

The 'sup' norm guarantees homogeneous convergence of a Cauchy sequence of functions $f^{(k)}(x)$ to a single function $f(x)$.

Unlike other norms we've encountered, the $\|f\|_{sup}$ does not bear a notion of inner product. Choosing a trivial example $f(x) = \cos x$, $g(x) = x$ in the interval $x \in [a, b]$, we have

$$\|f\|_{sup} = 1 \qquad\qquad \|g\|_{sup} = \pi$$

$$\|f+g\|_{sup} = 1+\pi \qquad\qquad \|f-g\|_{sup} = -1-\pi \ ,$$

which violates the parallelogram law:

$$\begin{aligned} \|f+g\|_{sup}^2 + \|f+g\|_{sup}^2 &\neq 2\left(\|f\|_{sup}^2 + \|g\|_{sup}^2\right) \\ (1+\pi)^2 + (-1-\pi)^2 &\neq 2\left(1+\pi^2\right) \\ 4\pi &\neq 0 \end{aligned}$$

Hilbert Space

A complete inner product space is called a *Hilbert space*, and we have shown that all finite-dimensional vector spaces are Hilbert spaces. The 'supremum' norm is a unique example of a complete space that is *not* a Hilbert space.

8 Countably Infinite Systems

The results of the previous section readily generalize to handle a *countably infinite* basis.

Fourier Series

Consider an orthonormal basis $\{|e_j\rangle\}$ containing an *infinite* number of basis vectors. An infinite linear combination

$$|x\rangle = \sum_{j=1}^{\infty} \langle e_j|x\rangle \, |e_j\rangle$$

is called the *Fourier series* of the vector $|x\rangle$ in the basis $\{|e_j\rangle\}$, where $\langle x|e_j\rangle$ are called *Fourier coefficients*.

Bessel Inequality

Let us show that any partial sum of a Fourier series is a Cauchy sequence. Truncating the series at the nth term gives

$$\left|x^{(n)}\right\rangle = \sum_{j=1}^{n} \langle e_j|x\rangle \, |e_j\rangle$$

as a partial sum. Another truncation of the series with $m > n$ can be written

$$\left|x^{(m)}\right\rangle = \left|x^{(n)}\right\rangle + \sum_{j=n+1}^{m} \langle e_j|x\rangle \, |e_j\rangle \ ,$$

where the norm of the difference of the two vectors reads

$$\left\|x^{(m)} - x^{(n)}\right\|^2 = \sum_{j=n+1}^{m} \left|\langle e_j|x\rangle\right|^2 \qquad\qquad m > n\,.$$

The above series is assured to be a positive real number, reducing the problem to showing that

$$\sum_{j=1}^{\infty} \left|\langle e_j|x\rangle\right|^2$$

does not diverge.

Proceed by writing out the $m \to \infty$ case, giving

$$\left\langle x - x^{(n)} \middle| x - x^{(n)} \right\rangle = \|x\|^2 - \sum_{j=1}^{n} \left|\langle e_j|x\rangle\right|^2 \geq 0\,.$$

Reading the equation form the right, we arrive at the *Bessel inequality*

$$\sum_{j=1}^{n} \left|\langle e_j|x\rangle\right|^2 \leq \|x\|^2\,,$$

proving that a partial sum of a Fourier series is a Cauchy sequence. Reading the above from left to right, we also have

$$\left\|x - x^{(n)}\right\| = \sqrt{\|x\|^2 - \sum_{j=1}^{n} \left|\langle e_j|x\rangle\right|^2}\,,$$

which in the case of convergence ($n \to \infty$), we get the *Parseval relation*:

$$\|x\|^2 = \sum_{j=1}^{\infty} \left|\langle e_j|x\rangle\right|^2$$

Inner Product Space of Functions

The space $\mathcal{C}_2\left[a,b\right]$ of the inner product of two complex functions was written as

$$\langle f|g\rangle = \int_a^b f^*\left(z\right) g\left(z\right)\, dz\,.$$

For certain cases of the function $f\left(z\right)$, for instance a discontinuous function, the space $\mathcal{C}_2\left[a,b\right]$ is easily shown to not be complete with respect to the usual notion of norm,

$$\|f\| = \sqrt{\int_a^b |f|^2\, dz}\,,$$

implying that a Hilbert space of functions must be carefully discerned.

According to the Riesz-Fisher theorem, we may define a Hilbert of functions with a countably infinite basis, denoted $\mathcal{L}_2\left[a,b\right]$. Furthermore, the Stone-Weierstrass theorem states that the set of polynomials $\{|x_j\rangle\}$ with $j = 1, 2, \ldots$ forms a basis in $\mathcal{L}_2\left[a,b\right]$. Of course, we found such vectors to be non-orthogonal, corrected by the Gram-Schmidt process. The resulting vectors are called *Legendre polynomials*.

9 Operators

An *operator* L is a function that 'acts on' a vector $|x\rangle \in \mathcal{V}$ to create a new vector

$$L|x\rangle = |Lx\rangle = |y\rangle$$

that may or may not live in $\mathcal{V}$.

Linear Operator

An operator that maps a vector to its own vector space $\mathcal{V}$ is said to be *linear* if the relation

$$L|\alpha u + \beta v\rangle = \alpha L|u\rangle + \beta L|v\rangle$$

is satisfied, where α, β are complex scalars. Needless to mention, scalar multiplication is a special case of a linear operator.

Interpreting vectors as functions, we can check whether certain operations for a function $f(x)$ qualify as linear operators. For example, the transformation

$$L(f(x)) = \sin(f(x))$$

fails when tested for linearity:

$$L(\alpha f(x)) = \sin(\alpha f(x)) = \sin\alpha\cos(f(x)) + \cos\alpha\sin(f(x)) \neq \alpha L(f(x))$$

The less trivial example

$$L(f(x)) = \int_0^1 \sin(xy)\, f(y)\, dy$$

does qualify as a linear operator, as the function f enters the integral linearly. Explicitly, we have

$$\begin{aligned} L(\alpha f(x) + \beta g(x)) &= \int_0^1 \sin(xy)\,(\alpha f(y) + \beta g(y))\, dy \\ &= \alpha \int_0^1 \sin(xy)\, f(y)\, dy + \beta \int_0^1 \sin(xy)\, g(y)\, dy \\ &= \alpha L(f(x)) + \beta L(g(x)) \ . \end{aligned}$$

Adjoint Operator

Given an operator L, the *adjoint* operator $L^\dagger$, is defined such that

$$\langle b| L |a\rangle = \overline{\langle a| L^\dagger |b\rangle}$$

readily implying

$$\langle b| L |a\rangle = \left\langle L^\dagger b \middle| a\right\rangle \qquad\qquad \left(L^\dagger\right)^\dagger = L \ .$$

Two linear operators A and B always obey the relation

$$(AB)^\dagger = B^\dagger A^\dagger \ ,$$

proven by writing $\langle u| AB |v\rangle$ two different ways and comparing each right-hand result:

$$\langle u| AB |v\rangle = \left\langle (AB)^\dagger u \middle| v\right\rangle$$
$$\langle u| AB |v\rangle = \left\langle A^\dagger u \middle| Bv\right\rangle = \left\langle B^\dagger A^\dagger u \middle| v\right\rangle$$

Hermitian Operator

An operator that is its own adjoint operator is called *self-adjoint*, also known as *Hermitian*:

$$L = L^\dagger \qquad \rightarrow \qquad \langle b|La\rangle = \langle Lb|a\rangle$$

For self-adjoint operators, the quantity

$$\langle b|\,L\,|a\rangle$$

is always real.

Let us inquire whether the product AB of two Hermitian operators is itself Hermitian. Starting with $(AB)^\dagger = B^\dagger A^\dagger$, let $A = A^\dagger$ and $B = B^\dagger$ to find

$$AB = \left(B^\dagger A^\dagger\right)^\dagger = (BA)^\dagger \ ,$$

telling us AB is Hermitian only if $AB = BA$.

Anti-Hermitian Operator

An *Anti-Hermitian* operator is one that obeys

$$L = -L^\dagger \ .$$

For two Hermitian operators A, B, it turns out that $AB - BA$ is anti-Hermitian:

$$(AB - BA)^\dagger = (AB)^\dagger - (BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger = BA - AB$$

Partial Derivative Operator

The partial defivative operator

$$A = \frac{\partial}{\partial x}$$

is an anti-Hermitian operator for certain boundary conditions. Consider two arbitrary function $f\,(x)$, $g\,(x)$ in the domain Ω where each function is zero on the boundary $\partial\Omega$. Then, writing out $\langle f|\,A\,|g\rangle$ two different ways gives

$$\langle f|Ag\rangle = \left\langle A^\dagger f\middle|g\right\rangle$$
$$\int_\Omega f^* \partial_x g \; dx = \int_\Omega (\partial_x)^\dagger \, f^* g \; dx \ .$$

Integrating the left side by parts, we write

$$f^* g\big|_{\partial\Omega} - \int_\Omega \partial_x f^* g \; dx = \int_\Omega (\partial_x)^\dagger \, f^* g \; dx \ ,$$

where the boundary term equals zero by construction, indicating A to be anti-Hermitian.

Projector

For any fixed vector $|a\rangle$, the combination

$$P_a = |a\rangle \langle a|$$

is called the *projector* of $|a\rangle$, where acting on a vector $|x\rangle$, we have

$$P_a |x\rangle = |a\rangle \langle a|x\rangle = \langle a|x\rangle |a\rangle \ ,$$

which is $|a\rangle$ multiplied by a scalar.

The projector is a linear operator, easily verified by

$$\begin{aligned} P_a |\alpha u + \beta v\rangle &= |a\rangle \langle a| (|\alpha u\rangle + |\beta v\rangle) \\ &= |a\rangle (\alpha \langle a|u\rangle + \beta \langle a|v\rangle) \\ &= \alpha \langle a|u\rangle |a\rangle + \beta \langle a|v\rangle |a\rangle \\ &= \alpha P_a |u\rangle + \beta P_a |v\rangle \ , \end{aligned}$$

and is also Hermitian:

$$\begin{aligned} \langle u|P_a v\rangle &= \langle u| (\langle a|v\rangle |a\rangle) \\ &= \langle a|v\rangle \langle u|a\rangle \\ &= (\langle u|a\rangle \langle a|) |v\rangle \\ &= \langle P_a u|v\rangle \end{aligned}$$

Identity Operator

Consider any vector $|x\rangle$ in the vector space $\mathcal{V}$ spanned by the basis vectors $\{|e_k\rangle\}$. Starting with any basis vector $|e_j\rangle$, apply a projector

$$P_{e_j} = |e_j\rangle \langle e_j|$$

to $|x\rangle$ to get

$$P_{e_j} |x\rangle = \langle e_j|x\rangle |e_j\rangle = x_j |e_j\rangle \ .$$

Summing over the index j, we find

$$\left(\sum_j P_{e_j} \right) |x\rangle = \sum_j x_j |e_j\rangle = |x\rangle \ .$$

The parenthesized quantity that leaves the vector unchanged is called the *identity* operator I. That is,

$$I |x\rangle = |x\rangle \ ,$$

where

$$I = \sum_j |e_j\rangle \langle e_j|$$

is also called the *completeness relation* for the basis. (It is possible for a basis to be *incomplete*, in which case the above sum is not equivalent to an identity operator.)

Commutator

For any two operators A and B, the difference

$$[AB] = AB - BA$$

defines their *commutator*, also known as the *commutation relation*. The result of $[AB]$ tells us which 'extra terms' emerge when swapping two operators. When the commutator evaluates to zero, the operators are said to *commute*.

For example, consider two operators

$$A = \frac{\partial}{\partial x} \qquad\qquad B = x$$

that can act on a function $f(x)$. Allowing the commutator to act on $f(x)$, we write

$$\begin{aligned}(AB - BA)\left|f\right\rangle &= \partial_x\left(x\left|f\right\rangle\right) - x\left(\partial_x\left|f\right\rangle\right) \\ &= \left|f\right\rangle + x\partial_x\left|f\right\rangle - x\partial_x\left|f\right\rangle \\ &= \left|f\right\rangle\ ,\end{aligned}$$

telling us that that operators on hand do not commute, but instead obey

$$AB - BA = I\ .$$

10 Eigenvectors

If an operator L applied to a vector $\left|u\right\rangle$ results in a scaled vector $\lambda\left|u\right\rangle$, then $\left|u\right\rangle$ is called an *eigenvector* of L, and λ is its *eigenvalue*:

$$L\left|u\right\rangle = \lambda\left|u\right\rangle$$

Real Eigenvalues

It's straightforward to show that eigenvalues are always real if L is self-adjoint:

$$\left\langle u\right| L\left|u\right\rangle = \lambda\left\langle u|u\right\rangle \qquad \rightarrow \qquad \lambda = \frac{\left\langle u\right| L\left|u\right\rangle}{\left\langle u|u\right\rangle}$$

Non-Equal Eigenvalues

For an operator L, we can show that two distinct eigenvalues λ_1, λ_2 correspond to two orthogonal eigenvectors. Start with

$$\left\langle u_2\right| L\left|u_1\right\rangle = \lambda_1\left\langle u_2|u_1\right\rangle \qquad\qquad \left\langle u_1\right| L\left|u_2\right\rangle = \lambda_2\left\langle u_1|u_2\right\rangle\ ,$$

and complex-conjugate the second term to eliminate the $\left\langle u_2|u_1\right\rangle$-term:

$$\left\langle u_2\right| L\left|u_1\right\rangle = \left(\frac{\lambda_1}{\lambda_2}\right)\overline{\left\langle u_1\right| L\left|u_2\right\rangle} = \left(\frac{\lambda_1}{\lambda_2}\right)\left\langle u_2\right| L\left|u_1\right\rangle$$

For $\lambda_1 \neq \lambda_2$, the only reasonable conclusion is

$$\left\langle u_2\right| L\left|u_1\right\rangle = 0 \qquad \rightarrow \qquad \left\langle u_1|u_2\right\rangle = 0\,,$$

meaning $\left|u_1\right\rangle$, $\left|u_2\right\rangle$ must be orthogonal.

Equal Eigenvalues

If two eigenvalues are equal, their corresponding eigenvectors are not necessarily orthogonal, in which case the vectors form a subspace of the original vector space that can admit an orthonormal basis.

Single Time Derivative Operator

Take a vector space $\mathcal{V}$ of dimension N admitting a fixed orthonormal basis $\{|e_j\rangle\}$ where $j = 1, \ldots, N$. A time-varying vector $|u(t)\rangle$ is a linear combination of time-varying coefficients such that

$$|u(t)\rangle = \sum_j u_j(t)\,|e_j\rangle \ .$$

Consider the single time-derivative operator $L = \partial_t$. If L acts on an arbitrary vector $|u(t)\rangle$, the result is

$$\partial_t\,|u(t)\rangle = |\dot{u}(t)\rangle = |\dot{u}\rangle = \sum_j \dot{u}_j(t)\,|e_j\rangle \ .$$

Of course, if L acts on an eigenvector $|x(t)\rangle$, the result is equal to $|x(t)\rangle$ multiplied by its eigenvalue λ. That is,

$$\sum_j \dot{x}_j\,|e_j\rangle = \sum_j \lambda x_j(t)\,|e_j\rangle \ ,$$

implying a separable differential equation

$$\partial_t x_j(t) = \lambda x_j(t)$$

for each index j. Elementary methods give the solution for each coefficient

$$x_j(t) = x(t=0)_j\, e^{\lambda t} = x_{0j}\, e^{\lambda t} \ ,$$

telling us

$$|x(t)\rangle = \sum_j x_{0j}\, e^{\lambda t}\,|e_j\rangle = e^{\lambda t}\sum_j x_{0j}\,|e_j\rangle = e^{\lambda t}\,|x_0\rangle \ ,$$

where the time dependence factors out of the sum. Perhaps not surprisingly, the eigenvectors evolve exponentially in time.

Double Time Derivative Operator

We also consider the double time-derivative operator $L = \partial_{tt}$. Using the same setup, it follows that each x_j is governed by the differential equation

$$\partial_{tt} x_j(t) = \lambda x_j(t) \ ,$$

whose solution is governed by λ. For $\lambda = 0$, the coefficients evolve linearly in time:

$$\lambda = 0 \qquad\qquad x_j(t) = x_{0j} + x_{1j}\, t$$

For $\lambda > 0$, we have a linear combination of exponential terms:

$$\lambda > 0 \qquad\qquad x_j\left(t\right) = x_{0j}\, e^{\sqrt{\lambda} t} + x_{1j}\, e^{-\sqrt{\lambda} t}$$

Spending a moment on the $\lambda < 0$ case, we have sines and cosines in general

$$\lambda < 0 \qquad\qquad x_j\left(t\right) = x_{0j}\, \cos\left(\sqrt{-\lambda} t + \phi_{0j}\right) + x_{1j}\, \sin\left(\sqrt{-\lambda} t + \phi_{1j}\right) \ ,$$

where the substitution

$$A_j = x_{0j} \cos\phi_{0j} + x_{1j} \sin\phi_{1j} \qquad\qquad B_j = -x_{0j} \sin\phi_{0j} + x_{1j} \cos\phi_{1j}$$

simplifies $x_j\left(t\right)$ to

$$x_j\left(t\right) = A_j \cos\left(\sqrt{-\lambda} t\right) + B_j \sin\left(\sqrt{-\lambda} t\right) \ .$$

The variables A_j and B_j can be re-expressed in polar coordinates as

$$A_j = R_j \cos\theta_j \qquad\qquad B_j = R_j \sin\theta_j \ ,$$

simplifying $x_j\left(t\right)$ down to

$$\lambda < 0 \qquad\qquad x_j\left(t\right) = R_j \cos\left(\sqrt{-\lambda} t - \theta_j\right) \ ,$$

having two unknown coefficients. In each case, of course the eigenvector is

$$\left|x\left(t\right)\right\rangle = \sum_j x_j\left(t\right) \left|e_j\right\rangle \ .$$

11 Matrix Operators

Consider a vector $\left|x\right\rangle$ living in vector space $\mathcal{V}$ that admits an orthonormal basis $\{\left|e_j\right\rangle\}$. As a linear combination of coefficients $\{x_j\}$, such a vector is written

$$\left|x\right\rangle = \sum_j x_j \left|e_j\right\rangle \ .$$

Another vector $\left|y\right\rangle$, which is itself a linear combination of coefficients $\{y_j\}$ in the same basis, may arise by applying a linear operator A onto $\left|x\right\rangle$:

$$\left|y\right\rangle = A\left|x\right\rangle = \sum_j y_j \left|e_j\right\rangle$$

Matrix Elements

Isolating any component y_j requires taking the inner product with a basis vector $|e_{k\neq j}\rangle$

$$\langle e_k| A |x\rangle = \sum_j \langle e_k| y_j |e_j\rangle = y_j \delta_{jk} = y_k \ ,$$

implying

$$y_j = \langle e_j| A |x\rangle = \langle e_j| A \sum_k x_k |e_k\rangle = \sum_k \langle e_j| A |e_k\rangle x_k \ .$$

That is, any y_j depends on each member of $\{x_j\}$ multiplied by a number

$$A_{jk} = \langle e_j| A |e_k\rangle$$

called a *matrix element*. The set of matrix elements $\{A_{jk}\}$ is the *matrix* represented by the operator A. To restore the operator A in terms of its elements, begin with the identity $A = IAI$ and write out each identity operator explicitly to get

$$A = IAI = \sum_j \sum_k |e_j\rangle \langle e_j| A |e_k\rangle \langle e_k| = \sum_j \sum_k |e_j\rangle A_{jk} \langle e_k| \ .$$

Matrix Addition

Two matrix operators A and B readily add to form a new operator C such that

$$A |x\rangle + B |x\rangle = C |x\rangle \qquad\qquad A_{jk} + B_{jk} = C_{jk} \ ,$$

which of course assumes that A and B are of equal dimension.

Scalar Multiplication

A scalar λ can be 'multiplied into' an operator A by scaling each component to create another operator B:

$$B = \lambda A \qquad\qquad B_{jk} = \lambda A_{jk}$$

Matrix Multiplication

Two operators A and B can 'multiply' to form a new operator C such that

$$A (B |x\rangle) = C |x\rangle \qquad\qquad AB = C \ ,$$

which is generally an associative operation, but not commutative:

$$(AB) C = A (BC) \qquad\qquad AB \neq BA$$

The formula for matrix multiplication is calculated by brute force

$$C = AB = \sum_j \sum_k \sum_{j'} \sum_m |e_j\rangle A_{jk} \langle e_k|e_{j'}\rangle B_{j'm} \langle e_m|$$

$$C = \sum_j \sum_m |e_j\rangle \left(\sum_k A_{jk} B_{km} \right) \langle e_m| \ ,$$

telling us

$$C_{jm} = \sum_k A_{jk}\, B_{km} \; .$$

The matrices A and B need not be perfectly equal in dimension - it's only required that the number of columns in A match the number of rows in B. For instance, the operation $A\,(2,4) \times B\,(4,3) = C\,(2,3)$, explicitly written as

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \\ b_{41} & b_{42} & b_{43} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix}$$

is perfectly valid, whereas the commutated product $B\,(4,3) \times A\,(2,4)$ is undefined.

12 Hermitian Matrix

A matrix operator A that is it's own adjoint operator, meaning $A = A^\dagger$ as appearing in the definition

$$\langle y|\, A\, |x\rangle = \overline{\langle x|\, A^\dagger\, |y\rangle} \qquad\qquad |x\rangle\,, |y\rangle\; \in \mathcal{V}$$

is a *Hermitian matrix*, where $A^\dagger$ is the *Hermitian conjugate*. In component form, we note that

$$\langle y|Ax\rangle = \sum_{ij} A_{ij}\, y_i^* x_j \; ,$$

where meanwhile, an equivalent statement is

$$\overline{\langle x|A^\dagger y\rangle} = \sum_{ij} \left(A_{ji}^\dagger\right)^* x_j y_i^* \; ,$$

telling us that the components of a Hermitian matrix obey

$$A_{ij}^* = A_{ji}^\dagger \; .$$

The notion of anti-Hermitian operators extends to matrices. An anti-Hermitian matrix satisfies

$$A^\dagger = -A \qquad\qquad A_{ij}^* = -A_{ij} \; .$$

As a corollary, we note that if a matrix A is Hermitian, then iA is anti-Hermitian, and vise-versa.

Commuting Operators

Now we derive the important fact that commuting Hermitian operators share an orthonormal basis. To begin, consider an operator A admitting an eigenvector $|\psi_n\rangle$ with corresponding eigenvalue a_n, and also a second operator B admitting an eigenvector $|\phi_n\rangle$ with corresponding eigenvalue b_n as

$$A\,|\psi_n\rangle = a_n\,|\psi_n\rangle \qquad\qquad B\,|\phi_n\rangle = b_n\,|\phi_n\rangle \; .$$

Our key assumption is that any $|\psi_n\rangle$ can be written as a linear combination of $\{|\phi_n\rangle\}$, and vice-versa, meaning each vector is a member of the same basis:

$$|\psi_n\rangle = \sum_m \gamma_{mn} |\phi_m\rangle \qquad\qquad |\phi_n\rangle = \sum_m \tilde{\gamma}_{mn} |\psi_m\rangle \ .$$

In the above, γ, $\tilde{\gamma}$ are matrix coefficients. We gain a restriction on γ, $\tilde{\gamma}$ by substitution

$$|\psi_n\rangle = \sum_m \gamma_{mn} \sum_{m'} \tilde{\gamma}_{m'm} |\phi_{m'}\rangle = \sum_{m'} \left(\sum_m \gamma_{mn}\tilde{\gamma}_{m'm} \right) |\phi_{m'}\rangle \ ,$$

giving a delta function relation

$$\delta_{m'n} = \sum_m \gamma_{mn}\tilde{\gamma}_{m'm} \ .$$

To proceed, compute $A|\psi_n\rangle$ two different ways to write

$$\begin{aligned} A|\psi_n\rangle &= A\sum_m \gamma_{mn} |\phi_m\rangle = A\sum_m \sum_{m'} \gamma_{mn}\tilde{\gamma}_{m'm} |\psi_{m'}\rangle \\ a_n |\psi_n\rangle &= \sum_{m'} \sum_m \gamma_{mn}\tilde{\gamma}_{m'm} a_{m'} |\psi_{m'}\rangle \ , \end{aligned}$$

evidently telling us, and similarly for $B|\phi_n\rangle$,

$$\frac{1}{a_n}\sum_m \gamma_{mn}\tilde{\gamma}_{m'm} a_{m'} = \delta_{m'n} \qquad\qquad \frac{1}{b_n}\sum_m \gamma_{mn}\tilde{\gamma}_{m'm} b_{m'} = \delta_{m'n} \ .$$

Now let the operator B act on $A|\psi_n\rangle$

$$\begin{aligned} BA|\psi_n\rangle &= B\sum_{m'} \sum_m \gamma_{mn}\tilde{\gamma}_{m'm} a_{m'} |\psi_{m'}\rangle = \sum_{m'} \sum_\alpha \sum_m \gamma_{mn}\tilde{\gamma}_{m'm} a_{m'} \gamma_{\alpha m'} b_\alpha |\phi_\alpha\rangle \\ &= \sum_{m'} \sum_\rho \sum_\alpha \tilde{\gamma}_{\rho\alpha}\gamma_{\alpha m'} b_\alpha \delta_{nm'} a_n |\psi_\rho\rangle = \sum_{m'} \sum_\rho \delta_{\rho m'}\delta_{nm'} a_n b_\rho |\psi_\rho\rangle \\ &= a_n b_n |\psi_n\rangle \ , \end{aligned}$$

which simplifies nicely. Similarly, we must let A act on $B|\psi_n\rangle$. Begin by calculating $B|\psi_n\rangle$ to get

$$B|\psi_n\rangle = B\sum_m \gamma_{mn} |\phi_m\rangle = \sum_m \gamma_{mn} b_m |\phi_m\rangle = \sum_m \sum_{m'} \gamma_{mn} b_m \tilde{\gamma}_{m'm} |\psi_{m'}\rangle \ .$$

Simplifying, we have

$$\begin{aligned} AB|\psi_n\rangle &= A\sum_m \sum_{m'} \gamma_{mn} b_m \tilde{\gamma}_{m'm} |\psi_{m'}\rangle = \sum_m \sum_{m'} \gamma_{mn} b_m \tilde{\gamma}_{m'm} a_{m'} |\psi_{m'}\rangle \\ &= \sum_{m'} \delta_{nm'} b_n a_{m'} |\psi_{m'}\rangle \\ &= a_n b_n |\psi_n\rangle \ . \end{aligned}$$

To finish, let us write the commutator of A and B to arrive at

$$[AB]|\psi_n\rangle = (AB - BA)|\psi_n\rangle = (a_n b_n - a_n b_n)|\psi_n\rangle = 0 \,,$$

which evaluates to *zero*. That is, we get a zero commutator of two operators whose eigenvectors share an orthonormal basis.

13 Matrix in Hilbert Subspace

Laplacian Operator in Hilbert Subspace

In the Hilbert space of functions $\mathcal{L}_2[-1,1]$, one can determine the components of the Laplacian operator $B = \partial_{xx}$ of a subspace spanned by an orthonormal basis. For instance, taking

$$
\begin{aligned}
|e_1\rangle &= \frac{1}{\sqrt{2}} \\
|e_2\rangle &= \frac{1}{\sqrt{2}}\left(\sin\left(\frac{\pi}{2}x\right) + \cos\left(\pi x\right)\right) \\
|e_3\rangle &= \frac{1}{\sqrt{2}}\left(\sin\left(\frac{\pi}{2}x\right) - \cos\left(\pi x\right)\right) ,
\end{aligned}
$$

the corresponding matrix B is calculated from

$$
B = \begin{bmatrix} \langle e_1| B |e_1\rangle & \langle e_1| B |e_2\rangle & \langle e_1| B |e_3\rangle \\ \langle e_2| B |e_1\rangle & \langle e_2| B |e_2\rangle & \langle e_2| B |e_3\rangle \\ \langle e_3| B |e_1\rangle & \langle e_3| B |e_2\rangle & \langle e_3| B |e_3\rangle \end{bmatrix} ,
$$

where

$$
\langle e_j| B |e_k\rangle = \int_{-1}^{1} e_j^*(x)\, \partial_{xx} e_k(x)\; dx .
$$

Carrying out each integral, find

$$
B = \frac{-\pi^2}{8}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 5 & -3 \\ 0 & -3 & 5 \end{bmatrix} ,
$$

which is a Hermitian matrix by inspection.

Two Operators in Hilbert Subspace

In the (same) Hilbert space of functions $\mathcal{L}_2[-1,1]$, one can determine the components of a derivative operator $A = \partial_x$ and a Laplacian operator $B = \partial_{xx}$ of a subspace spanned by an orthonormal basis.

Using an example orthonormal basis

$$
\begin{aligned}
|e_1\rangle &= \frac{1}{\sqrt{2}} \\
|e_2\rangle &= \frac{1}{\sqrt{2}}\sin\left(\pi x\right) \\
|e_3\rangle &= \frac{1}{\sqrt{2}}\cos\left(\pi x\right) ,
\end{aligned}
$$

we first check that the subset is closed under operations A:

$$A\left|e_1\right\rangle = \partial_x\left(\frac{1}{\sqrt{2}}\right) = 0$$
$$A\left|e_2\right\rangle = \partial_x\left(\frac{1}{\sqrt{2}}\sin(\pi x)\right) = \frac{\pi}{\sqrt{2}}\cos(\pi x) = \pi\left|e_3\right\rangle$$
$$A\left|e_3\right\rangle = \partial_x\left(\frac{1}{\sqrt{2}}\cos(\pi x)\right) = -\frac{\pi}{\sqrt{2}}\sin(\pi x) = -\pi\left|e_2\right\rangle$$

Each nontrivial result can be written in terms of the original basis vectors. Therefore $A\left|x\right\rangle$, where $\left|x\right\rangle$ is an arbitrary linear combination of the basis vectors, will result in a vector in the same subspace. Expressing A as a matrix requires calculating $A_{jk} = \left\langle e_j\right| A \left|e_k\right\rangle$, resulting in

$$A = \pi\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} ,$$

which is a *not* Hermitian matrix.

Repeating the same exercise using $B = \partial_{xx}$ results (of course) in a Hermitian matrix

$$B = -\pi^2\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} .$$

Interestingly, since the operation ∂_{xx} is two instances of ∂_x, it should follow that $B = AA$, easily checked by matrix multiplication:

$$B = \pi^2\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} = -\pi^2\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Pauli Matrices

Consider the set of three 2×2 Hermitian matrices

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \qquad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

called the *Pauli matrices*. Interestingly, each matrix ($k = 1, 2, 3$) has the property

$$\sigma_k\sigma_k = I ,$$

where I is the two-dimensional identity matrix. As a corollary, we know

$$\sigma_k^{2m} = I \qquad \sigma_k^{2m+1} = \sigma_k$$

for integer m. Furthermore, we have

$$\sigma_2\sigma_3 = i\sigma_1 \qquad \sigma_3\sigma_1 = i\sigma_2 \qquad \sigma_1\sigma_2 = i\sigma_3 .$$

Eigenvectors and eigenvalues of the Pauli matrices are governed by

$$\sigma_k \left|x\right\rangle = \lambda_k \left|x\right\rangle \ .$$

Working with σ_1 as an example, we write

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda_1 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} ,$$

giving two equations

$$x_2 = \lambda_1 x_1 \qquad\qquad x_1 = \lambda_1 x_2 ,$$

having two nontrivial branches $\lambda_1 = 1$ and $\lambda_1 = -1$, implying either $x_1 = x_2$, or respectively, $x_1 = -x_2$. We thus have one normalized eigenvector per eigenvalue:

$$\lambda_1 = +1 \qquad\qquad \left|x^{(+)}\right\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\lambda_1 = -1 \qquad\qquad \left|x^{(-)}\right\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Repeating the exercise on $\sigma_2 \left|y\right\rangle = \lambda_2 \left|y\right\rangle$, and a third time on $\sigma_3 \left|z\right\rangle = \lambda_3 \left|z\right\rangle$, we find the eigenvalues are always $\lambda_k = \pm 1$. The corresponding normalized eigenvectors turn out to be

$$\left|y^{(+)}\right\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} i \\ 1 \end{bmatrix} \qquad\qquad \left|y^{(-)}\right\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix}$$

$$\left|z^{(+)}\right\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad\qquad \left|z^{(-)}\right\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \end{bmatrix} .$$

14 Functions of Operators

Recall that the single time derivative operator $L = \partial_t$ as it appears in the problem $L \left|x\right\rangle = \lambda \left|x\right\rangle$ has exponentially-evolving eigenvectors $\left|x(t)\right\rangle = e^{\lambda t} \left|x_0\right\rangle$, where $\left|x_0\right\rangle$ is the initial state in a fixed orthonormal basis. The more general statement

$$\left|x(t)\right\rangle = e^{Lt} \left|x_0\right\rangle$$

leads to the same solution, easily shown by taking a time derivative:

$$\partial_t \left|x(t)\right\rangle = L\, e^{Lt} \left|x_0\right\rangle = L \left|x(t)\right\rangle$$

The combination e^{Lt} is generally known as a *time evolution* operator. We now have an operator as the argument of a function, which is tractable if the function is equivalent to matrix addition and/or multiplication.

The exponential function, despite having an operator in its argument, readily expands as

$$e^{Lt} = \sum_{k=0}^{\infty} \frac{1}{k!} (Lt)^k = I + (Lt) + \frac{1}{2!} (Lt)^2 + \frac{1}{3!} (Lt)^3 + \cdots .$$

Grouping even terms and odd terms separately, the above reads

$$e^{Lt} = \left(I + \frac{1}{2!}(Lt)^2 + \frac{1}{4!}(Lt)^4 + \cdots\right) + \left(Lt + \frac{1}{3!}(Lt)^3 + \frac{1}{5!}(Lt)^5 + \cdots\right) .$$

Making the substitution $L = -i\tilde{H}$, the above becomes

$$e^{-i\tilde{H}t} = \left(I - \frac{1}{2!}\left(\tilde{H}t\right)^2 + \frac{1}{4!}\left(\tilde{H}t\right)^4 - \cdots\right) - i\left(\tilde{H}t - \frac{1}{3!}\left(\tilde{H}t\right)^3 + \frac{1}{5!}\left(\tilde{H}t\right)^5 - \cdots\right) ,$$

simplifying nicely to

$$e^{-i\tilde{H}t} = \cos\tilde{H}t - i\sin\tilde{H}t ,$$

giving away two more functions where an operator is natural to show up.

Example: Pauli Matrices

In the special case that $\tilde{H}$ is equal to any of the Pauli matrices $\{\sigma_k\}$ with $k = 1, 2, 3$ up to a proportionality constant μ such that $\tilde{H} = \mu\sigma_k$, the above reduces to

$$e^{-i\sigma_k\mu t} = I\cos\mu t - i\sigma_k\sin\mu t ,$$

which removes the operator from any infinite series. Explicitly:

$$e^{-i\sigma_1\mu t} = \begin{bmatrix} \cos\mu t & -i\sin\mu t \\ -i\sin\mu t & \cos\mu t \end{bmatrix}$$
$$e^{-i\sigma_2\mu t} = \begin{bmatrix} \cos\mu t & -\sin\mu t \\ -\sin\mu t & \cos\mu t \end{bmatrix}$$
$$e^{-i\sigma_3\mu t} = \begin{bmatrix} e^{-i\mu t} & 0 \\ 0 & e^{i\mu t} \end{bmatrix}$$

15 Unitary Operators

Consider a vector space $\mathcal{V}$ admitting two different sets of basis vectors $\{|e_j\rangle\}$ and $\{|\tilde{e}_j\rangle\}$. In terms of coefficients x_j and $\tilde{x}_j$, a given vector $|x\rangle$ is a linear combination in each basis:

$$|x\rangle = \sum_j x_j |e_j\rangle = \sum_j \tilde{x}_j |\tilde{e}_j\rangle$$

Any coefficient(s) x_j, which exist in the vector space $\tilde{\mathcal{V}}$, can be isolated by exploiting the orthogonality between each $|e_j\rangle$:

$$x_j = \langle e_j|x\rangle = \sum_k \langle e_j| x_k |\tilde{e}_k\rangle = \sum_j \tilde{x}_k \langle e_j|\tilde{e}_k\rangle ,$$

which applies similarly to $\tilde{x}_j$:

$$\tilde{x}_j = \langle\tilde{e}_j|x\rangle = \sum_k \langle\tilde{e}_j| x_k |e_k\rangle = \sum_j x_k \langle\tilde{e}_j|e_k\rangle$$

In component notation, we have two matrices

$$U_{jk} = \langle \tilde{e}_j | e_k \rangle \qquad\qquad \tilde{U}_{jk} = \langle e_j | \tilde{e}_k \rangle \ ,$$

which can be directly interpreted as a set of coordinate transformation matrices that carry $\{x_j\} \in \tilde{\mathcal{V}} \rightarrow \{\tilde{x}_j\} \in \tilde{\mathcal{V}}$, and vice versa. These matrices are intricately related, which we first notice by the observation

$$\tilde{U}_{jk} = \langle e_j | \tilde{e}_k \rangle = \overline{\langle \tilde{e}_k | e_j \rangle} = U^*_{kj} = U^\dagger_{jk} \ ,$$

telling us each matrix is the other's Hermitian conjugate (dropping component notation):

$$\tilde{U} = U^\dagger \qquad\qquad U = \tilde{U}^\dagger$$

Now, the double transformation $\{x_j\} \rightarrow \{\tilde{x}_j\} \rightarrow \{x_j\}$ (and vice versa) tells us that each combination $\tilde{U}U$ and $U\tilde{U}$ is an identity matrix:

$$\tilde{U}U = I \qquad\qquad U\tilde{U} = I$$

Combining the two previous results, we find the identities:

$$U^\dagger U = I \qquad\qquad UU^\dagger = I$$

$$\tilde{U}\tilde{U}^\dagger = I \qquad\qquad \tilde{U}^\dagger \tilde{U} = I$$

Any operator satisfying the above equations is called *unitary*.

An important consequence of unitary operators arises when dealing with the inner product of two vectors $|x\rangle$, $|y\rangle \in \mathcal{V}$. Calculating the inner product of $U|x\rangle$ and $U|y\rangle$, we find

$$\langle Ux | Uy \rangle = \left\langle U^\dagger U x \middle| y \right\rangle = \langle Ix | y \rangle = \langle x | y \rangle \ ,$$

indicating that the inner product is unaffected by the operations. In most applications (namely in two- and three-dimensional space), unitary operations correspond to rotations and reflections of the coordinates $\{x_j\} \in \tilde{\mathcal{V}}$.

Next, we examine how operator U acts directly on the basis vectors $\{|e_j\rangle\} \in \mathcal{V}$. Supposing we set $|x\rangle = |e_j\rangle$ and $|y\rangle = |e_k\rangle$, we find

$$\langle Ue_j | Ue_k \rangle = \langle e_j | e_k \rangle = \delta_{jk} \ ,$$

telling us that the combinations $|Ue_j\rangle$ and $|Ue_k\rangle$ are members of a second orthogonal basis $\{|g_j\rangle\}$ that can be generated from the original $\{|e_j\rangle\}$ via

$$|g_j\rangle = U |e_j\rangle \qquad\qquad j = 1, 2, 3, \ldots, N \ .$$

As a matrix, recall that the operator U can be written as

$$U = \sum_{jk} |e_j\rangle \, U_{jk} \, \langle e_k| = \sum_{jk} |e_j\rangle \, \langle \tilde{e}_j | e_k \rangle \, \langle e_k| \ ,$$

where the completeness relation

$$I = \sum_k |e_k\rangle \langle e_k|$$

reduces the above to

$$U = \sum_j |e_j\rangle \langle \tilde{e}_j| \ .$$

Right away, we find

$$U |\tilde{e}_j\rangle = \sum_j |e_j\rangle \langle \tilde{e}_j|\tilde{e}_j\rangle = |e_j\rangle \ .$$

Similarly, we deduce

$$\tilde{U} = \sum_j |\tilde{e}_j\rangle \langle e_j| \ ,$$

which leads to

$$\tilde{U} |e_j\rangle = |\tilde{e}_j\rangle \ .$$

In component form, these results read:

$$|e_k\rangle = U |\tilde{e}_k\rangle = \sum_{ij} |\tilde{e}_i\rangle U_{ij} \langle \tilde{e}_j|\tilde{e}_k\rangle = \sum_j U_{jk} |\tilde{e}_j\rangle$$

$$|\tilde{e}_k\rangle = \tilde{U} |e_k\rangle = \sum_{ij} |e_i\rangle \tilde{U}_{ij} \langle e_j|e_k\rangle = \sum_j \tilde{U}_{jk} |e_j\rangle$$

As a sanity check, let us apply U to a vector $|x\rangle \in \mathcal{V}$. Calculating this, we have

$$U |x\rangle = \sum_{jk} x_k |e_j\rangle \langle \tilde{e}_j|e_k\rangle = \sum_j \left(\sum_k U_{jk} x_k \right) |e_j\rangle = |x'\rangle \ ,$$

where

$$x'_j = \sum_k U_{jk} x_k$$

is the rotated vector component in the same basis $\{|e_j\rangle\}$.

Rotations in Two Dimensions

The simplest nontrivial case involving unitary operators addresses rotations in the two-dimensional plane. Consider a Cartesian space spanned by the orthonormal basis $|e_1\rangle = \hat{x}$, $|e_2\rangle = \hat{y}$. A second orthonormal basis $\{|\tilde{e}_j\rangle\}$ is oriented at angle ϕ with respect with respect to the original. Using the above formula for U_{jk} applied to standard two-dimensional geometry, we quickly find the components of U to be

$$U_{xx} = \langle \tilde{e}_1|e_1\rangle = \cos\phi \qquad U_{xy} = \langle \tilde{e}_1|e_2\rangle = \cos\left(\frac{\pi}{2} - \phi\right) = \sin\phi$$

$$U_{yx} = \langle \tilde{e}_2|e_1\rangle = \cos\left(\frac{\pi}{2} + \phi\right) = -\sin\phi \qquad U_{yy} = \langle \tilde{e}_2|e_2\rangle = \cos\phi \ ,$$

and similarly for $U^\dagger = \tilde{U}$:

$$\tilde{U}_{xx} = \langle e_1 | \tilde{e}_1 \rangle = \cos\phi \qquad \tilde{U}_{xy} = \langle e_1 | \tilde{e}_2 \rangle = \cos\left(\frac{\pi}{2} + \phi\right) = -\sin\phi$$

$$\tilde{U}_{yx} = \langle e_2 | \tilde{e}_1 \rangle = \cos\left(\frac{\pi}{2} - \phi\right) = \sin\phi \qquad \tilde{U}_{yy} = \langle e_2 | \tilde{e}_2 \rangle = \cos\phi$$

That is, the the matrix U calculates the components of a rotated vector in a fixed basis, whereas the matrix $U^\dagger = \tilde{U}$ transforms the coordinates of a fixed vector when the basis is rotated.

Rotations in Three Dimensions

In three dimensions, we extend the two-dimensional case to write three matrices (presumably named after aviation terms)

$$\begin{aligned} R_z(\alpha) &= \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} = \text{`yaw'} \\ R_y(\beta) &= \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} = \text{`pitch'} \\ R_x(\gamma) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma \\ 0 & \sin\gamma & \cos\gamma \end{bmatrix} = \text{`roll'}\,, \end{aligned}$$

where a general rotation of a vector in three dimensions is the product

$$R = R_z(\alpha)\, R_y(\beta)\, R_x(\gamma)\ .$$

Due to commutativity rules, the order in which the matrices are applied *does* affect the result. Formally, the above matrices correspond to an *intrinsic* rotation, having *Tait–Bryan* angles α, β, γ.

Matrix Transformations